GAUGE MOMENTA AS CASIMIR FUNCTIONS OF NONHOLONOMIC SYSTEMS

LUIS GARCÍA-NARANJO & JAMES MONTALDI

ABSTRACT. We consider nonholonomic systems with symmetry possessing a certain type of first integrals that are linear in the velocities. We develop a systematic method for modifying the standard nonholonomic almost Poisson structure that describes the dynamics so that these integrals become Casimir functions after reduction. This explains a number of recent results on Hamiltonization of nonholonomic systems, and has consequences for the study of relative equilibria in such systems.

1. INTRODUCTION

The search for geometric structures that are invariant by the flow of nonholonomic mechanical systems has driven a good part of the research in nonholonomic systems from the time of Chaplygin to the present day. These range from first integrals and invariant measures, to symplectic or Poisson structures.

The difficulty of the problem arises because, in the nonholonomic case, the addition of the D’Alembert constraint reaction forces destroys the Hamiltonian nature of the mechanical equations of motion. The resulting equations preserve the energy of the system but only allow a formulation in terms of almost Poisson structures \([49, 43, 18, 35]\) which fail to satisfy the Jacobi identity and are not preserved by the flow.

A look at concrete nonholonomic examples (see e.g. the tables in \([14, 15]\) for a good overview), suggests that the presence of symmetries could lead to the existence of some geometric invariants in concrete problems, but precise results in this direction are missing.

In recent years, much effort has been devoted to the generalisation of Noether’s Theorem (which in general terms is a statement linking symmetries to first integrals) to the nonholonomic setting (see for example \([1, 37, 26, 29]\)). Of particular interest is the case in which the symmetries of the system arise as the lift of the action of a Lie group \(G\) on the configuration space. In analogy with the Hamiltonian counterpart of this situation, one would expect that if such first integrals exist, they are linear on the velocities and remain constants of motion in the presence of \(G\)-invariant force potentials. First integrals for nonholonomic systems having these properties are known in the literature as horizontal gauge momenta \([8, 27, 28]\). In this paper we do not consider other types of linear first integrals, so we will simply call them gauge momenta.

A separate line of research in the field of nonholonomic systems with symmetry has focused on the geometry of their reduction (see \([41, 9, 11, 19, 39]\) and others). As was already known to Chaplygin, it is possible that the reduced equations of motion allow a Hamiltonian formulation, possibly after a time reparametrisation. In this scenario one says that the system admits a Hamiltonization \([24, 31, 16, 3]\). The theorem of Chaplygin on the reducing multiplier \([21]\) is one of the fundamental results in this area but its direct applicability is limited to systems with a small number of degrees of freedom, possessing an invariant measure, and with a specific type of symmetry.

Date: September 21, 2017.
This research was made possible by a Newton Advanced Fellowship from the Royal Society.
Contributions. In this paper, we bring together ideas and results from the theory of gauge momenta, from singular nonholonomic reduction, and from Poisson geometry, to produce new results in the area of Hamiltonization. Our research was inspired by the results of Borisov, Mamaev and Kilin \cite{13, 14, 15} who, by explicitly finding a Poisson bracket, showed that numerous systems—the Chaplygin ball, a solid of revolution rolling without slipping on a plane or on a sphere, or a uniform sphere rolling on a surface of revolution—admit a Hamiltonization. All of these examples share a fundamental property: they possess $G$-invariant gauge momenta that are Casimir functions of the Poisson bracket on the reduced space.

Our main result shows that the above situation holds in considerable generality. Namely, we prove that a $G$-invariant nonholonomic system possessing $\ell$ independent $G$-invariant gauge momenta, admits a description in almost Poisson form with respect to a $G$-invariant bracket which upon reduction has the gauge momenta as Casimirs. This result gives a positive answer to a question that was originally raised in \cite{32} (see also Section 7 in \cite{33}). Moreover, our proof is constructive and we give explicit formulae for the bracket in terms of a choice of configuration coordinates $q^i$ and $G$-invariant momenta $\pi_\alpha$ (defined by an equivariant moving frame $\{X_\alpha\}$ for the constraint distribution).

In our approach, the crucial object used to obtain the almost Poisson bracket $\Pi^A_{\text{nh}}$ with the aforementioned properties, is a $G$-invariant 3-form $\Lambda$ defined on the configuration space. The construction of $\Lambda$ uses the kinetic energy metric and vector fields that generate the gauge momenta in a way that closely resembles the definition of the Cartan 3-form on a compact Lie group. We then consider the almost Poisson bracket $\Pi_{\text{nh}}$ introduced in \cite{49} and use $\Lambda$ to perform a gauge transformation (as in \cite{3, 47}).

The need to perform a gauge transformation of $\Pi_{\text{nh}}$ to guarantee that the gauge momenta are Casimir functions of the reduced bracket is related to a fundamental property of their generators: they are vector fields on configuration space which are tangent to the group orbits but they are usually not infinitesimal generators of the action. Instead, they are so-called gauge symmetries which are configuration-dependent linear combinations of the infinitesimal generators. Because of the variation of the coefficients, the corresponding gauge momenta in general fail to be Casimirs of the reduction of $\Pi_{\text{nh}}$. This scenario is not encountered in holonomic mechanics. See Section 5.4.2.

If the level sets of the gauge momenta on the reduced space are 2-dimensional, our result leads to a direct Hamiltonization of the problem in terms of a rank 2 Poisson structure (Corollary 5.8 in the text). Two examples where this arises are the motion of a convex solid of revolution that rolls on a plane or on a sphere, and the motion of a uniform sphere rolling on a convex surface of revolution. We stress that these rank 2 Poisson structures on the reduced space arise as the projection of the almost Poisson bracket $\Pi^A_{\text{nh}}$ defined on the unreduced space, and some dynamical properties may be deduced from this fact. In particular, certain components of the tensor $\Pi^A_{\text{nh}}$ are preserved by the flow of the unreduced equations (Corollary 5.8).

Another dynamical consequence may be drawn from our construction. If the characteristic distribution defined by the bracket induced by $\Pi^A_{\text{nh}}$ on the reduced space is integrable and the leaves coincide with the level sets of the resulting Casimirs, as is the case in all of the examples mentioned above, then the relative equilibria of the system are characterised as the critical points of the energy restricted to the level sets of the gauge momenta.

We mention finally that it would be interesting to extend the results of our paper to produce Casimir functions of the reduced bracket in the case where the gauge momenta are not $G$-invariant. Such Casimir functions would necessarily be nonlinear functions of the velocities of the system. This type of construction could lead to a geometric understanding of the Hamiltonization of other examples.
like the multi-dimensional versions of the Chaplygin sphere and the Veselova system [40, 31], or the nonholonomic hinge [10]. We will consider this question separately.

**Previous work.** The origin of the rank 2 Poisson brackets given in [14, 15] for the reduced dynamics of the nonholonomic systems mentioned above has been considered by a number of authors using different approaches. In [25] Fassò, Giacobbe and Sansonetto indicate that their existence is a consequence of the generic periodicity of the reduced dynamics. From their perspective, no insight about the dynamics of the unreduced system can be obtained from these structures. This contrasts with our discussion above and with the content of Corollary 5.8.

Ramos [45] studied these brackets from an algebraic perspective. He correctly indicates that there is no need to perform a rescaling to satisfy the Jacobi identity, and notices that they allow an extension to the singular strata of the reduced space (that correspond to certain relative equilibria of the system). However, no link is made with any brackets for the system on the unreduced space.

The treatment by Tsiganov [48] proceeds by doing the reduction in two steps and by proposing an ansatz for a bracket at the intermediate stage. Then the author applies ‘brute force calculations’ hoping to obtain brackets for the system with specific properties. We mention that there is certain correspondence between his ansatz and with equation (3.6) that gives the form of a bracket for the system that is obtained via a gauge transformation.

In [4] Balseiro studies the Jacobiator of almost Poisson brackets that are obtained by gauge transformations of the nonholonomic bracket $\Pi_{nh}$ introduced in [49]. An emphasis is given to the behaviour of this Jacobiator under reduction. A link between the constructions of [4] and gauge momenta is suggested in [5].

**Outline.** The structure of the paper is the following. In Section 2 we give a quick review of the structure of the equations of motion using quasi-velocities defined by moving frames. This allows us to introduce notation and obtain working expressions for the almost Poisson bracket $\Pi_{nh}$ defined by van der Schaft and Maschke [49], and known as the nonholonomic bracket. In Section 3 we show how, to a given a 3-form $\Lambda$ on the configuration space $Q$ one can associate an almost Poisson structure $\Pi_{nh}^{\Lambda}$ for the nonholonomic dynamics that possesses the same properties as $\Pi_{nh}$. Our discussion follows the ideas of [3, 47] and gives explicit formulae for $\Pi_{nh}^{\Lambda}$. In Section 4 we review some existing results about gauge momenta and reduction. For our purposes, the most convenient point of view for reduction is the one developed in [19] that applies in the case of non-free actions and follows a Poisson perspective.

Our principal contributions are contained in Section 5. The material in Section 5.1 establishes the preliminaries for the formulation of Lemma 5.4 that states the existence of the $G$-invariant 3-form $\Lambda$ that appears in the statement of Theorem 5.6. This theorem states that the gauge momenta of the system are Casimir functions of the reduction of $\Pi_{nh}^{\Lambda}$ and is our main result. Our proof is valid for free actions of $G$ on $Q$. For more general actions, we need to assume that the 3-form in Lemma 5.4 admits a smooth extension to the points in $Q$ having non-trivial isotropy (a situation that is encountered in all examples that we consider). The proof of Lemma 5.4 is somewhat technical and is presented in the Appendix. The special case in which the difference between the rank of $D$ and the number of gauge momenta is less than 3, is encountered in a great number of examples. In this case the 3-form in Lemma 5.4 is unique and a local formula for the corresponding bracket $\Pi_{nh}^{\Lambda}$ is presented in Section 5.4. This section also contains a discussion of the need to modify the nonholonomic bracket $\Pi_{nh}$ by $\Lambda$ and describes an open direction of research.

We treat two different examples in Sections 6 and 7. The first deals with the Chaplygin sphere and is meant to be an illustration of how the different elements in our construction come together. In Section 7 we consider the rolling of a body of revolution on the plane. This example is more involved
and shows how our method works in the case of a non-free action. Up to an unnecessary conformal factor, we recover the bracket for the reduced system presented in [14].

We would like to thank Francesco Fassò for comments on an early draft of this paper, and a referee for comments helping to improve the exposition.

Since completing this work, we have learned of recent work of Balseiro [6] on the same example we treat in Section 7 where she shows that the bracket found in [14] (also without the conformal factor) can be found using gauge momentum methods similar to ours, although her method for modifying the nonholonomic bracket is chosen for that specific example. There is also ongoing work of the same author [7] on the hamiltonization of a homogeneous ball rolling on a surface of revolution.

2. Preliminaries

2.1. The equations of motion — Lagrangian approach. Consider a nonholonomic system on an n-dimensional configuration manifold \( Q \). This consists of a Lagrangian \( L : TQ \to \mathbb{R} \) which we assume to be of mechanical type, kinetic minus potential energy:

\[
L(q, \dot{q}) = \frac{1}{2} \langle \dot{q}, \dot{q} \rangle - V(q),
\]

where the kinetic term is given by a Riemannian metric \( \langle \cdot, \cdot \rangle \) on \( Q \), and \( V(q) \) is the potential energy, and a regular non-integrable distribution \( D \subset TQ \) of rank \( r < n \) that determines the nonholonomic constraints. We assume that both the Lagrangian and the constraint distribution are time independent.

In accordance with the Lagrange-D’Alembert principle, the equations of motion are given in bundle coordinates \((q^i, \dot{q}^i)\) by

\[
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}^i} \right) - \frac{\partial L}{\partial q^i} = R_i, \quad i = 1, \ldots, n. \tag{2.1}
\]

Here \( R_i : D \to \mathbb{R} \), denote the components of the constraint reaction force. Such a reaction force is assumed to be ideal, namely

\[
R_i(q, \dot{q})\dot{q}^i = 0, \quad \text{whenever} \quad \dot{q} \in D_q, \tag{2.2}
\]

and is determined uniquely by the condition that the nonholonomic constraints are satisfied. For convenience we consider \( R \) as smoothly extended to \( TQ \) (such an approach is taken in [1] for example). Then the system (2.1) is defined on all of \( TQ \) and its flow leaves \( D \) invariant. The restriction of this system to \( D \) determines a well defined vector field \( Y_{nh} \) on \( D \) whose integral curves satisfy the equations of motion of our problem.

As indicated by many different authors (e.g. [24, 12, 34, 17] and references therein), the equations of motion are conveniently written by introducing moving frames and the associated quasi-velocities adapted to the problem. This approach goes back to Hamel and allows one to write the equations on \( D \) explicitly, without any reference to the constraint force \( R \), and will be very useful for our purposes. Following the treatment in [34], let \( \{X_1, \ldots, X_r\} \) be vector fields that form a local basis of sections of \( D \) and let \( \{X_{r+1}, \ldots, X_n\} \) be a local basis of sections of \( D^\perp \) where the orthogonal complement is taken with respect to the kinetic energy metric. Any tangent vector \( \dot{q} \in T_qQ \) can be written uniquely as

\[
\dot{q} = v^\alpha X_\alpha(q) + v^A X_A(q),
\]

for coefficients \( v^1, \ldots, v^n \), that are the aforementioned quasi-velocities. Here and throughout, we use the following convention on the indices:

- greek indices \( \alpha, \beta, \gamma, \ldots \) run from 1 to \( r \),
- latin indices \( A, B, C, \ldots \) run from \( r+1 \) to \( n \).

\(^1\)Unless the sum over the indices is explicitly indicated, in this paper we use the Einstein summation convention.
• latin indices \( i, j, k, \ldots \) run from 1 to \( n \).

We can now use \((q^i, v^\alpha, v^A)\) as coordinates for \( TQ \) and write \( L = L(q^i, v^\alpha, v^A) \). Note that \( D \subset TQ \) is specified by the condition \( v^A = 0 \) for all \( A = r + 1, \ldots, n \), so we can take \((q^i, v^\alpha)\) as coordinates for \( D \).

After a lengthy calculation using the chain rule, one can show that the equations (2.1) on \( TQ \) are equivalent to

\[
\frac{d}{dt} \left( \frac{\partial L}{\partial q^a} \right) - \rho^i_a \frac{\partial L}{\partial \dot{q}^i} = -C^k_i \nu^k \frac{\partial L}{\partial v^k}, \quad \alpha = 1, \ldots, r, \\
\frac{d}{dt} \left( \frac{\partial L}{\partial v^\alpha} \right) - \rho^i_A \frac{\partial L}{\partial \dot{v}^i} = -C^i_A \nu^i \frac{\partial L}{\partial v^i} + \rho^i_A R_i, \quad A = r + 1, \ldots, n. \tag{2.3}
\]

In the above equation, the \( q \)-dependent coefficients \( \rho^i_A \) and \( C^k_i \) are defined by the relations

\[
X_j = \rho^i_j \partial_{\dot{q}^i}, \quad \{X_i, X_j\} = C^k_{ij} X_k,
\]

where \( \{,\} \) is the Lie bracket of vector fields. In the derivation of (2.3) one has to make use of the relation

\[
\rho^i_A R_i = 0, \quad \alpha = 1, \ldots, r,
\]

which follows from (2.2).

The restriction of the system (2.1) on \( TQ \) to \( D \) can now be performed by substituting \( v^A = 0 \) for all \( A = r + 1, \ldots, n \) on the first set of equations in (2.3). Note that our assumption that the Lagrangian is of mechanical type, and that \( X_\alpha \) and \( X_A \) are orthogonal with respect to the kinetic energy metric, implies

\[
\frac{\partial L}{\partial v^\alpha} \bigg|_{v^\alpha = 0} = 0.
\]

If we write \( L_c = L |_{D} \) for the constrained (or restricted) Lagrangian, the desired system on \( D \) becomes

\[
\frac{d}{dt} \left( \frac{\partial L_c}{\partial q^a} \right) - \rho^i_a \frac{\partial L_c}{\partial \dot{q}^i} = -C^k_i \nu^k \frac{\partial L_c}{\partial v^k}, \quad \alpha = 1, \ldots, r. \tag{2.4}
\]

In coordinates the constrained Lagrangian is given by \( L_c(q^i, v^\alpha) = L(q^i, v^\alpha, 0) \) and satisfies

\[
L_c(q^i, v^\alpha) = \frac{1}{2} \Sigma^\alpha_{\beta \gamma} (q^i) v^\beta v^\gamma - V(q),
\]

where \( \Sigma^\alpha_{\beta \gamma} := \langle X_\alpha, X_\beta \rangle \).

Equations (2.4) appear in [34] and are complemented by the kinematic relations

\[
\dot{q}^i = \rho^i_\alpha \nu^\alpha, \quad i = 1, \ldots, n, \tag{2.5}
\]

that follow from the definition of the quasi-velocities. Note that the second set of equations in (2.3) does not give any information about the dynamics in \( D \). Instead it serves to determine the reaction force \( R \).

It is useful to note that, since the vector fields \( X_\alpha \) and \( X_A \) are orthogonal, the coefficients \( C^\gamma_{\alpha \beta} \) satisfy

\[
\Sigma^\gamma_{\alpha \beta} C^\delta_{\alpha \beta} = \langle [X_\alpha, X_\beta], X_\gamma \rangle. \quad \text{The coefficients}
\]

\[
C^\gamma_{\alpha \beta} := \langle [X_\alpha, X_\beta], X_\gamma \rangle = \Sigma^\gamma_{\alpha \beta} C^\delta_{\alpha \beta} \tag{2.6}
\]

will play an important role in our construction.
2.2. The equations of motion — Hamiltonian approach. The Hamiltonian approach is defined on the dual bundle $D^*$. For our treatment it is convenient to realise this abstract bundle as a vector subbundle of $T^*Q$, but, as is well known from basic linear algebra, there is no canonical way to do this. However, if $W$ is any fixed subbundle of $T^*Q$ complementary to $D$, that is, $T^*Q = D \oplus W$, then there is a natural identification of $D^*$ with the annihilator $W^*$ which is a subbundle of $T^*Q$. One possible choice for $W$ is of course $D^\perp$, but in examples other choices may simplify coordinate calculations. In what follows we assume that a choice of $W$ has been made. The sections of $D^*$ can therefore be interpreted as 1-forms on $Q$ that annihilate $W$, and we can operate with them as one normally does.

Let $\{\mu^\alpha\}$ be a local basis of sections of $D^*$ that is dual to $\{X_\alpha\}$. Namely, $\mu^\alpha$ are locally defined 1-forms on $Q$ that vanish along $W$ and satisfy $\mu^\alpha(X_\beta) = \delta^\alpha_\beta$ (Kronecker delta). Any element of $D^*$ on the fibre $D^*_q$ over $q \in Q$ can be written uniquely as a linear combination $\pi_\alpha \mu^\alpha(q)$ for certain scalars $\pi_1, \ldots, \pi_n$. In this way one can use $(q^i, \pi_\alpha)$ as coordinates on $D^*$.

The restriction of the kinetic energy to $D$ is non-degenerate and allows one to define a constrained Legendre transform $\text{Leg}_c$ that is a vector bundle isomorphism between $D$ and $D^*$. It is the restriction to $D$ of the standard Lagrange transform in mechanics. Explicitly we have

$$\text{Leg}_c : D \to D^*, \quad (q^i, v^\alpha) \mapsto \left(q^i, \pi_\alpha = \frac{\partial L_c}{\partial \dot{q}^i} = G^\alpha_\beta \dot{v}^\beta \right).$$

The constrained Hamiltonian $H_c : D^* \to \mathbb{R}$ is the energy of the system and is defined by

$$H_c(q^i, \pi_\alpha) = \pi_\beta \dot{v}^\beta - L_c(q^i, v^\alpha),$$

where it is understood that $v^\beta$ is written in terms of $\pi_\alpha$ via the inverse Legendre transform $\text{Leg}_c^{-1}$. Explicitly we have

$$H_c(q^i, \pi_\alpha) = \frac{1}{2} G^\alpha_\beta \pi_\alpha \pi_\beta + V(q), \quad \text{(2.7)}$$

where $G^\alpha_\beta$ denotes the inverse matrix of $G_{\alpha\beta}$; namely, $G^\alpha_\beta G_{\gamma\beta} = \delta^\alpha_\gamma$. It is a straightforward calculation to show that

$$v^\alpha = \frac{\partial H_c}{\partial \pi_\alpha}, \quad \dot{q}^i = \rho^i_\alpha \frac{\partial H_c}{\partial \pi_\alpha}, \quad \pi_\alpha = -\rho^i_\alpha \frac{\partial H_c}{\partial \dot{q}^i} - C^i_\alpha_\beta \pi_\beta \frac{\partial H_c}{\partial \pi_\beta}, \quad \text{(2.8)}$$

and hence, the equations (2.4), (2.5), are equivalent to the following first order system on $D^*$:

$$\dot{q}^i = \rho^i_\alpha \frac{\partial H_c}{\partial \pi_\alpha}, \quad \pi_\alpha = -\rho^i_\alpha \frac{\partial H_c}{\partial \dot{q}^i} - C^i_\alpha_\beta \pi_\beta \frac{\partial H_c}{\partial \pi_\beta}. \quad \text{(2.8)}$$

We denote the corresponding vector field by $X_{\alpha nh}$, which is nothing other than the push-forward of $Y_{\alpha}$ by $\text{Leg}_c$. It is readily seen that the above equations can be written as

$$\{q^i, H_c\}_{nh} = 0, \quad \{q^i, \pi_\alpha\}_{nh} = \rho^i_\alpha, \quad \{\pi_\alpha, \pi_\beta\}_{nh} = -C^i_\alpha_\beta \pi_\gamma \frac{\partial H_c}{\partial \pi_\gamma}.$$

This bracket coincides with the one that was first introduced by van der Schaft and Maschke in [49] where it is shown that the Jacobi identity is satisfied if and only if the distribution $D$ is integrable. For a non-integrable $D$ one speaks of an almost Poisson bracket or a pseudo-Poisson bracket. Several intrinsic constructions and geometric interpretations of the bracket $\{\cdot, \cdot\}_{nh}$ are available in the literature, see e.g. [43, 18, 35, 42, 19] and the references therein. The presentation given above follows roughly the approach and notation of [42].
Denote by $\Pi_{nh}$ the bivector on $D^*$ defined by the bracket $\{\cdot, \cdot\}_{nh}$. Its expression in coordinates is

$$\Pi_{nh} = \rho^j_\alpha \partial_d \eta_j / \partial \pi^\alpha - \frac{1}{2} C_{\alpha \beta}^\gamma \pi^\alpha \partial_{\pi^\beta} \wedge \partial_{\pi^\gamma}. \quad (2.9)$$

The Hamiltonian\(^2\) vector field $X_f$ associated to $f \in C^\infty(D^*)$ is defined by $X_f := \Pi_{nh}^\sharp(df)$. Then clearly $X_{nh} = X_H$. Recall the standard notation: if $\Pi$ is a bivector field on $M$ then $\Pi^\sharp : T^*M \to TM$ is the associated bundle map; similarly if $\Xi$ is a 2-form on $M$ then $\Xi^\star$ is the bundle map $\Xi^\star : TM \to T^*M$.

2.3. Second order vector fields. The vector field $Y_{nh}$ on $D$ defined by the Lagrangian equations of motion (2.4) and (2.5) is second order with respect to the vector bundle structure of $D$ over the configuration manifold $Q$. That is, $Y_{nh}$ satisfies

$$T\tau_D \circ Y_{nh} = \text{id}_D,$$

where $\tau_D : D \to Q$ is the bundle projection. This condition naturally extends the standard definition of second order vector fields on $TQ$, see e.g. [44]. In local coordinates we have $T\tau_D(\dot{q}, \dot{v}) = \dot{q}$, so the condition of being second order is written as

$$\dot{q}(t) = v^\alpha(t) X_{\alpha}(q(t)) \in D_{\dot{q}(t)},$$

for any integral curve $(q(t), v^\alpha(t))$ of $Y_{nh}$.

The identification of $D$ with $D^*$ via the constrained Legendre transform suggests the following.

**Definition 2.1.** A vector field on $Z$ on $D^*$ is second order if its pull-back to $D$ by $\text{Leg}_q$ is second order.

In coordinates this means that a second order vector field $Z$ on $D^*$ satisfies

$$T\tau_{D^*}(Z(q, \pi)) = \sum^\alpha \beta \pi^\alpha X_{\alpha}(q)$$

where $\tau_{D^*} : D^* \to Q$ is the projection. The vector field $X_{nh}$ on $D^*$ defined by the equations (2.8) is the push forward of $Y_{nh}$ by the constrained Legendre transform and it is clearly second order.

Let $\beta^A$ be linearly independent 1-forms on $Q$ that span the annihilator $D^*$ and let $\mathcal{E}$ be the distribution on $D^*$ defined by the joint annihilator of $\{\tau_{D^*}^\sharp \beta^A\}$ on $D^*$. It is clear that $\mathcal{E}$ is well defined (independent of the choice of basis of $D^*$) and is a regular distribution of rank $2r$ on $D^*$. (A regular distribution is one of constant rank.)

The importance of the distribution $\mathcal{E}$ in the geometric formulation of nonholonomic mechanics seems to have first been noticed in [50] where it is shown that its fibres are symplectic (with respect to the canonical symplectic form on the ambient space $T^*Q$). $\mathcal{E}$ is a crucial ingredient in the formulations in e.g. [9], [19], [35], [39]. In coordinates we have

$$\mathcal{E}_{(q, \pi)} = \text{span} \{ \rho^j_\alpha \partial_d \eta_j, \partial_{\pi^\alpha} \}. \quad (2.10)$$

We collect some properties of $\mathcal{E}$ in the following proposition whose proof can be given using the above expression.

**Proposition 2.2.** The distribution $\mathcal{E}$ has the following properties:

(i) it is integrable if and only if $D$ is integrable;

(ii) a vector field $Z$ on $D^*$ is second order if and only if it is a section of $\mathcal{E}$;

(iii) it coincides with the characteristic distribution of $\Pi_{nh}$: that is, $\Pi_{nh}^\sharp(T^*D^*) = \mathcal{E}$.

As a consequence, all of the Hamiltonian vector fields on $D^*$ generated by $\Pi_{nh}$ are second order.

\(^2\)strictly speaking, $X_f$ is not a Hamiltonian vector field, since $\Pi_{nh}$ is not a Poisson bracket; it is only an ‘almost’ Hamiltonian vector field. We will however call these Hamiltonian vector fields throughout.
3. Gauge transformations associated to 3-forms

A particular procedure to construct a family of almost Poisson structures that describe the dynamics of a nonholonomic system was presented in [33]. Later, this method was put into a solid geometric context in [3] by relating it to the gauge transformations of Poisson brackets by 2-forms as introduced by Ševera and Weinstein [47]. Here we modify this construction by starting from a 3-form \( \Lambda \) on \( Q \). In Section 5 we apply this construction using specific 3-forms defined by the nonholonomic geometry in the presence of symmetry.

Let \( \Lambda \) be a section of \( \wedge^3(D^*) \). Using our identification of \( D^* \) with the annihilator \( W^\circ \subset T^*Q \) of a complement \( W \) of \( D \) on \( TQ \), we can interpret \( \Lambda \) as a 3-form on \( Q \) that vanishes upon contraction with tangent vectors to \( W \). Fix a basis of local basis of sections \( \{X_\alpha\} \) of \( D \). The local expression for the 3-form \( \Lambda \) is

\[
\Lambda = \frac{1}{6} B_{\alpha \beta \gamma} \mu^\alpha \wedge \mu^\beta \wedge \mu^\gamma,
\]

where, as before, \( \{\mu^\beta\} \) are locally defined 1-forms on \( Q \) annihilating \( W \), such that \( \mu^\beta(X_\alpha) = \delta^\beta_\alpha \). Here, the coefficients \( B_{\alpha \beta \gamma} \) are alternating (that is, skew-symmetric with respect to transpositions of \( \alpha, \beta, \gamma \)) and are given by

\[
B_{\alpha \beta \gamma} := \Lambda(X_\alpha, X_\beta, X_\gamma).
\]

We define the 2-form \( \Xi \) on \( D^* \) as the contraction of the pull-back \( \tau_\rho^* \Lambda \) of \( \Lambda \) to \( D^* \) with any second order vector field \( Z \) on \( D^* \). The 2-form \( \Xi \) is independent of the choice of \( Z \), it is semi-basic, and has local expression

\[
\Xi = \frac{1}{2} B_{\alpha \beta \gamma} \pi_\tau \mu^\alpha \wedge \mu^\beta,
\]

where \( B_{\alpha \beta \gamma} := \Xi^\gamma_{\alpha \beta} \).

**Lemma 3.1.** Given any 3-form \( \Lambda \) as above and writing \( \Xi \) for the resulting 2-form, the map

\[
\text{Id}_{TD^*} + \Pi^\gamma_{nh} \circ \Xi^\gamma
\]

is an invertible endomorphism of \( TD^* \).

**Proof.** We work in local coordinates with the notation introduced in Section 2. We have

\[
\mu^\alpha = \bar{\rho}^\alpha_i \, dq^i, \quad \alpha = 1, \ldots, r,
\]

where the duality between \( \{\mu^\alpha\} \) and \( \{X_\beta\} \) implies \( \bar{\rho}^\alpha_i \rho_j^\beta = \delta^\alpha_\beta \). We shall denote by \( \rho \) the \( n \times r \) matrix with entries \( \rho_i^\alpha \), and by \( \bar{\rho} \) the \( r \times n \) matrix with entries \( \bar{\rho}^\alpha_i \). The duality condition becomes \( \bar{\rho} \rho = \text{Id}_{r \times r} \).

The matrix representations of \( \Xi^\gamma \) and \( \Pi^\gamma_{nh} \) with respect to the respective bases \( \{\partial_q^i, \partial_{\pi_\alpha}\} \) and \( \{dq^i, d\pi_\alpha\} \) of \( T_{(q,\pi)}D^* \) and \( T_{(q,\pi)}D^* \) are given in block form by

\[
\Xi^\gamma = \begin{pmatrix} \bar{\rho}^T B \bar{\rho} & 0 \\ 0 & 0 \end{pmatrix}, \quad \Pi^\gamma_{nh} = \begin{pmatrix} 0 & \rho \\ -\rho^T & -\bar{\rho} \end{pmatrix},
\]

where \( C \) and \( B \) are \( r \times r \) skew-symmetric matrices with entries

\[
C_{\alpha \beta} = C^\gamma_{\alpha \beta} \pi_\tau, \quad B_{\alpha \beta} = B^\gamma_{\alpha \beta} \pi_\tau.
\]

Performing the matrix algebra, one finds

\[
\text{Id}_{TD^*} + \Pi^\gamma_{nh} \circ \Xi^\gamma = \begin{pmatrix} I & 0 \\ -B \bar{\rho} & I \end{pmatrix}
\]

which is clearly invertible. \( \square \)
Following [3] and [47] we define the bivector $\Pi^\Lambda_{nh}$ by

$$(\Pi^\Lambda_{nh})^\sharp := (\text{Id}_{TD^*} + \Pi^\sharp_{nh} \circ \Xi^\flat)^{-1} \circ \Pi^\sharp_{nh}. \tag{3.4}$$

This exists by virtue of the lemma above. We say that the bracket defined by $\Pi^\Lambda_{nh}$ is obtained by a 
\textit{gauge transformation} of the nonholonomic bracket $\{\cdot, \cdot\}_{nh}$ by the 3-form $\Lambda$. (Note that in [3] these would be denoted $\Pi^\sharp_{nh} \circ \Xi^\flat$, but in the present context the fundamental object is the 3-form $\Lambda$ rather than its contraction $\Xi$.)

\textbf{Theorem 3.2.} The bivector field $\Pi^\Lambda_{nh}$ has the following properties:

(i) its characteristic distribution is $\mathcal{E}$, that is $(\Pi^\Lambda_{nh})^\sharp(TD^*) = \mathcal{E}$,

(ii) it describes the nonholonomic dynamics; namely

$$(\Pi^\Lambda_{nh})^\sharp(dH_c) = X_{nh}, \tag{3.5}$$

(iii) it is given in local coordinates by

$$\Pi^\Lambda_{nh} = \rho^i_a \partial_{q^i} \wedge \partial_{\pi^a} + \frac{1}{2} \left( B^\gamma_{\alpha \beta} - C^\gamma_{\alpha \beta} \right) \pi^\gamma \partial_{\pi^\alpha} \wedge \partial_{\pi^\beta}. \tag{3.6}$$

\textbf{Proof.} (i): this follows directly from Proposition 2.2 and equation (3.4) that defines $\Pi^\Lambda_{nh}$.

(ii): since the 2-form $\Xi$ was defined as the contraction of the 3-form $\tau^\flat \circ \Lambda$ with any second order vector field $Z$ on $D^*$, and $X_{nh}$ is a second order vector field on $D^*$, we can write

$$\Xi(\cdot, \cdot) = \tau^\flat \circ \Lambda(X_{nh}, \cdot, \cdot),$$

and therefore

$$\Xi^\flat(X_{nh}) = 0.$$

It follows that $(\text{Id}_{TD^*} + \Pi^\sharp_{nh} \circ \Xi^\flat)(X_{nh}) = X_{nh}$, and hence (3.5) holds since $\Pi^\sharp_{nh}(dH_c) = X_{nh}$.

(iii): Finally, in the notation of the proof of Lemma 3.1, and in view of (3.3) we have

$$(\text{Id}_{TD^*} + \Pi^\sharp_{nh} \circ \Xi^\flat)^{-1} = \begin{pmatrix} I & 0 \\ \mathcal{B} & 0 \end{pmatrix},$$

which, combined with (3.2), gives

$$(\Pi^\Lambda_{nh})^\sharp = \begin{pmatrix} I & 0 \\ \mathcal{B} & I \end{pmatrix} \begin{pmatrix} 0 & \rho^\flat \\ -\rho^\flat & -\mathcal{C} \end{pmatrix} = \begin{pmatrix} 0 & \rho^\flat \\ -\rho^\flat & -\mathcal{C} \end{pmatrix},$$

and this is equivalent to (3.6). \hfill \Box

The above theorem shows that the equations of motion (2.8) can be formulated in Hamiltonian form with respect to the bivector $\Pi^\Lambda_{nh}$, for any 3-form $\Lambda$. The Jacobi identity fails for $\Pi^\Lambda_{nh}$ since the characteristic distribution $\mathcal{E}$ is non-integrable. Note that all of the Hamiltonian vector fields associated to $\Pi^\Lambda_{nh}$ are tangent to $\mathcal{E}$ and therefore are second order vector fields on $D^*$.

\textbf{Remark 3.3.} At many points in this section, we liberally refer to $\Lambda$ as a 3-form, and we will continue to do so throughout the paper. Strictly speaking $\Lambda$ is a section of $\Lambda^3(D^*)$ and it is only in virtue of the identification $D^* = W^\circ$, that we may interpret it as a 3-form on $Q$ (that annihilates $W$). This issue is taken into consideration in the construction in Lemma 5.4 of the specific $\Lambda$ that is used to prove our main result.
4. Gauge momenta and symmetry reduction

The existence of first integrals that are linear in velocity, or momentum variables, for nonholonomic systems has received a great deal of attention (see e.g. [29] and the references therein). In this section we outline some of the known results in the field and prove Theorem 4.3 which will be used in our construction in Section 5 below. Similar versions of this theorem are available in the literature (see the discussion in [29]).

4.1. Linear first integrals of nonholonomic systems. Consider first the Hamiltonian formulation in terms of the equations of motion (2.8) on $D^\ast$. Linear functions on $D^\ast$ are sections of $D^{\ast\ast} = D$ so we can naturally make a one to one correspondence between vector fields that take values on $D$ and linear functions on the phase space $D^\ast$. This observation goes back to Iliev [36, 37] and to some extent Agostinelli [1]. Let $Z$ be a vector field on $Q$ taking values on $D$. It can be written as a linear combination of the basis of sections of $D$ as

$$Z = Z^\alpha(q)X_\alpha(q),$$

for certain functions $Z^\alpha \in C^\infty(Q)$. We denote by $p_Z \in C^\infty(D^\ast)$ the linear function associated to $Z$. In terms of the coordinates $(q^i, \pi_\alpha)$ we have

$$p_Z(q, \pi) = Z^\alpha(q)\pi_\alpha.$$

In particular note that $p_{X_\alpha} = \pi_\alpha$ for all $\alpha = 1, \ldots, k$.

Now consider the Lagrangian formulation given by equations (2.4), (2.5) defined on $D$. This approach is followed by many recent references [26, 27]. With a slight abuse of notation we denote the function $p_Z \circ \text{Leg}_c \in C^\infty(D)$ also by $p_Z$. It has the local expression

$$p_Z(q, v) = Z^\alpha(q)\frac{\partial L_c}{\partial q^\alpha} = Z^\alpha(q)\Sigma_{\alpha\beta}(q)v^\beta.$$

**Definition 4.1.** The vector field $Z$ taking values on $D$ is called the *generator* of the linear function $p_Z$.

We stress that the generator of a linear function is uniquely determined by the condition that it is a section of $D$. Also, linear functions on $D$ (or $D^\ast$) are independent over the set of points where the generating vector fields are linearly independent.

As it is shown in [37], the evolution of $p_Z$ along the nonholonomic Lagrangian system (2.4), (2.5) on $D$ is given by

$$p_Z = Z^{TQ}[L]|_D$$

(4.1)

where $Z^{TQ}$ denotes the *tangent lift* of the generator vector field $Z$ (the expression for $Z^{TQ}$ in bundle coordinates $(q^i, \dot{q}^i)$ is given in (4.2) below).

We now present an alternative characterisation of the condition that $p_Z$ is a first integral. Recall that a vector field $Z$ on $Q$ acts by infinitesimal isometries if the Lie derivative $\mathcal{L}_Z \Sigma$ of the metric along $Z$ vanishes. This suggests the following definition.

**Definition 4.2.** A vector field $Z$ on $Q$ acts by *infinitesimal isometries on $D$* if $(\mathcal{L}_Z \Sigma)(u, u) = 0$ for all $u \in D$.

Note that this does not assume that the flow defined by $Z$ preserves the distribution $D$. The following is a result from [26], where they allow $Z$ to be a section of the *reaction annihilator distribution*, which contains $D$.

**Theorem 4.3.** Consider a nonholonomic Lagrangian system with constraint distribution $D$ and Lagrangian $L(q, u) = \Sigma(u, u) - V(q)$. Let $Z$ be a section of the distribution $D$. Then the momentum
$p_Z$ generated by $Z$ is a first integral of the Lagrangian system if and only if $Z$ acts by infinitesimal isometries on $D$ and annihilates the potential energy $V$.

**Proof.** We have $\dot{p}_Z = Z^T Q[L] \big|_D$, so $p_Z$ is conserved if and only if the right hand side vanishes. We show this is equivalent to annihilating the potential and kinetic energies separately, and then show the latter is equivalent to being an infinitesimal isometry on $D$ (which follows from the lemma below). One implication is clear a fortiori. For the converse, if $Z^T Q$ annihilates $L$ on $D$, then restricting to the zero section shows that $Z[V] = 0$. Since $V$ is a function on $Q$, it follows that $Z^T Q$ annihilates its pull-back to $TQ$ (and hence to $D$). Since $L$ and $V$ are both annihilated by $Z^T Q$ (on $D$), it follows that $Z^T Q$ also annihilates the metric (kinetic energy) restricted to $D$. That this is equivalent to $Z$ being an isometry on $D$ follows from the following lemma. □

**Lemma 4.4.** Let $Z$ be any vector field on $Q$, and $G$ the metric tensor. Then as functions on $TQ$,

$$(\xi_Z G)(u,u) = Z^T Q[G(u,u)].$$

**Proof.** In local bundle coordinates, the tangent lift of $Z$ is

$$Z^T Q = Z^l \partial_{q^l} + \frac{\partial Z^k}{\partial q^j} \dot{q}^j \partial_{\dot{q}^k}. \quad (4.2)$$

Applying this to $G(u,u) = G_{ij} u^i u^j$ gives

$$Z^T Q[G(u,u)] = Z^k \frac{\partial G_{ij}}{\partial q^k} u^i u^j + 2Z^k \frac{\partial Z^k}{\partial q^j} u^j.$$  

But this is exactly the expression for $(\xi_Z G)(u,u)$ — see for example [38, p.55]. □

### 4.2. Gauge momenta.

An important class of linear first integrals that may exist in the presence of a symmetry group is that of gauge momenta. This terminology was first used in [8] where the authors indicate the existence of this kind of integrals in some classical examples of nonholonomic systems. Further research on their properties can be found in [27], [28], where they are called horizontal gauge momenta.

Let $G$ be a Lie group that acts properly on $Q$, and suppose that the lift of $G$ to $TQ$ leaves the Lagrangian $L$ and the constraint distribution $D$ invariant. It follows that $G$ acts by isometries on $Q$ (with respect to the kinetic energy metric), and that the potential $V \in C^\infty(Q)$ is $G$-invariant. Define by $\mathcal{S}$ the, possibly non-regular, distribution on $Q$ defined by

$$\mathcal{S}_q := D_q \cap g \cdot q.$$  

Here $g$ is the group’s Lie algebra and $g \cdot q$ denotes the tangent space to the group orbit at $q$. We assume that this distribution is regular on $Q_f$, and only changes rank possibly at points where the action of $G$ fails to be free.

**Definition 4.5.** A linear first integral of a nonholonomic system is called a **gauge momentum** associated to the $G$-action if its unique generator on $D$ is a section of $\mathcal{S}$.

We will be especially interested in **$G$-invariant gauge momenta**. In addition to being a section of $\mathcal{S}$, the generator of such integrals is $G$-equivariant.
4.3. Reduction. We present a basic outline of the almost Poisson reduction of nonholonomic systems with possibly non-free actions. We continue to work under the assumptions that were introduced above. Namely, there is an action of the Lie group $G$ on the configuration space $Q$ whose lift to $TQ$ preserves the Lagrangian $L$ and the constraint distribution $D$.

Via the constrained Legendre transform, the restricted action on $D$ defines a $G$-action on $D^*$ that leaves the constrained Hamiltonian $H_c$ (2.7) and the bivector $\Pi_{nh}$ (2.9) invariant. The reduced Hamiltonian is the unique function, which we also denote $H_c$, on the orbit space $D^*/G$ whose pull-back to $D^*$ is $H_c$. If the $G$-action on $Q$ is free then this orbit space can be expressed as a bundle of rank $r$ over the ‘shape space’ $Q/G$ (where $r$ is the rank of $D$). In general, if the action is not free, the quotient $D^*/G$ is a stratified space.

As is usual, one identifies smooth functions on $D^*/G$ with smooth $G$-invariant functions on $D^*$. Since the nonholonomic (almost) Poisson structure is invariant, it follows that $\{f,g\}_{nh}$ is invariant whenever $f$ and $g$ are, and hence $\{\cdot,\cdot\}_{nh}$ descends to an almost Poisson structure on $D^*/G$, which we continue to denote $\{\cdot,\cdot\}_{nh}$. One can, by the usual formula, define Hamiltonian ‘vector fields’ on $D^*/G$ from this bracket, and it turns out that these are genuine vector fields on, or tangent to, each stratum. In particular, the Hamiltonian vector field defined by the reduced Hamiltonian $H_c$ is the reduced Hamiltonian vector field. The flow of this reduced Hamiltonian vector field is the projection of the flow of the original Hamiltonian vector field on $D^*$ given by equations (2.8).

Further details about the reduction procedure when the action is not free can be found in the book of Cushman, Duistermaat and Śniatycki [19].

The main point of our paper is that given any $G$-invariant 3-form $\Lambda$ on $Q$ as constructed in Section 3, the reduction outlined above follows mutatis mutandis for the bracket on $D^*$ defined by the bivector $\Pi_{nh}^\Lambda$. In the next section we see how to choose $\Lambda$ so that the invariant gauge momenta are Casimir functions of the reduced system.

5. Invariant gauge momenta are Casimirs of a reduced bracket

In this section we continue to work under the assumption that there is a Lie group $G$ acting on $Q$ and preserving both the mechanical Lagrangian $L$ and the constraint distribution $D$. Moreover, we shall assume that the $G$-action on $Q$ is proper and that the isotropy of a generic point $q \in Q$ is trivial. We denote by $Q_f$ the set of points in $Q$ having trivial isotropy. Then $Q_f$ is a $G$-invariant open dense subset of $Q$ and the restriction of the $G$ action to $Q_f$ is free and proper.

We suppose that, on $Q_f$ there exist $\ell$ linearly independent gauge momenta with $G$-equivariant generators $Z_b$, for $b = 1, \ldots, \ell$. We note that linear combinations of these vector fields with constant coefficients are also generators of gauge momenta.

5.1. Adapted bases and preliminary results.

**Definition 5.1.** A basis $\{X_\alpha\}$ of sections of $D$ defined on an open subset $U$ of $Q_f$ is said to be symmetry-momentum adapted (or simply adapted) if it can be written as $\{X_\alpha\} = \{Z_b, Y_l\}$ where:

- (i) the vector fields $\{Z_b\}$ generate gauge momenta,
- (ii) the vector fields $Z_b$ and $Y_l$ are $G$-equivariant.

Here we refine the convention on the indices introduced in Section 2.1 where $\alpha, \beta, \ldots$ run from 1, \ldots, $r$ (where $r = \text{rank}(D)$) by using

- lower case latin indices $b, c, d, \ldots$ running from 1 to $\ell$,
- upper case latin indices $I, J, K, \ldots$ running from $\ell + 1$ to $r$.

\[3\text{we do not use the letter } a \text{ since its typography is very similar to that of } \alpha.\]
The following lemma shows that it is possible to extend locally the gauge momentum generators \( \{ Z_b \} \) to a minimal set of generators of \( D \) that is adapted at points of \( Q_f \). Note that if a distribution is regular, then a minimal set of generators is a set of vector fields that defines a basis of the distribution at each point.

**Lemma 5.2.** Let \( Z_b \), for \( b = 1, \ldots, \ell \), be given linearly independent equivariant vector fields in \( D \) and let \( q \in Q_f \). There is a \( G \)-invariant neighbourhood \( U \) of \( q \) in \( Q_f \) on which there exist equivariant vector fields \( Y_1 \) (\( I = \ell + 1, \ldots, r \) such that \( \{ Z_b, Y_1 \} \) generate sections of \( D \) on \( U \).

**Proof.** Let \( S \) be a submanifold of \( Q_f \) for which \( T_q S \oplus T_q (G \cdot q) = T_q Q \) (that is, \( S \) is a slice to the orbit). Then the natural map \( G \times S \to Q \) defined by \( (g, s) \mapsto g \cdot s \) defines an equivariant diffeomorphism in a neighbourhood of \( Q \) and hence in a neighbourhood of \( G \cdot q \) (by the inverse function theorem). The image \( U \) of such a neighbourhood is called a *tubular neighbourhood* of \( q \) (or of the orbit \( G \cdot q \)).

Let \( V \) be any vector field on \( Q \) defined in a neighbourhood of \( q \). Then its restriction to \( S \) can be extended to an equivariant vector field \( Y \) on the tubular neighbourhood simply by the formula

\[
Y(g, s) = TgV(s).
\]

If \( V \) is a section of a \( G \)-invariant distribution, then so is the resulting equivariant vector field \( Y \).

Since \( D \) is a smooth distribution, the given set of vector fields \( \{ Z_b \} \) can be extended to a minimal set of generators \( \{ Z_b, V_I \} \) of \( D \). For each of the \( V_I \), restrict to \( S \) and extend by equivariance to define \( Y_I \) as above. These span \( D \) by dimension count: they remain linearly independent in a neighbourhood of \( q \) (as they have the same value at \( q \) as the \( V_I \), and they are sections of \( D \)).

The proof of the lemma uses in an essential way that the basis point \( q \) has trivial isotropy. At points with non-trivial isotropy one would not expect there to be a set of equivariant vector fields that span \( D \).

The following theorem shows that, for an adapted basis \( \{ Z_b, Y_\alpha \} \), the coefficients

\[
C_{\alpha \beta \gamma} := \langle [X_\alpha, X_\beta], X_\gamma \rangle
\]

satisfy

\[
C_{b \alpha \beta} = -C_{b \beta \alpha}.
\]  

(5.1)

This is central to the definition of the 3-form \( \Lambda \) in Lemma 5.4 below. Moreover, this property serves to characterise gauge momenta.

**Theorem 5.3.** Let \( \{ X_\alpha \} \) (\( \alpha = 1, \ldots, r = \dim(D) \)) be globally defined equivariant vector fields on \( Q \) which on \( Q_f \) generate \( D \). Suppose moreover that \( X_1 \) is a section of \( \mathcal{S} \). Then \( X_1 \) generates a gauge momentum if and only if, for all \( \alpha, \beta \),

\[
C_{1 \alpha \beta} = -C_{1 \beta \alpha}.
\]

**Proof.** Consider the quantities \( F_{\alpha \beta} = \langle X_\alpha, X_\beta \rangle \). Since the \( X_\alpha \) are equivariant and the metric is \( G \)-invariant, the \( F_{\alpha \beta} \) are invariant functions. Since \( X_1 \) is tangent to group orbits, it follows that \( X_1 [F_{\alpha \beta}] = 0 \). Thus

\[
0 = X_1 [F_{\alpha \beta}]
= \mathcal{L}_{X_1}(\mathcal{S}) (X_\alpha, X_\beta) + \langle [X_1, X_\alpha], X_\beta \rangle + \langle X_\alpha, [X_1, X_\beta] \rangle
= \mathcal{L}_{X_1}(\mathcal{S}) (X_\alpha, X_\beta) + C_{1 \alpha \beta} + C_{1 \beta \alpha}.
\]

Here \( \mathcal{L}_{X_1}(\mathcal{S}) \) is the Lie derivative of the metric along \( X_1 \). Since \( X_1 \) is a section of \( \mathcal{S} \) it follows that it annihilates the potential energy. Consequently, by Theorem 4.3, \( X_1 \) generates a gauge momenta.
if and only if it is an infinitesimal isometry on \( D \). Over \( Q_f \) this latter condition is equivalent, by definition, to \( \mathcal{L}_{X_i}(S)(X_\alpha,X_\beta) = 0 \) for all \( \alpha, \beta = 1, \ldots, r \). This proves the theorem over \( Q_f \).

Finally, note that if \( p_{X_i} \) is a gauge momentum on \( D^* \) restricted to \( Q_f \), it is also one over all of \( D^* \), since \( Q_f \) is an open dense subset of \( Q \). \( \square \)

This theorem can be used to find gauge momenta in examples. See the treatment of the solid of revolution that rolls without slipping on the plane in Section 7.

5.2. The 3-forms \( \Lambda \). From now on, we assume that the subbundle \( W \) of \( TQ \) with the property \( TQ = D \oplus W \) is \( G \)-invariant; this implies that the identification of \( D^* \) with a subbundle of \( T^*Q \) is equivariant. A possibility to achieve this is to take \( W = D^\perp \) but, as indicated before, other choices may simplify coordinate calculations in concrete examples.

**Lemma 5.4.** There exists a globally defined \( G \)-invariant 3-form \( \Lambda \) on \( Q_f \) with the following properties

(i) \( \Lambda \) vanishes upon contraction with elements of \( W \) (i.e. it can be interpreted as a section of \( \wedge^3(D^*) \))

(ii) if \( \{ X_\alpha \} = \{ Z_b, Y_I \} \) is an adapted basis of sections of \( D \) on \( Q_f \) then

\[
\Lambda(Z_b, X_\alpha, X_\beta) = \langle [Z_b, X_\alpha], X_\beta \rangle,
\]

(5.2)

for \( 1 \leq b \leq \ell, 1 \leq \alpha, \beta \leq r \).

**Remarks 5.5.**

(i) It follows from Theorem 5.3 that condition (ii) in the above Lemma is consistent with the antisymmetry relation \( \Lambda(Z_b, X_\alpha, X_\beta) = -\Lambda(Z_b, X_\beta, X_\alpha) \). On the other hand, in general, \( \Lambda(X_\alpha, X_\beta, Z_b) \) does not equal \( \langle [X_\alpha, X_\beta], Z_b \rangle \).

(ii) Conditions (i) and (ii) of the lemma impose no restriction on the values of \( \Lambda(Y_I, Y_J, Y_K) \) which leads to the possible non-uniqueness of \( \Lambda \). However, if \( r - \ell < 3 \), given that the indices \( I, J, K \) run from 1 to \( r - \ell \), then \( \Lambda(Y_I, Y_J, Y_K) = 0 \) and formula (5.2) completely characterises \( \Lambda \). Therefore, in this case Lemma 5.4 defines a unique 3-form on \( Q_f \).

(iii) Equation (5.2) bears a similarity with the definition of the Cartan 3-form on Lie groups. However the vector fields involved are only equivariant and not necessarily tangent to the group orbit.

(iv) It is worth pointing out that while the construction of \( \Lambda \) depends explicitly on the kinetic energy (metric), it does not depend on the potential part of the Lagrangian provided it is \( G \)-invariant. This \( G \)-invariance is required because the vector fields \( Z_b \) must annihilate the potential (see Theorem 4.3).

The proof of Lemma 5.4 can be found in the Appendix (Proposition A.3). Note that the local expression for a 3-form \( \Lambda \) satisfying the conditions in the lemma may be given in terms of an adapted basis of sections \( \{ X_\alpha \} = \{ Z_b, Y_I \} \) of \( D \) by

\[
\Lambda = \frac{1}{b} B_{\alpha\beta\gamma} \mu^\alpha \land \mu^\beta \land \mu^\gamma,
\]

(5.3)

where the coefficients \( B_{\alpha\beta\gamma} \) are \( G \)-invariant functions, alternating in the indices, that satisfy

\[
B_{b\beta\gamma} = \langle [Z_b, X_\beta], X_\gamma \rangle.
\]

(5.4)

As usual, in (5.3), \( \{ \mu^\alpha \} \) is a basis of sections of \( D^* = W^* \) that is dual to \( \{ X_\alpha \} \). Namely, they are locally defined 1-forms on \( Q \) that annihilate \( W \) and satisfy \( \mu^\alpha(X_\beta) = \delta^\alpha_\beta \). The \( G \)-invariance of \( W \) guarantees that \( \mu^\alpha \) are also \( G \)-invariant.
5.3. **Almost Poisson brackets having gauge momenta as Casimirs.** We assume for the remainder of this section that among the 3-forms of Lemma 5.4, there exists at least one that admits a smooth extension to the points of $Q$ having non-trivial isotropy. We point out that this property is satisfied in all of the examples that we considered. Denote the resulting 3-form on $Q$ by $\Lambda$.

We now use this 3-form $\Lambda$ to apply the construction outlined in Section 3 to construct a bracket $\Pi_{\text{nh}}^\Lambda$ for our nonholonomic system to obtain our main result.

**Theorem 5.6.** The bivector $\Pi_{\text{nh}}^\Lambda$ on $D^*$ is $G$-invariant and the gauge momenta $\pi_b$, $b = 1, \ldots, \ell$, are Casimir functions of the induced bracket on the reduced space $D^*/G$.

**Proof.** That $\Pi_{\text{nh}}^\Lambda$ is $G$-invariant follows from the invariance of the 3-form $\Lambda$ and of the bivector $\Pi_{\text{nh}}$. For the proof that the gauge momenta are Casimir functions, we first prove this on $Q_J$ and then deduce the full statement by continuity.

We may obtain a local expression for $\Pi_{\text{nh}}^\Lambda$ on $Q_J$ using formula (3.6) in Theorem 3.2. Recall that $B^T_{\alpha \beta} := \Omega^{\gamma \delta} B_{\alpha \beta \delta}$, where $B_{\alpha \beta \delta} = \Lambda(X_\alpha, X_\beta, X_\delta)$ as indicated in (3.1). Using (5.4) and (2.6) we find

\[
B^T_{\alpha \beta} = C^T_{\alpha \beta}, \quad B^T_{\beta \alpha} = C^T_{\beta \alpha}, \quad B^T_{ij} = \Omega^{bc} C_{ij} + \Omega^{bk} B_{ijk}, \quad B^T_{ij} = \Omega^{bc} C_{ij} + \Omega^{bk} B_{ijk}.
\]

Therefore, the expression for $\Pi_{\text{nh}}^\Lambda$ becomes

\[
\Pi_{\text{nh}}^\Lambda = \rho^i \partial_i \Lambda \wedge \partial \pi_a + \frac{1}{2} \left( \left( \Omega^{bc} C_{ij} + \Omega^{bk} B_{ijk} - C^T_{ij} \right) \pi_b + \left( \Omega^{bk} C_{bkl} + \Omega^{kl} B_{ijkl} - C^T_{ij} \right) \pi_k \right) \partial \pi_i \wedge \partial \pi_j.
\]

A direct calculation using the above formulae for $\Pi_{\text{nh}}^\Lambda$ gives

\[
(\Pi_{\text{nh}}^\Lambda)^T(d\pi_b) = -\rho^i \partial_i \Lambda.
\]

We claim that this vector field on $D^*$ is tangent to the group orbits of the lifted action of $G$ to $D^*$. Indeed, given that $Z_b = \rho^i \partial_i \Lambda$ is a section of $\mathcal{S}$, for any $q \in Q_J$ there exists a Lie algebra element $\xi(q) \in \mathfrak{g}$ such that $Z_b(q)$ coincides with the infinitesimal generator of $\xi(q)$ at $q$. Namely,

\[
Z_b(q) = \left. \frac{d}{ds} \right|_{s=0} \exp(s \xi(q)) \cdot q.
\]

Now, since all vector fields in the basis $\{X_\alpha\} = \{Z_b, Y_j\}$ are equivariant, the corresponding momenta $\pi_a$ are invariant functions. Therefore, in coordinates $(q^i, \pi_a)$, the local expression for the infinitesimal generator of $\xi(q)$ of the lifted action of $G$ to $D^*$ is

\[
\left. \frac{d}{ds} \right|_{s=0} \exp(s \xi(q)) \cdot (q, \pi) = \rho^i \partial_i \Lambda(q).
\]

which establishes the claim.

It follows that the Hamiltonian vector field associated to $\pi_b$ via the induced bracket on the reduced space vanishes; i.e. $\pi_b$ is a Casimir of the reduced bracket, over all points of the open dense subset $Q_J$. It then follows by continuity that $\pi_b$ (which is a globally defined invariant smooth function) is a Casimir everywhere. \qed

**Remark 5.7.** We stress that the theorem assumes that among the 3-forms of Lemma 5.4, there exists at least one that admits a smooth extension to the points of $Q$ having non-trivial isotropy, and that this property is satisfied in all of the examples that we considered. However, we do not have a proof that such an extension always exists. One may of course restrict the construction to $Q_J$, but the differential equation may then fail to be complete.
We conclude this section with some dynamical consequences of the above theorem. For free actions, "directions" that annihilate the group orbits are spanned by the pull-backs to $D^*$ of 1-forms on $D^*/G$. For more general actions, one replaces such forms by basic 1-forms, which are those $G$-invariant 1-forms $\beta$ on $D^*$ satisfying $\beta(u) = 0$ for all $u$ tangent to the group orbit. Note that the differential of any invariant function is basic.

**Corollary 5.8.** Suppose that the generic level sets of the gauge momenta on the reduced space $D^*/G$ are 2-dimensional. Then

(i) the reduced bracket defined on $D^*/G$ induced by $\Pi_{n_h}^\Lambda$ satisfies the Jacobi identity and the system is Hamiltonizable,

(ii) the flow of the system on $D^*$ preserves the restriction of $(\Pi_{n_h}^\Lambda)^\sharp$ to directions that annihilate the group orbits. More precisely,

$$\mathcal{L}_{X_{n_h}}(\Pi_{n_h}^\Lambda)^\sharp(\beta) = 0$$

for every basic 1-form $\beta$ on $D^*$. In particular this holds when $\beta = d\pi_b$, for any gauge momentum $\pi_b$ and therefore, for each $b = 1, \ldots, \ell$,

$$[X_{n_h}, X_{\pi_b}] = 0,$$

where $X_{\pi_b} = (\Pi_{n_h}^\Lambda)^\sharp(d\pi_b)$.

**Proof.** (i) The proof of this is simply that on any 2-dimensional manifold, every almost-Poisson structure is in fact Poisson (as the Schouten-Nijenhuis bracket of the structure with itself, which is an alternating 3-tensor, must vanish).

(ii) The proof proceeds stratum by stratum: the flow of the vector field $X_{n_h}$ preserves the strata because it is equivariant. On the stratum where the action is free, any basic 1-form $\beta$ is the pull-back $\beta = \tau^*\alpha$ for some 1-form $\alpha$ on $(D^*/G)_f$, where $\tau: D^* \to D^*/G$ is the projection. Then,

$$\mathcal{L}_{X_{n_h}}(\Pi_{n_h}^\Lambda)^\sharp(\beta) = \mathcal{L}_{X_{n_h}}(\Pi_{n_h}^\Lambda)^\sharp(\tau^\ast\alpha) = (\tau_\ast(\mathcal{L}_{X_{n_h}}(\Pi_{n_h}^\Lambda)))^\sharp(\alpha)$$

and by the natural properties of Lie derivatives, this is

$$\mathcal{L}_{\tau_\ast X_{n_h}}(\tau_\ast(\Pi_{n_h}^\Lambda))^\sharp(\alpha)$$

and this vanishes by part (i), and the fact that Hamiltonian flows are Poisson.

The argument for other strata proceeds in the same way, since the restriction of $\tau$ to an orbit-type stratum is a submersion [23].

5.4. **Final observations.**

5.4.1. The case $r - \ell < 3$. If the difference between the rank of the distribution $r$ and the number of independent gauge momenta $\ell$ is less than 3, then as indicated in Remark 5.5(ii), the 3-form $\Lambda$ of Lemma 5.4 is uniquely determined (given the choice of $W$) and, therefore, the same is true for the bracket $\Pi_{n_h}^\Lambda$ appearing in Theorem 5.6. In this case, since the indices $I, J, K$ run from 1 to $r - \ell$, all coefficients $B_{IJK}$ are 0 by the alternating property, and (5.5) simplifies to

$$\Pi_{n_h}^\Lambda = \rho'_a \partial_{q^a} \wedge \partial_{\pi_a} + \frac{1}{2} \left( \gamma_{bIJ} C_{bIJ} - C'_{IJ} \right) \pi_i \partial_{\pi_i} \wedge \partial_{\pi_i}. \quad (5.6)$$

This formula gives an expression for the bracket whose invariant gauge momenta pass to the reduced space as Casimir functions for a number of classical nonholonomic problems, such as the nonholonomic particle, the Chaplygin sphere, the problem of solids of revolution that roll without slipping on a horizontal plane or on a sphere, a homogeneous ball that rolls without slipping on a surface of...
revolution, and also the class of rigid bodies subject to generalised rolling constraints introduced in [3].

In order to use the above formula in examples one needs to obtain an expression for an adapted basis of sections to compute the coefficients $\rho^I$, $c^I_{\alpha\beta}$, $s^{I\alpha\beta}$. In sections 6 and 7 we illustrate how this is done for the Chaplygin sphere and for a solid of revolution that rolls without slipping on the plane. For these examples we also compute the 3-form $\Lambda$, both of which are given by interesting expressions involving the Cartan 3-form on $SO(3)$.

5.4.2. Why do we need to modify the brackets? We now address the following fundamental question: why does one need to modify the nonholonomic bracket $\Pi_{nh}$ by $\Lambda$ to guarantee that the invariant gauge momenta are Casimir functions of the reduced bracket?

First we note that the question is only relevant if the rank $r$ of the constraint distribution $D$ is greater than 2. Indeed, if $r \leq 2$ then the section $\Lambda$ of $\wedge^3(D^*)$ of Lemma 5.4 vanishes and the bracket $\Pi_{nh}^\Lambda$ in Theorem 5.6 coincides with $\Pi_{nh}$. An instance of this general situation is encountered by Balseiro [6] in her treatment of the nonholonomic particle.

Consider then the case where $r > 2$, and assume for simplicity that there is only one (invariant) gauge momentum whose unique generator taking values in $D$ is the equivariant vector field $Z$. A fundamental property of $Z$, that contrasts with the situation encountered in holonomic systems, is that even though $Z$ is tangent to the group orbits, it may not coincide with the infinitesimal generator $\xi_Q$ of a constant Lie algebra element $\xi \in g$ [8]. It is this property of $Z$, that is not encountered in holonomic mechanics, that leads to the need of the modification of the nonholonomic bracket. Indeed, if $Z$ were equal to $\xi_Q$ for a fixed $\xi \in g$, and $\{Z,Y_I\}$ is an adapted basis of sections of $D$, then $[Z,Y_I] = 0$ by equivariance of $Y_I$. Therefore the 3-form $\Lambda = 0$ satisfies the conditions of Lemma 5.4 and the corresponding bracket $\Pi_{nh}^\Lambda$ in Theorem 5.6 coincides with $\Pi_{nh}$.

Examples of nonholonomic systems where $Z = \xi_Q$ for a fixed $\xi \in g$ are rare and one usually must perform the modification of the nonholonomic bracket developed above to guarantee that the invariant gauge momenta pass to the quotient space as Casimir functions. An (somewhat artificial) example where this condition holds and no modification of the nonholonomic bracket $\Pi_{nh}$ is needed to accomplish the aforementioned goal, is the “rank 1” case of a rigid body subject to generalised rolling constraints treated in [3].

5.4.3. Future work. As mentioned in the introduction, it would be interesting to extend the results of our paper to the case where the gauge momenta of a nonholonomic system are not $G$-invariant. In this case it may be possible to find modifications of the nonholonomic bracket to guarantee that the $G$-invariant (nonlinear) functions of the gauge momenta pass to the quotient space as Casimir functions. This seems to be possible to accomplish for some examples, and in particular for the nonholonomic hinge [10].

6. The Chaplygin sphere

This problem was considered by Chaplygin in [20] and concerns the motion of an inhomogeneous sphere, whose center of mass coincides with its geometric centre, which is the rolling without slipping on a fixed plane. The vertical component of the angular momentum of the sphere is a gauge momentum for this problem. Using our method it is possible to construct an almost Poisson bracket for the system that upon reduction has this angular momentum as Casimir. This bracket was first found by Borisov and Mamaev [13].

We assume that two of the moments of inertia of the sphere coincide to simplify the algebra and better illustrate how our construction works but a similar approach works in the general case.
The configuration space for the problem is \( Q = SO(3) \times \mathbb{R}^2 \). The attitude matrix \( R \in SO(3) \) specifies the orientation of the sphere by relating a body frame centred at the centre of the sphere, with a space frame whose third axis is perpendicular to the fixed plane. Let \((x,y) \in \mathbb{R}^2\) be the spatial coordinates of the contact point of the sphere with the plane.

We will use Euler angles as local coordinates for \( SO(3) \). In accordance with the \textit{x-convention}, see e.g. [44], we write a matrix \( R \in SO(3) \) as

\[
R = \begin{pmatrix}
\cos \psi \cos \phi - \cos \theta \sin \phi \sin \psi & -\sin \psi \cos \phi - \cos \theta \sin \phi \cos \psi & \sin \theta \sin \phi \\
\cos \psi \sin \phi + \cos \theta \cos \phi \sin \psi & -\sin \psi \sin \phi + \cos \theta \cos \phi \cos \psi & -\sin \theta \cos \phi \\
\sin \theta \sin \psi & \sin \theta \cos \psi & \cos \theta 
\end{pmatrix}, \tag{6.1}
\]

where the Euler angles \( 0 < \phi, \psi < 2\pi, 0 < \theta < \pi \). According to this convention, we obtain the following expressions for the angular velocity in space coordinates \( \dot{\omega} \), and in body coordinates \( \Omega \) (see e.g. [44]):

\[
\omega = \begin{pmatrix}
\dot{\theta} \cos \phi + \dot{\psi} \sin \phi \sin \theta \\
\dot{\theta} \sin \phi - \dot{\psi} \cos \phi \sin \theta \\
\dot{\psi} + \dot{\phi} \cos \theta
\end{pmatrix}, \quad \Omega = \begin{pmatrix}
\dot{\theta} \cos \phi + \dot{\psi} \sin \phi \sin \theta \\
-\dot{\theta} \sin \phi + \dot{\psi} \cos \phi \sin \theta \\
\dot{\phi} \cos \theta + \dot{\psi}
\end{pmatrix}. \tag{6.2}
\]

The constraints of rolling without slipping are

\[
\dot{x} = R \omega_2 = R(\theta \sin \phi - \dot{\psi} \cos \phi \sin \theta), \quad \dot{y} = -R \omega_1 = -R(\dot{\theta} \cos \phi + \dot{\psi} \sin \phi \sin \theta), \tag{6.3}
\]

where \( R \) is the radius of the sphere.

Assuming that the third axis of the body frame is the axis of symmetry of the sphere, the Lagrangian is

\[
L = \frac{1}{2} \left( I_1 (\dot{\theta} \cos \phi + \dot{\psi} \sin \phi \sin \theta)^2 + I_1 (-\dot{\theta} \sin \phi + \dot{\psi} \cos \phi \sin \theta)^2 + I_3 (\dot{\phi} \cos \theta + \dot{\psi})^2 + m(\dot{x}^2 + \dot{y}^2) \right) - V(\theta, \psi)
\]

where \( I_1, I_1, I_3 \) are the principal moments of inertia and \( m \) is the total mass of the sphere. Here \( V \) is a potential energy chosen to be invariant under the symmetry group action defined below. Expanding the above expression, one finds that the kinetic energy metric is

\[
\mathcal{G} = (I_1 \sin^2 \theta + I_3 \cos^2 \theta) d\phi^2 + I_1 d\theta^2 + I_3 d\psi^2 + m d^2 x + m d^2 y + 2 I_3 \cos \theta \cos \phi d \phi d \psi.
\]

The symmetry group is \( G = SE(2) \). The action of \((\vartheta, a, b) \in SE(2) \) on \( Q \) is free and proper and is given in the above coordinates by

\[
(\vartheta, a, b) : (\phi, \theta, \psi, x, y) \mapsto (\vartheta + \varphi, \theta, \psi, x \cos \vartheta - y \sin \vartheta + a, y \cos \vartheta + x \sin \vartheta + b).
\]

One checks that both the constraints and the Lagrangian are invariant under the lift of this action to \( TQ \). The constraint distribution \( D \) has rank 3 and is spanned by the \( G \)-equivariant vector fields

\[
Z_1 = \partial_\varphi, \quad Y_2 = \partial_\vartheta + R \sin \phi \partial_\alpha - R \cos \phi \partial_\gamma, \quad Y_3 = \partial_\psi - R \cos \phi \sin \theta \partial_\alpha - R \sin \phi \sin \theta \partial_\gamma. \tag{6.4}
\]

We note that the basis of sections \( \{X_\alpha\} = \{Z_1, Y_2, Y_3\} \) is adapted in the sense of Definition 5.1. Indeed, its elements are equivariant vector fields and \( Z_1 \) generates a gauge momentum. This can be seen by noticing that \( Z_1 \) is tangent to the group orbits and that its tangent lift in bundle coordinates is \( Z_1^{TQ} = \partial_\varphi \). Hence \( Z_1^{TQ}[L] = 0 \) and therefore \( p_{Z_1} = 0 \) by (4.1).

The non-zero coefficients \( C_{\alpha \beta \gamma} \) (with \( \alpha < \beta \)) are computed to be

\[
C_{123} = -m R^2 \sin \theta, \quad C_{132} = m R^2 \sin \theta, \quad C_{233} = m R^2 \sin \theta \cos \theta. \tag{6.5}
\]

All other terms may be determined by the skew-symmetry \( C_{\alpha \beta \gamma} = -C_{\beta \alpha \gamma} \). Note that \( C_{123} = -C_{132} \) which serves as a double check that \( p_{Z_1} \) is a gauge momentum in view of Theorem 5.3.
We select the dual 1-forms \( \{u^\alpha\} = \{d\varphi, d\theta, d\psi\} \) which amounts to identifying \( D^* = (\text{span}\{\partial_x, \partial_y\})^\circ = T^*SO(3) \times \mathbb{R}^2 \). This is allowed in our construction since \( W = \text{span}\{\partial_x, \partial_y\} \subset TQ \) is invariant under the \( \text{SE}(2) \) action. The formula (5.3) defines a unique 3-form \( \Lambda \) since the rank of \( D \) is \( r = 3 \) and we have \( l = 1 \) gauge momentum, and so \( r - l < 3 \). Such unique 3-form is given by

\[
\Lambda = -mR^2 \sin \theta d\varphi \wedge d\theta \wedge d\psi.
\]

A coordinate independent expression for \( \Lambda \) may be given in terms of the unique left-invariant 1-forms \( \lambda^1, \lambda^2, \lambda^3 \) on \( SO(3) \) at the group identity are dual to the canonical basis of \( \mathfrak{so}(3) \cong \mathbb{R}^3 \). These 1-forms have local expressions:

\[
\lambda^1 = \sin \psi \sin \theta d\varphi + \cos \psi d\theta, \quad \lambda^2 = \cos \psi \sin \theta d\varphi - \sin \psi d\theta, \quad \lambda^3 = \cos \theta d\varphi + d\psi. \tag{6.6}
\]

So we can write

\[
\Lambda = mR^2 \lambda^1 \wedge \lambda^2 \wedge \lambda^3.
\]

Therefore, up to the constant factor of \( mR^2 \), \( \Lambda \) equals the Cartan bi-invariant volume form on \( SO(3) \) normalised to have volume one on the unit cube of \( \mathfrak{so}(3) \cong \mathbb{R}^3 \). This also holds for the general Chaplygin sphere with arbitrary moments of inertia as had been indicated in [33].

An expression for the bracket \( \{\cdot, \cdot\} \) defined by \( \Pi^\alpha_{\beta\gamma} \), whose reduction has \( \pi_1 \) as a Casimir, and describes the nonholonomic dynamics, can be obtained using (5.6). All the brackets between the coordinate functions \( \varphi, \theta, \psi, x, y \) are zero. The non-zero brackets between the coordinates and the momenta \( \pi_{\alpha} \) are obtained using (6.4):

\[
\{\varphi, \pi_1\} = 1, \quad \{\theta, \pi_2\} = 1, \quad \{x, \pi_2\} = R \sin \varphi, \quad \{y, \pi_2\} = -R \cos \varphi,
\]

\[
\{\psi, \pi_3\} = 1, \quad \{x, \pi_3\} = -R \cos \varphi \sin \theta, \quad \{y, \pi_3\} = -R \sin \varphi \sin \theta.
\]

Also, we have

\[
\{\pi_1, \pi_2\} = 0, \quad \{\pi_1, \pi_3\} = 0,
\]

that must hold since \( \pi_2 \) and \( \pi_3 \) are invariant functions and \( \pi_1 \) is a Casimir of the reduced bracket. To compute \( \{\pi_2, \pi_3\} \) one needs to compute the inverse of the \( 3 \times 3 \) symmetric matrix \( G \) with entries \( G_{\alpha\beta} = (X_\alpha, X_\beta) \). One gets

\[
G^{-1} = \frac{1}{K(\theta) \sin^2 \theta} \begin{pmatrix}
I_3 + mR^2 \sin^2 \theta & 0 & -I_3 \cos \theta \\
0 & \frac{K(\theta) \sin^2 \theta}{I_1 + mr} & 0 \\
-I_3 \cos \theta & 0 & I_1 \sin^2 \theta + I_3 \cos^2 \theta
\end{pmatrix} \tag{6.7}
\]

where

\[
K(\theta) = I_1 mR^2 \sin^2 \theta + I_3 mR^2 \cos^2 \theta + I_1 I_3.
\]

With the aid of these formulae and (6.5) one obtains

\[
C_{23}^1 = -\frac{mR^2 I_3 \cos \theta}{K(\theta) \sin \theta}, \quad C_{23}^2 = 0, \quad C_{23}^3 = \frac{mR^2 \cos \theta (I_1 \sin^2 \theta + I_3 \cos^2 \theta)}{K(\theta) \sin \theta},
\]

and hence, by (5.6) we get

\[
\{\pi_2, \pi_3\} = -\frac{mR^2 (I_3 + mR^2) \sin \theta}{K(\theta)} \pi_1 - \frac{mR^2 (I_1 - I_3) \cos \theta \sin \theta}{K(\theta)} \pi_3.
\]

Let us now write the bracket in terms of more standard physical variables for the problem. First, we introduce the Poisson vector \( \gamma := R^T e_3 \) that gives the coordinates on the body frame of the vector normal to the plane on which the rolling takes place. Its components are

\[
\gamma_1 = \sin \theta \sin \psi, \quad \gamma_2 = \sin \theta \cos \psi, \quad \gamma_3 = \cos \theta. \tag{6.8}
\]
Next, the angular momentum vector about the contact point, expressed in the body frame is given by

$$M = I\Omega + mR^2\gamma \times (\Omega \times \gamma),$$

where $I = \text{diag}(I_1, I_1, I_3)$ is the tensor of inertia, and $\times$ denotes the vector product in $\mathbb{R}^3$. To write $M$ in terms of our coordinates and the momenta $\pi_a$ start by noticing the quasi-velocities $v^a$ defined by the basis $\{Z_1, Y_2, Y_3\}$ given by (6.4) satisfy

$$\dot{\phi} = v^1, \quad \dot{\theta} = v^2, \quad \dot{\psi} = v^3. \quad (6.9)$$

Next, write $v^a = \delta^{ab} \pi_b$ using the expression for $\delta^{-1}$ given above in (6.7). Combining this with the expression for $\Omega$ given in (6.2) and the expression for $\gamma$ on (6.8) one gets

$$M_1 = \frac{\sin \psi \pi_1 + \cos \psi \sin \theta \pi_2 - \sin \psi \cos \theta \pi_3}{\sin \theta}, \quad M_2 = \frac{\cos \psi \pi_1 - \sin \psi \sin \theta \pi_2 - \cos \psi \cos \theta \pi_3}{\sin \theta},$$

and $M_3 = \pi_3$. Note that both vectors $M$ and $\gamma$ are $SE(2)$-invariant and its components drop down to the quotient $D^*/SE(2)$. In fact, as a manifold $D^*/SE(2) = \mathbb{R}^3 \times S^2$. The entries of $M$ serve as coordinates on the $\mathbb{R}^3$ factor while the components of $\gamma$ are redundant coordinates on $S^2$. In particular notice that $\pi_1$ equals the vertical component of the angular momentum vector; i.e. $\pi_1 = (M, \gamma)$, where $(\cdot, \cdot)$ is the euclidean scalar product in $\mathbb{R}^3$.

By direct calculation using the above formulae for the bracket one gets

$$\{\gamma, \gamma\} = 0, \quad \{M_i, \gamma\} = -\epsilon_{ijk} \gamma_k, \quad \{M_i, M_j\} = -\epsilon_{ijk} (M_k - mR^2 (\Omega, \gamma) \gamma_k), \quad (6.10)$$

where the alternating tensor $\epsilon_{ijk}$ equals 0 if two of its indices are equal, it equals 1 if $(i, j, k)$ is a cyclic permutation of $(1, 2, 3)$ and it equals $-1$ otherwise. The term $(\Omega, \gamma)$ is the spinning speed of the ball. It can be written in terms of $M$ and $\gamma$ using the expression

$$\Omega = AM + mR^2 \frac{(AM, \gamma)}{1 - mR^2 (A\gamma, \gamma)} A\gamma,$$

where $A = (I + mR^2)^{-1}$. (For the calculations it is useful to note that $K(\theta) = (I_1 + mR^2) (I_3 + mR^2) (1 - (A\gamma, \gamma))$).

The above formulae determine the reduced bracket in the quotient space $D^*/SE(2)$. The reduced equations of motion (see e.g. [30])

$$\dot{M} = M \times \Omega, \quad \dot{\gamma} = \gamma \times \Omega,$$

are Hamiltonian with respect to $H = \frac{1}{2} (M, \Omega) + V(\gamma)$.

The bracket (6.10) has rank 4 and, even though it does not satisfy the Jacobi identity, its characteristic distribution is integrable - the leaves of the corresponding foliation are the level sets of the Casimir function $(M, \gamma)$. As was first noticed in [13], the bracket obtained by multiplication by the conformal factor $\sqrt{1 - mR^2 (A\gamma, \gamma)}$ does satisfy the Jacobi identity and Hamiltonizes the problem. This kind of multiplication is commonly interpreted as a *time reparametrisation*. We mention that the conformal factor is intimately related with the preserved measure of the problem [30].

7. A SOLID OF REVOLUTION ROLLING WITHOUT SLIPPING ON A FIXED PLANE

We consider a convex body of revolution with a smooth surface that rolls without slipping on a fixed plane. This problem was originally considered by Routh [46] in the case of a spherical body, and by Chaplygin [22] and Appel [2] in the general case; see [14] for historical details. Our treatment and notation is close to the one used in [14].

As for the Chaplygin sphere, the configuration space is $Q = SO(3) \times \mathbb{R}^2$. We denote by $u = (x, y, z)$ the coordinates of the centre of mass $O$ of the body with respect to an inertial frame. Our choice of
inertial frame is such that the fixed plane where the rolling takes place corresponds to \( z = 0 \). On the other hand, the body frame is chosen to be centred at the centre of mass, and having third axis \( E_3 \) along the symmetry axis of the body.

Denote by \( R \in SO(3) \) the attitude matrix that relates the two frames. The constraints of rolling without slipping are given by

\[
u = R \rho,
\]

where \( \rho \) is the vector from contact point \( P \) to the centre of mass of the body \( O \) written in the body frame (see Figure 1). The last component in the above equation is in fact the holonomic constraint

\[
z = (\rho, \gamma),
\]

where, just like in the previous section, \( \gamma = R^T e_z \) is the vector normal to the fixed plane written in body coordinates. We assume that the orientation of \( e_z \) is such that that \( \gamma \) is the inward normal vector of the body at \( P \). Hence, the inverse of the classical Gauss map from differential geometry of surfaces, allows us to express \( \rho \) as a function of \( \gamma \) in the form

\[
\rho_1 = f_1(\gamma_1) \gamma_1, \quad \rho_2 = f_1(\gamma_2) \gamma_2, \quad \rho_3 = f_2(\gamma_3).
\]

By writing \( \gamma_3 = \cos \theta \) in accordance with the Euler angles introduced in Section 6, then we may write \( f_1, f_2 \) as functions of \( \theta \). Their geometric meaning can be read off from Figure 1 that depicts the curve that generates the surface of revolution in the perspective of the body frame. Notice that \( \pi - \theta \) is the angle between the \( E_3 \) axis and the outer normal vector to the surface and \( a_1(\theta) := f_1(\theta) \sin \theta \) is the distance between \( \mathcal{P} \) and the \( E_3 \) axis. The figure also illustrates the height \( z \) of the centre of mass. Note that \((z, \theta - \pi/2)\) are polar coordinates with respect to \( E_1 \) for the pedal curve of the generating curve about \( O \).

Recall that the principal lines of curvature of a surface of revolution are the parallels (the circles \( \theta = c t \)) and the meridians (perpendicular to the parallels and having shape equal to the rotating curve). All of the meridians meet at the poles where \( \gamma_3 = \pm 1 \ (\theta = 0, \pi) \). It is clear from Figure 1 that \( f_1, f_2 \) can be extended as even \( 2\pi \)-periodic functions of \( \theta \) that, consequently, have local extrema at the poles. The radii of curvature along the parallels \( R_p \) and the meridians \( R_m \) are given by

\[
R_p = f_1, \quad R_m = -\frac{1}{\sin \theta} \frac{df_2}{d\theta} = \frac{df_2}{d\gamma_3},
\]

Note that \( R_p \) and \( R_m \) are smooth, and positive and their values coincide at the poles. As functions of \( \theta \), they are \( 2\pi \)-periodic and even.

Now, using (6.8) and (7.2) we have \( z(\theta) = \sin^2 \theta f_1(\theta) + \cos \theta f_2(\theta) \). In view of the identity

\[
\sin \theta \frac{d a_1}{d\theta} + \cos \theta \frac{d f_2}{d\theta} = 0,
\]

that follows from the definition of \( \theta \) in Figure 1, we have

\[
a_2(\theta) := \frac{dz}{d\theta} = \sin \theta (\cos \theta f_1(\theta) - f_2(\theta)),
\]

and we can write \( \dot{z} = a_2(\theta) \dot{\theta} \) (which is consistent with (7.1)). The kinetic energy of the system is therefore

\[
\mathcal{K} = \frac{1}{2} \langle I \Omega, \Omega \rangle + \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + a_2(\theta)^2 \dot{\theta}^2).
\]
The assumption that the body is axisymmetric implies that the inertia tensor has the form $I = \text{diag}(I_1, I_1, I_3)$. In terms of Euler angles (using (6.2)) we get the expression for the Lagrangian

$$L = \frac{1}{2} \left( (I_1 \sin^2 \theta + I_3 \cos^2 \theta) \dot{\phi}^2 + (I_1 + ma_2(\theta)^2) \dot{\theta}^2 + I_3 \dot{\psi}^2 + 2I_3 \cos \theta \dot{\phi} \dot{\psi} + m(x^2 + y^2) \right) - V(\theta),$$

(7.4)

where $m$ is the total mass of the body and the potential $V$ is an even function of $\theta$ that is invariant under the symmetries of the system that are discussed ahead. If the potential is gravitational then $V(\theta) = mgz(\theta)$. The constraints (7.1) of rolling without slipping are expressed in coordinates as

$$\dot{x} = -a_2(\theta) \cos \phi \phi + z(\theta) \sin \phi \dot{\theta} - a_1(\theta) \cos \phi \psi, \quad \dot{y} = -a_2(\theta) \sin \phi \phi - z(\theta) \cos \phi \dot{\theta} - a_1(\theta) \sin \phi \psi.$$

The symmetry group is $G = SE(2) \times S^1$ corresponding to translations and rotations on the rolling plane, and to the internal rotational symmetry of the body. This action is not free. The configurations for which the point of contact $P$ is one of the poles have $S^1$ isotropy. These configurations lie outside the Euler angle chart that has $0 < \theta < \pi$. All points in our chart have trivial isotropy, and the action of $((\vartheta, a, b), \Psi) \in SE(2) \times S^1$ on $Q$ is represented by

$$((\vartheta, a, b), \Psi) : (\phi, \theta, \psi, x, y) \mapsto (\phi + \vartheta, \theta, \psi + \Psi, x \cos \phi - y \sin \phi + a, y \cos \phi + x \sin \phi + b).$$

One checks that both the constraints and the Lagrangian are invariant under the lift of this action to $TQ$. 

**Figure 1.** The generator of a solid of revolution; see text for details.
The constraint distribution has rank 3. Along the points covered by our chart, it is spanned by the $G$-equivariant vector fields

\[
\begin{align*}
W_1 &= \partial_\theta - a_2(\theta)(\cos \phi \partial_\phi + \sin \phi \partial_\psi), \\
W_2 &= \partial_\phi - a_1(\theta)(\cos \phi \partial_\phi + \sin \phi \partial_\psi), \\
Y_3 &= \partial_\theta + z(\theta)(\sin \phi \partial_\phi - \cos \phi \partial_\psi).
\end{align*}
\]

Although the vector fields $W_1$ and $W_2$ span the space $\delta_q = D_q \cap g \cdot q$, they generally do not generate gauge momenta and hence $\{W_1, W_2, Y_3\}$ is not an adapted basis in the sense of Definition 5.1 (the vector fields $W_1, W_2$ should not be confused with the subbundle $W$ complementary to $D$). A generator $Z_1$ of a gauge momentum may be found using the ansatz

\[ Z_1 = g(\theta)W_1 + k(\theta)W_2 \tag{7.5} \]

and using Theorem 5.3 to determine the functions $g, k$. According to the theorem, if $Z_1$ generates a gauge momentum then the functions $g, k$ should be such that

\[
\begin{align*}
\langle [Z_1, Y_2], Z_1 \rangle &= \langle [Z_1, Y_2], Y_2 \rangle = \langle [Z_1, Y_3], Y_3 \rangle = 0, \\
\langle [Z_1, Y_3], Z_1 \rangle &= 0, \\
\langle [Z_1, Y_2], Y_3 \rangle &= -\langle [Z_1, Y_3], Y_2 \rangle,
\end{align*}
\]

where $\langle \cdot, \cdot \rangle$ is the kinetic energy Riemannian metric defined by the Lagrangian (7.4), and $Y_2$ is any $G$-equivariant linear combination of $W_1, W_2$ (i.e. with coefficients that are functions of $\theta$).

Independently of the choice made for $Y_2$, a short direct calculation shows that the first three identities in (7.6) hold for any functions $g, k$. On the other hand, after a long calculation, it is seen that the other two are satisfied by any solution of the following system of linear ODE’s:

\[
\frac{d}{d\theta} \begin{pmatrix} g(\theta) \\ k(\theta) \end{pmatrix} = L(\theta) \begin{pmatrix} g(\theta) \\ k(\theta) \end{pmatrix},
\]

where the $2 \times 2$ matrix $L(\theta) = \frac{1}{k(\theta)} \tilde{L}(\theta)$ with

\[ K(\theta) := I_1 I_3 + m f_1(\theta)^2 + m f_2(\theta)^2, \]

and where the entries of $\tilde{L}$ are given by

\[
\begin{align*}
L_{11}(\theta) &= m f_2(\theta) \left( \frac{R_m(\theta) - R_p(\theta)}{\sin \theta} \right) - m a_2(\theta) f_1(\theta)(I_3 + m z(\theta)f_1(\theta)), \\
L_{12}(\theta) &= m f_1(\theta) \cos \theta \left( \frac{R_m(\theta) - R_p(\theta)}{\sin \theta} \right) - m f_1(\theta) a_1(\theta)(I_3 + m z(\theta)f_1(\theta)), \\
L_{21}(\theta) &= m f_1(\theta)(I_1 \sin^2 \theta + I_3 \cos^2 \theta) \left( \frac{R_p(\theta) - R_m(\theta)}{\sin \theta} \right) + \frac{ma_2(\theta)}{\sin^3 \theta} (ma_1(\theta) a_2(\theta) z(\theta) + (R_m(\theta) - R_p(\theta)) I_3 \cos \theta + (I_3 - I_1) a_1(\theta) \sin \theta \cos \theta), \\
L_{22}(\theta) &= m \cos \theta (I_1 a_1(\theta) \sin \theta + I_3 f_2(\theta) \cos \theta) \left( \frac{R_p(\theta) - R_m(\theta)}{\sin \theta} \right) + m f_1(\theta)^2 (m z(\theta) a_2(\theta) + (I_3 - I_1) \sin \theta \cos \theta).
\end{align*}
\]

We point out that the function $K$ defined above is always positive and may be written as $K = I_1 I_3 + m \langle \dot{p}, p \rangle$ where, as in Section 6, $\langle \cdot, \cdot \rangle$ denotes the standard scalar product in $\mathbb{R}^3$. As explained in [14], the function $K$ is related to the density of an invariant measure for the system.
Proposition 7.1. There exist two independent gauge momenta of the system on $D^s$.

Proof. The local existence of the two independent gauge momenta follows from the application of the existence theorem for ODE’s to the system (7.7) to obtain two linearly independent solutions (this can be done since the matrix $L(\theta)$ is smooth, see below).

In order to show that these integrals may be extended outside of our chart, we need to argue that the generator $Z_1$ given by (7.5) admits an extension to all of $Q$. This is certainly true if all the solutions to (7.7) are $2\pi$-periodic, even functions of $\theta$. To show that this is the case, we will prove that the matrix $L(\theta)$ is smooth, odd, and $2\pi$-periodic (and therefore vanishes at $\theta = n\pi$, $n \in \mathbb{Z}$) and we will apply the following lemma.

Lemma 7.2. Let $L(t)$ be an $n \times n$ matrix depending smoothly on $t$, $T$-periodic and odd (that is, $L(-t) = -L(t)$). Then any solution to the differential equation $\dot{x}(t) = L(t)x(t)$ is even (i.e., $x(-t) = x(t)$) and $T$-periodic.

Proof. Let $x_0 \in \mathbb{R}^n$ and consider the initial value problem

$$x(t) = L(t)x(t), \quad x(0) = x_0.$$ 

The evenness of $x(t)$ follows from uniqueness of solutions: if $x(t)$ were not even then it is easy to check that $y(t) := x(-t)$ would be a different solution to the same initial value problem.

Since $L$ is also $T$-periodic, it follows that

$$L(T/2 + t) = -L(-T/2 - t) = -L(T/2 - t)$$

and the argument above showing $x(t)$ is even also shows $x$ is "even about $T/2$: $x(T/2 + t) = x(T/2 - t)$. Consequently,

$$x(T/2 + t) = x(T/2 - t) = x(t - T/2)$$

whence $x$ is $T$-periodic. \hfill \Box

To show that $L(\theta)$ has the required properties to apply the lemma, recall that $f_1$ and $f_2$ are even and $2\pi$-periodic. Consequently, the same is true about the functions $z$, $R_p$, $R_m$ and $K$, while $a_1$ and $a_2$ are odd and $2\pi$-periodic. Taking this into account, and in view of (7.8), we conclude that $L$ is odd. Finally, the entries of $L$ are seen to be smooth and vanish at $\theta = n\pi$ by using again (7.8). For this matter note that these points correspond to the poles where, as a function of $\theta$, $R_p - R_m$ vanishes to second order (since both $R_p$ and $R_m$ are even) and $a_1$ and $a_2$ also vanish. \hfill \Box

The differentials of the gauge momenta of the proposition above become dependent along the points of $D^s$ where the lifted $G$-action is not free. As will be explained below, these points correspond to a special kind of relative equilibrium.

The 3-form $\Lambda$ given by (5.3) that defines the desired gauge transformation can be computed without explicitly solving (7.7). We will first find its expression on our chart and then give its expression on all of $Q$. Let $(g(\theta), h(\theta))$, be a solution of (7.7) with $h$ not identically zero. Then $\{Z_1, Y_2, Y_3\}$ is an adapted basis at all points where $h(\theta) \neq 0$, where we have put $Y_2 = W_1$. A simple calculation yields

$$C_{123} = \langle [Z_1, Y_2], Y_3 \rangle = -mh(\theta)z(\theta)a_1(\theta).$$

We consider the dual basis $\{\mu^{\alpha}\}$ of $\{Z_1, Y_2, Y_3\}$ given by

$$\mu^1 = \frac{1}{h(\theta)} d\psi, \quad \mu^2 = d\varphi - \frac{g(\theta)}{h(\theta)} d\psi, \quad \mu^3 = d\theta.$$
As for the Chaplygin sphere, this amounts to identifying $D^* = W^\circ$ where $W = \text{span}\{\partial_x, \partial_y\}$. The 3-form $\Lambda$ defined by (5.3) is unique since $r - \ell < 3$ (item (ii) of Remark 5.5). It is given in our coordinates by

$$\Lambda = -mz(\theta)R_p(\theta)\sin \theta \, d\phi \wedge d\theta \wedge d\psi,$$

where we have written $a_1(\theta) = R_p(\theta)\sin \theta$. This 3-form may be written in terms of the invariant 1-forms $\lambda^\alpha$ for $SO(3)$ defined by (6.6) as

$$\Lambda = mz(\gamma_1)R_p(\gamma_3)\lambda^1 \wedge \lambda^2 \wedge \lambda^3.$$ 

Two facts about $\Lambda$ should be remarked at this point. Firstly, $\Lambda$ admits a smooth extension to all of $Q$ including configurations where the action is not free. Secondly, $\Lambda$ is independent of our choice of solution of the system (7.7). Therefore, any gauge momentum of the system will be a Casimir function of the bracket that the corresponding bivector $\Pi_{nh}^\lambda$ induces on the quotient space $D^*/G$.

The explicit expressions for the bracket $\Pi_{nh}^\lambda$ on $D^*$ may be obtained working with the basis of sections $\{W_1, W_2, Z_3\}$ and using the formulae in Section 3. This is analogous to what was done in Section 6 for the Chaplygin ball. We do not give the details of this calculation. Instead we give expressions for the brackets of the entries of $\gamma$ and of the angular momentum about the contact point

$$M = I\Omega + m\rho \times (\Omega \times \rho).$$

We have

$$\{\gamma, \gamma\} = 0, \quad \{M, \gamma\} = -\varepsilon_{ijk}\gamma_k,$$

$$\{M_i, M_j\} = \varepsilon_{ijk}\left(-M_k + mR_m(\gamma_3)(\Omega, \gamma)\rho_k + \frac{m(R_p(\gamma_3) - R_m(\gamma_3))(\gamma_3)(\Omega, \gamma)\rho_k + T_k)}{K(\gamma_3)}\right),$$

where

$$T_k = \frac{I_3(M_1\gamma_1 + M_2\gamma_2)\gamma_k}{1 - \gamma_3^2}, \quad k = 1, 2, \quad T_3 = I_3M_3.$$ 

(Notice that the bracket has no singularity at the poles $\gamma_3 = \pm 1$ since $R_p - R_m$ vanishes there). In these expressions, we think that $\Omega$ is written in terms of $M$ and $\gamma$ as

$$\Omega = AM + m\frac{(AM, \rho)}{1 - m(\rho, \rho)}A\rho,$$

where $A := (I + m||\rho||^2)^{-1}$.

The reduction of the system by $G = SE(2) \times S^1$ can be performed in two steps since the individual actions of $SE(2)$ and $S^1$ commute. Analogous to the case of the Chaplygin sphere, the orbit space $D^*/SE(2)$ is smooth and isomorphic to $S^2 \times \mathbb{R}^3$, and points in this space are labeled by the pair $\gamma, M$. The formulae (7.9) can be interpreted as the reduction of the bracket $\Pi_{nh}^\lambda$ to $D^*/SE(2)$. This partially reduced bracket does not satisfy the Jacobi identity except for very particular cases, like a perfectly homogeneous sphere. In more general cases, like the so-called Routh sphere where the body is spherical but the centre of mass does not coincide with the geometric centre, its characteristic distribution is not even integrable.

As indicated in [14], the dynamics in this intermediate reduced space is given by

$$\dot{M} = M \times \Omega + m\dot{\rho} \times (\Omega \times \rho) + \frac{dV}{d\gamma_3}(\gamma \times E_3), \quad \gamma = \gamma \times \Omega.$$ 

The above equations are Hamiltonian with respect to the bracket (7.9) and the Hamiltonian $H(M, \gamma) = \frac{1}{2}(M, \Omega) + V(\gamma_3).$
The ultimate reduction of the system is achieved by noticing that the action of $S^1$ on the orbit space $D^*/SE(2)$ is by simultaneous rotation on the planes $\gamma_1, \gamma_2$ and $M_1, M_2$. Note that the points having $\gamma_1 = \gamma_2 = 0$ and $M_1 = M_2 = 0$ are fixed by the action. The locus of these points corresponds to relative equilibria where the body of revolution is steadily spinning about its axis of symmetry touching the plane at one of the poles. Along these points, the differentials of any two gauge momenta of the system are dependent.

In order to see how the gauge momenta descend to Casimir functions on the ultimately reduced space $D^*/G = (D^*/SE(2))/S^1$, we note that the linear functions on $D^*$ generated by the vector fields $W_1$ and $W_2$ may be written as

$$p_{W_1} = (M, \gamma), \quad p_{W_2} = M_3.$$  

It follows that the gauge momenta of the system are of the form

$$C = g(\gamma_3)(M, \gamma) + k(\gamma_3)M_3,$$

where $g$ and $k$ are solutions to the system (7.7), expressed as functions of $\gamma_3 = \cos \theta$ (this is possible since the solutions to this system are periodic and even functions of $\theta$ as shown in Proposition 7.1).

Using the differential equation satisfied by $g, k$ one can show that the Hamiltonian vector field of $C$ is

$$X_C = k(\gamma_3)(\gamma_2 \partial_{\gamma_2} - \gamma_1 \partial_{\gamma_1} + M_2 \partial_{M_2} - M_1 \partial_{M_1}),$$

which is clearly vertical with respect to the action of $S^1$ defined above.

The ultimate reduced space $D^*/G = (D^*/SE(2))/S^1$ can be described by introducing generators of the ring of $S^1$-invariant polynomials on $D^*/SE(2)$. For example

$$\sigma_1 = \gamma_3, \quad \sigma_2 = \gamma_1 M_2 - \gamma_2 M_1, \quad \sigma_3 = \gamma_1 M_1 + \gamma_2 M_2, \quad \sigma_4 = M_3, \quad \sigma_5 = M_1^2 + M_2^2.$$  

These functions identically satisfy

$$\sigma_2^2 + \sigma_3^2 = \sigma_5 (1 - \sigma_1^2), \quad \sigma_5 \geq 0.$$  

The reduced space $D^*/G$ is then isomorphic to the four dimensional, semi-algebraic variety $M \subset \mathbb{R}^5$ defined by

$$M := \{ \sigma \in \mathbb{R}^5 : \sigma_2^2 + \sigma_3^2 = \sigma_5 (1 - \sigma_1^2), \quad \sigma_5 \geq 0 \}.$$  

This space is not smooth having singularities along the two lines

$$L^\pm = \{ \sigma \in \mathbb{R}^5 : \sigma = (\pm 1, 0, 0, \sigma_4, 0) \}$$

that correspond to the relative equilibria mentioned above. Each of these lines is a one dimensional stratum of $M$.

By $G$-invariance of $\Pi_{nh}$, there is an induced bracket $\{ \cdot, \cdot \}_M$ on $M$ having

$$C_j(\sigma) = g_j(\sigma_1) \sigma_1 + (g_j(\sigma_1) \sigma_1 + k_j(\sigma_1)) \sigma_4, \quad j = 1, 2,$$

as Casimir functions. Here $(g_j, k_j)$, $j = 1, 2$, are two linearly independent solutions of (7.7) written as functions of $\sigma_1 = \gamma_3 = \cos \theta$. These Casimir functions are independent everywhere on $M$ but their differentials are linearly dependent along the singular strata $L^\pm$. The Hamiltonian $H$ can be written in terms of $\sigma$ as

$$H(\sigma) = \frac{1}{2}\left( \frac{\sigma_5}{K_1(\sigma_1)} + \frac{\sigma_4^2}{K_3(\sigma_1)} \right) + m \frac{(\sigma_3 f_1(\sigma_1) K_3(\sigma_1) + \sigma_4 f_2(\sigma_1) K_1(\sigma_1))^2}{K(\sigma_1) K_1(\sigma_1) K_3(\sigma_1)}$$

where $K_j(\sigma_1) := I_j + m(1 - \sigma_1^2) f_1(\sigma_1)^2 + m f_2(\sigma_1)^2$, $j = 1, 3$. The ultimately reduced equations of motion can be formulated as

$$\sigma_j = \{ \sigma_j, H \}_M, \quad j = 1, \ldots, 5.$$  

(7.10)
The crucial point of our construction is that the above equations are true Hamiltonian. Namely, the bracket \{\cdot,\cdot\}_M satisfies the Jacobi identity. This follows immediately from Corollary 5.8.

Equations (2.8) in [14] give explicit expressions for the bracket \(\chi(\sigma_i)\{\cdot,\cdot\}_M\) in terms of a different family of \(S^1\)-invariant functions on \(D^*/SE(2)\), where the function \(\chi(\sigma_i) = \frac{\delta_i}{1-\sigma_i^2}\). As was already pointed out in [45], it is not necessary to introduce this conformal factor in order to satisfy the Jacobi identity. There is no explanation in [14] about the origin of this bracket, and it is likely that it was found by the authors using an ad hoc approach.

The restriction of the system to the 2-dimensional symplectic leaf determined as the level set of the Casimir functions, defines a one degree of freedom, and hence integrable, Hamiltonian system. Apparently (see [14]) the reduction of the integration of the reduced system to a set of 2 linear ODE’s was known to Chaplygin. More details about the explicit integration of the reduced system can be found in [14] or [19]. We simply mention that the generic solutions are periodic.

We stress that our approach for the reduction of the system follows the philosophy and treatment in [19] but with a fundamental difference: we are performing the Poisson reduction of the system with respect to the bracket \(\Pi^A_{nh}\) and not with respect to \(\Pi_{nh}\). By following the reduction of \(\Pi_{nh}\), the authors of [19] arrive at an equation ((180) in their text) analogous to (7.10) but with respect to a bracket of functions that does not satisfy the Jacobi identity. The authors do seem to notice that for fixed values of the integrals \(C_1, C_2\) one has a one-degree of freedom Hamiltonian system (section 6.3.7.4), but they do not provide a link between this observation and their reduction.

**Appendix A. Proof of Lemma 5.4**

The proof of Theorem 5.6 depends in a crucial way on the existence of the 3-form \(\Lambda\) that satisfies the conditions of Lemma 5.4. Here we present the a proof that such a \(\Lambda\) always exists. As explained at several points of the text, we will only prove that \(\Lambda\) is well defined on the open dense subset \(Q_f \subset Q\) where the \(G\)-action is free.

Recall the assumption made at the beginning of Section 5.2: the chosen subbundle \(W\) of \(TQ\) with the property \(TQ = D \oplus W\) is \(G\)-invariant.

A.1. **Local considerations.** Consider an adapted basis \(\{X_\alpha\}\) of sections of \(D\). Recall that this means that all vector fields \(X_\alpha\) are equivariant sections of \(D\) and that \(X_\alpha\) are gauge momentum generators for \(b = 1, \ldots, \ell\). That is, \(X_\alpha\) is a section of the distribution \(\mathcal{B}\) on \(Q\) defined point-wise as \(\mathcal{B}_q = D_q \cap (g \cdot q)\), and satisfies \(\tilde{p}_{X_\alpha} = 0\). In virtue of Lemma 5.2, we may assume that our adapted basis of sections is defined on a \(G\)-invariant open subset \(U \subset Q_f\).

In the above paragraph, as in the remainder of this section, we do not distinguish \(Z_\alpha\) and \(Y_{\alpha}\) as in Definition 5.1, as there would be too much notation—we rely on the indices to distinguish the type of generator. Recall that our convention is \(b,c,\ldots \in \{1,\ldots, \ell\}\) (corresponding to gauge momentum generators), \(I,J,\ldots \in \{1,\ldots, \ell\}\) and \(\alpha,\beta,\ldots \in \{1,\ldots, r\}\), where \(r\) is the rank of \(D\).

A 3-form \(\Lambda\) satisfying conditions (i) and (ii) of Lemma 5.4 may be defined on \(U\) by the formula

\[
\Lambda = \frac{1}{6} B_{\alpha\beta\gamma} \mu^\alpha \wedge \mu^\beta \wedge \mu^\gamma, \tag{A.1}
\]

where the coefficients \(B_{\alpha\beta\gamma}\) are any \(G\)-invariant functions, alternating in the indices and satisfying

\[
B_{\alpha\beta\gamma} = \langle [X_\alpha, X_\beta], X_\gamma \rangle. \tag{A.2}
\]

Here \(B_{IJK}\) are arbitrary smooth \(G\)-invariant functions, alternating in the indices. Note that the skew-symmetry condition \(B_{\alpha\beta\gamma} = -B_{\beta\alpha\gamma}\) holds in virtue of Theorem 5.3. As usual, \(\{\mu^\alpha\}\) in (A.1) is the dual basis of \(\{X_\alpha\}\). Namely, they are 1-forms defined on \(U\) that annihilate \(W\) and satisfy \(\mu^\alpha(X_\beta) = \delta^\alpha_\beta\). Moreover, they are \(G\)-invariant by \(G\)-invariance of \(W\), and hence \(\Lambda\) is also \(G\)-invariant.
Now let $V$ be a possibly different $G$-invariant open subset of $Q_f$, and let $\{Y_\alpha\}$ be a (new) adapted basis of $D$ on $V$, with $Y_\alpha = M_\alpha^\beta X_\beta$ on the intersection $U \cap V$. In order to be gauge momentum generators, we require $Y_\alpha = M_\alpha^\beta X_\beta$, with $M_\alpha^\beta \in \mathbb{R}$ (constants), and in order that the $Y_\beta$ are equivariant we require that all the coefficients $M_\alpha^\beta$ be $G$-invariant functions.

(Here we are assuming the distribution is strongly nonholonomic on the configuration space, otherwise the $M_\alpha^\beta$ are only annihilated by every vector field tangent to the distribution $D$; the proof in the more general case proceeds in the same way.)

The dual basis $\{\mu^\alpha\}$ transforms into a new basis $\{\nu^\beta\}$ dual to the $Y_\beta$ satisfying $\nu^\beta = M_\alpha^\beta \mu^\alpha$, where $\bar{M}$ is the inverse matrix of $M$; that is, $\bar{M}_\alpha^\beta M_\beta^\gamma = \delta^\alpha_\gamma$ (Kronecker $\delta$). Let us emphasise that

$$M_\alpha^\beta = \bar{M}_\alpha^\beta = 0, \quad \forall b, I.$$  \hfill (A.3)

We define

$$\Lambda' = \frac{1}{6} B_{\alpha\beta\gamma}' \nu^\alpha \wedge \nu^\beta \wedge \nu^\gamma,$$

where the $G$-invariant alternating coefficients $B_{\alpha\beta\gamma}'$ are related by $\Lambda$. Namely, $B_{\alpha\beta\gamma}'$ is given by (A.2), using the $Y_\alpha$ in place of the $X_\alpha$, and $B_{\alpha\beta\gamma}'$ are arbitrary.

We wish to compare $\Lambda$ and $\Lambda'$ on the intersection $U \cap V$.

**Lemma A.1.** The coefficients $B_{\alpha\beta\gamma}$ and $B_{\alpha\beta\gamma}'$ are related by

$$B_{\beta\gamma} = M_\alpha^\beta M_\beta^\gamma M_\gamma^\rho B_{\rho\sigma\tau}$$  \hfill (A.4)

$$B_{\alpha\beta\gamma}' = M_\alpha^\rho M_\beta^\sigma M_\gamma^\tau B_{\rho\sigma\tau} + E_{\alpha\beta\gamma},$$  \hfill (A.5)

for some $G$-invariant functions $E_{\alpha\beta\gamma}$ on $Q$, alternating in the indices. The forms themselves are related more simply by,

$$\nu^\alpha \wedge \nu^\beta \wedge \nu^\gamma = M_\alpha^\rho M_\beta^\sigma M_\gamma^\tau \mu^\delta \wedge \mu^\epsilon \wedge \mu^\eta.$$  \hfill (A.6)

Note that since $M_\alpha^\beta = 0$ (A.3), all the non-zero terms on the right-hand side of (A.4) involve only coefficients of the form $B_{\sigma\tau\gamma}$, and not the $B_{\alpha\beta\gamma}'$.

**Proof.** The expression (A.6) follows immediately from the relation $\nu^\alpha = M_\alpha^\rho \mu^\rho$ given above.

For the $B'$ coefficients, expand $B'$ in terms of the $X_\alpha$:

$$B_{\beta\gamma} = \langle [Y_\beta, Y_\gamma], Y_\tau \rangle = \langle M_\alpha^\beta X_\alpha, M_\gamma^\tau X_\tau \rangle = M_\alpha^\beta M_\gamma^\tau B_{\sigma\tau\gamma} + M_\alpha^\beta M_\gamma^\tau (M_\sigma^\rho X_\sigma) (X_\tau) - M_\alpha^\beta M_\gamma^\tau (M_\sigma^\rho X_\sigma) (X_\tau) \langle X_\tau, X_\tau \rangle.$$

However, the final two terms both vanish because firstly $X_\alpha$ is tangent to the group orbit and $M_\sigma^\rho$ is invariant (so constant on group orbits), and secondly $M_\tau^\beta$ is constant. Thus, $B_{\alpha\beta\gamma}' = M_\alpha^\rho M_\beta^\sigma M_\gamma^\tau B_{\sigma\tau\gamma}$. The first equation (A.4) then follows from (A.3).

Equation (A.5) can be taken as a definition of $E_{\alpha\beta\gamma}$, where the invariance and the alternating structure is clear. \hfill $_\Box$

Recall that for each $q$ the subspace $S_q \subset T_q Q$ is defined to be $S_q = D_q \cap g \cdot q$. We also write $S_0 \subset S$ to be the sub-distribution spanned by the generators of the gauge momenta, which is therefore assumed to be of constant rank. Recall that $W$ is the distribution complementary to $D$ on which all $\mu^\alpha$ vanish.
**Proposition A.2.** On $U \cap V$, the difference $\Psi := \Lambda - \Lambda'$ is an invariant 3-form annihilating $S_0 \oplus W$, so is of the form

$$\Psi = E_{IJK} \mu^I \wedge \mu^J \wedge \mu^K.$$

**Proof.** Using the lemma above,

$$\Lambda' = \frac{1}{6} B_{\alpha \beta \gamma} V^\alpha \wedge V^\beta \wedge V^\gamma + \sum_{\sigma \tau} \Lambda_\sigma^I M^\sigma_R M^\tau_S M^I_R T^K - \frac{1}{6} \delta_\sigma^\rho \delta_\tau^\sigma B_{\rho \sigma \tau} \mu^\rho \wedge \mu^\sigma \wedge \mu^\tau + E_{IJK} M^I_R M^J_S M^K_T \mu^R \wedge \mu^S \wedge \mu^T$$

as required (the final step involves relabelling $RST$ to $IJK$).

**A.2. Global considerations.** As was pointed out in Remark 5.5, if $k - \ell < 3$ then $\Lambda = \Lambda'$ and the 3-form is uniquely defined. There remains the question of whether when $k - \ell \geq 3$, the locally-defined 3-forms $\Lambda$ can be chosen to agree everywhere, by choosing the $B_{IJK}$ suitably. The answer is yes, by a standard partition of unity argument used in Čech cohomology, as follows.

**Proposition A.3.** Let $\mathcal{U} = \{U_i\}$ be a cover of $Q_f$ by G-invariant open sets, and on each $U_i$ suppose a 3-form $\Lambda_i$ is selected, of the form (A.1). There exist G-invariant 3-forms $\Psi_i$ on $U_i$ annihilating $S_0 \oplus W$, such that the forms

$$\Lambda_i := \Lambda_i + \Psi_i$$

define a global 3-form on $Q$: that is on each intersection $U_i \cap U_j$, $\Lambda_i = \Lambda_j$.

**Proof.** Let $\{\phi_i\}$ be a partition of unity subordinate to the cover $\mathcal{U}$, by G-invariant functions. On each $U_i \cap U_j$ let $\Psi_{ij} = \Lambda_i - \Lambda_j$. By Proposition A.2, the $\Psi_{ij}$ are invariant forms annihilating $S_0 \oplus W$, and they clearly satisfy the cocycle condition,

$$\Psi_{ij} + \Psi_{jk} + \Psi_{ki} = 0, \quad \text{wherever } U_i \cap U_j \cap U_k \neq \emptyset.$$

Now, for each $i$ define

$$\Psi_i = \sum_k \phi_k \Psi_{ik}.$$

(Note that $\Psi_{ii} = 0$, and the sum is over all $k$.) Then with $\Lambda_i = \Lambda_i + \Psi_i$

$$\Lambda_i - \Lambda_j = \Lambda_i - \Lambda_j + \Psi_i - \Psi_j = \Psi_{ij} + \sum_k \phi_k (\Psi_{ik} - \Psi_{jk})$$

as required. 

**Funding:** This research was made possible by a Newton Advanced Fellowship from the Royal Society, ref: NA140017.

**Conflict of Interest:** The authors declare that they have no conflict of interest.
REFERENCES

[1] AGOSTINELLI, C., Nuova forma sintetica delle equazioni del moto di un sistema anolonomo ed esistenza di un integrale lineare nelle velocita, *Boll. Un. Mat. Ital.*, 11 (1956), 1–9.

[2] APPELL, P., Sur l’intégration des équations du mouvement d’un corps pesant de révolution roulant par une arête circulaire sur un plan horizontal; cas particulier du cercle. *Rendiconti del circolo matematico di Palermo*, 14, (1900), 1–6.

[3] BALSEIRO, P. AND L.C. GARCÍA-NARANJ O, Gauge transformations, twisted Poisson brackets and Hamiltonization of nonholonomic systems. *Arch. Rat. Mech. Anal.*, 205 (2012), no. 1, 267–310.

[4] BALSEIRO, P., The Jacobiator of nonholonomic systems and the geometry of reduced nonholonomic brackets. *Arch. Ration. Mech. Anal.*, 214 (2014), 453–501.

[5] BALSEIRO, P. AND N. SANSONETTO, A geometric characterisation of certain first integrals for nonholonomic systems with symmetries. *SIGMA* 12 (2016), 14 pages.

[6] BALSEIRO, P., Hamiltonization of solids of revolution through reduction. *J. Nonlinear Sci.* https://doi.org/10.1007/s00332-017-9394-1, (2017).

[7] P. BALSEIRO, P., AND L.P. YAPU-QUISPE, in preparation.

[8] BATES, L., H. GRAUMANN AND C. MACDONnell, Examples of gauge transformation laws in nonholonomic systems *Rep. Math. Phys.*, 37 (1996), 295–308.

[9] BATES L., AND J. SNATYCKI, Nonholonomic reduction, *Rep. Math. Phys.*, 32, (1993) 99–115.

[10] BIZYAEv, I. A., A. BOLSINOv, A. BORISOv. AND I. MAMAEv, Topology and bifurcations in nonholonomic mechanics. Internat. J. Bifur. Chaos Appl. Sci. Engrg. 25 (2015), no. 10, 1530028, 21 pp.

[11] BLOCH, A.M., P.S. KRISHNAPRASAD, J.E. MARSDEN AND R.M. MURRAY, Nonholonomic mechanical systems with symmetry. *Arch. Ration. Mech. Anal.*, 136 (1996), 21–99.

[12] BLOCH, A.M., J.E. MARSDEN AND D.V. ZENkOV, Quasivelocities and symmetries in non-holonomic systems. *Dyn. Syst.* 24 (2009), 187–222.

[13] BORISOv A. V. AND I. S. MAMAEv, Chaplygin’s Ball Rolling Problem Is Hamiltonian. *Math. Notes*, (2001), 70, 793–795.

[14] BORISOv A.V. AND I.S. MAMAEv, Rolling of a rigid body on a plane and sphere. Hierarchy of dynamics. *Regul. Chaotic Dyn.* 7 (2002), 177–200.

[15] BORISOV A. V., I. S. MAMAEv AND A. A. KILIN, Rolling of a ball on a surface. New integrals and hierarchy of dynamics. *Regul. Chaotic Dyn.*, 7 (2002), 201–219.

[16] BORISOV A.V. AND I.S. MAMAEv, Isomorphism and Hamilton Representation of Some Non-holonomic Systems, *Siberian Math. J.*, 48 (2007), 33–45 See also: arXiv: nl-nlin.SI/0509036 v. 1 (Sept. 21, 2005).

[17] BORISOV A.V. AND I. S. MAMAEv, Symmetries and Reduction in Nonholonomic Mechanics, *Regul. Chaotic Dyn.*, 20 (2015), 553–604.

[18] CANTRIN, F., M. DE LEÓN, AND D. MARTÍN DE DIEGO, On almost-Poisson structures in nonholonomic mechanics. *Nonlinearity*, 12, (1999) 721–737.

[19] CUSHMAN R., J.J. DUISTERMAAT AND J. ŠNATYCKI, Geometry of nonholonomically constrained systems. Advanced Series in Nonlinear Dynamics, 26. World Scientific Publishing, 2010.

[20] CHAPLYGIN, S.A., On a ball’s rolling on a horizontal plane. *Regul. and Chaotic Dyn.*, 7, (2002), 131–148; [original paper in *Mathematical Collection of the Moscow Mathematical Society*, 24 (1903), 139–168]

[21] CHAPLYGIN, S.A., On the theory of the motion of nonholonomic systems. The reducing-multiplier Theorem. *Regul. Chaotic Dyn.*, 13, 369–376 (2008) [Translated from Matematicheskii sbornik (Russian) 28 (1911), by A. V. Getling]

[22] CHAPLYGIN, S.A., On a motion of a heavy body of revolution on a horizontal plane. *Collected works. Vol. I. Theoretical mechanics. Mathematics (Russian)*, 51–57, Gos. Izd. Tekhn.-Teoret. Lit., Moscow, 1948 [see MR0052352 (14,609i)]. English translation in: *Regul. Chaotic Dyn.*, 7 (2002), no. 2, 119–130.

[23] DUISTERMAAT, J.J. AND J.A.C. KOLK, *Lie Groups*, Springer (2000).

[24] EHLERS, K., J. KOILLER, R. MONTGOMERY AND P. M. RIOS, Nonholonomic Systems via Moving Frames: Cartan Equivalence and Chaplygin Hamiltonization. in *The breath of Symplectic and Poisson Geometry*, Progress in Mathematics Vol. 232, (2004), 75–120.

[25] FASSO, F., A. GIacobbe AND N. SANSONETTO, Periodic flows, rank-two Poisson structures, and nonholonomic mechanics. *Regul. Chaotic Dyn.*, 10 (2005), 267–284.

[26] FASSO, F., A. RAMOS AND N. SANSONETTO, The reaction annihilator distribution and the nonholonomic Noether theorem for lifted actions, *Regul. Chaot. Dyn.*, 12 (2007), 579–588.
measure and flows on homogeneous spaces. J. Nonlinear Sci. Transl. (2) 3 (2010), 1343–1355.

[30] FEDOROV, Y. N., and V.V. KOZLOV, Various aspects of n-dimensional rigid body dynamics. Amer. Math. Soc. Transl. (2), 168 (1995), 141–171.

[31] FEDOROV, Y. N. and B. JOVANOVIĆ, Nonholonomic LR systems as generalized Chaplygin systems with an invariant measure and flows on homogeneous spaces. J. Nonlinear Sci. 14 (2004), 341–381.

[32] GARCÍA-NARANJIO, L.C., Almost Poisson brackets for nonholonomic systems on Lie groups, Ph.D. dissertation, University of Arizona, 2007

[33] GARCÍA-NARANJIO, L.C., Reduction of almost Poisson brackets and Hamiltonization of the Chaplygin sphere. Discrete Contin. Dyn. Syst. Ser. S 3 (2010), 37–60.

[34] GRABOWSKI, J., DE LEÓN, M., MARRERO, J. C. and D. MARTÍN DE DIEGO, Nonholonomic constraints: a new viewpoint. J. Math. Phys. 50 (2009), 013520, 17 pp.

[35] IBORT, A., M. DE LEÓN, J. C. MARRERO, and D. MARTÍN DE DIEGO, Dirac Brackets In Constrained Dynamics, Fortschr. Phys. 47 (1999), 459–492.

[36] ILIEV IL. and KHR. SEMERDZHIEV, Relations between the first integrals of a nonholonomic mechanical system and of the corresponding system freed of constraints. J. Appl. Math. Mech., 36 (1972), 381–388.

[37] ILIEV, I.L., On first integrals of a nonholonomic mechanical system, J. Appl. Math. Mech., 39 (1975), 147–150.

[38] JOST, J., Riemannian Geometry and Geometric Analysis. Springer, 2000.

[39] JOTZ, M. and T. S. RATIU, Dirac structures, nonholonomic systems and reduction. Rep. Math. Phys. 69 (2012), 5–56.

[40] JOVANOVIĆ, B., Hamiltonization and integrability of the Chaplygin sphere in $\mathbb{R}^n$. J. Nonlinear Sci. 20 (2010), 569–593.

[41] KOHLER, J., Reduction of some classical nonholonomic systems with symmetry. Arch. Ration. Mech. Anal. 118 (1992), 113–148.

[42] DE LEÓN, M., MARRERO, J. C. and D. MARTÍN DE DIEGO, Linear almost Poisson structures and Hamilton-Jacobi equation. Applications to nonholonomic mechanics. J. Geom. Mech. 2 (2010), 159–198.

[43] MARLE, CH. M., Various approaches to conservative and nonconservative nonholonomic systems. Proc. Workshop on Non-Holonomic Constraints in Dynamics (Calgary, August 26-29, 1997) Rep. Math. Phys. 42 (1998) 211–229.

[44] MARSDEN J.E. and T.S. RATIU, Introduction to Mechanics with symmetry, Texts in Applied Mathematics 17 Springer-Verlag 1994.

[45] RAMOS, A., Poisson structures for reduced non-holonomic systems. J. Phys. A 37 (2004), 4821–4842.

[46] ROUTH, E.D., Dynamics of a system of rigid bodies, 7th ed., revised and enlarged. Dover Publications, Inc., New York 1960.

[47] SEVERA, P. and A. WEINSTEIN, Poisson geometry with a 3-form background. Noncommutative geometry and string theory (Yokohama, 2001). Progr. Theoret. Phys. Suppl. No. 144 (2001), 145–154.

[48] TSIGANOV, A., Poisson structures for two nonholonomic systems with partially reduced symmetries. J. Geom. Mech. 6 (2014), no. 3, 417–440.

[49] VAN DER SCHAFT, A.J., and B.M. MASCHKE, On the Hamiltonian formulation of nonholonomic mechanical systems. Reports on Math. Phys. 34 (1994), 225–233.

[50] WEBER, R.W., Hamiltonian systems with constraints and their meaning in mechanics. Arch. Ration. Mech. Anal. 91 (1986), 309–335.

E-mail address: luis@mym.iimas.unam.mx, j.montaldi@manchester.ac.uk

LGN: IIMAS, UNAM, Mexico City, Mexico. JM: SCHOOL OF MATHEMATICS, UNIVERSITY OF MANCHESTER, UK.