

HAMILTON–JACOBI THEORY FOR DEGENERATE LAGRANGIAN SYSTEMS WITH HOLONOMIC AND NONHOLONOMIC CONSTRAINTS

MELVIN LEOK, TOMOKI OHSAWA, AND DIANA SOSA

ABSTRACT. We extend Hamilton–Jacobi theory to Lagrange–Dirac (or implicit Lagrangian) systems, a generalized formulation of Lagrangian mechanics that can incorporate degenerate Lagrangians as well as holonomic and nonholonomic constraints. We refer to the generalized Hamilton–Jacobi equation as the *Dirac–Hamilton–Jacobi equation*. For non-degenerate Lagrangian systems with nonholonomic constraints, the theory specializes to the recently developed nonholonomic Hamilton–Jacobi theory. We are particularly interested in applications to a certain class of degenerate nonholonomic Lagrangian systems with symmetries, which we refer to as *weakly degenerate Chaplygin systems*, that arise as simplified models of nonholonomic mechanical systems; these systems are shown to reduce to non-degenerate almost Hamiltonian systems, i.e., generalized Hamiltonian systems defined with non-closed two-forms. Accordingly, the Dirac–Hamilton–Jacobi equation reduces to a variant of the nonholonomic Hamilton–Jacobi equation associated with the reduced system. We illustrate through a few examples how the Dirac–Hamilton–Jacobi equation can be used to exactly integrate the equations of motion.

1. INTRODUCTION

1.1. Degenerate Lagrangian Systems and Lagrange–Dirac Systems. Degenerate Lagrangian systems are the motivation behind the work of Dirac [16, 17, 18] on constrained systems, where degeneracy of Lagrangians imposes constraints on the phase space variables. The theory gives a prescription for writing such systems as Hamiltonian systems, and is used extensively for gauge systems and their quantization (see, e.g., Henneaux and Teitelboim [26]).

Dirac’s theory of constraints was geometrized by Gotay et al. [25] (see also Gotay and Nester [22, 23, 24] and Künzle [35]) to yield a constraint algorithm to identify the solvability condition for presymplectic systems and also to establish the equivalence between Lagrangian and Hamiltonian descriptions of degenerate Lagrangian systems. The algorithm is extended by de León and Martín de Diego [14] to degenerate Lagrangian systems with nonholonomic constraints.

On the other hand, Lagrange–Dirac (or implicit Lagrangian) systems of Yoshimura and Marsden [52, 53] provide a rather direct way of describing degenerate Lagrangian systems that do not explicitly involve constraint algorithms. Moreover, the Lagrange–Dirac formulation can address more general constraints, particularly nonholonomic constraints, by directly encoding them in terms of Dirac structures, as opposed to symplectic or Poisson structures.

1.2. Hamilton–Jacobi Theory for Constrained Degenerate Lagrangian Systems. The goal of this paper is to generalize Hamilton–Jacobi theory to Lagrange–Dirac systems. The challenge in doing so is to generalize the theory to simultaneously address degeneracy and nonholonomic constraints. For degenerate Lagrangian systems, some work has been done, built on Dirac’s theory of constraints, on extending Hamilton–Jacobi theory (see, e.g., Henneaux and Teitelboim [26, Section 5.4] and Rothe and Scholtz [44]) as well as from the geometric point of view by Cariñena et al. [8]. For nonholonomic systems, Iglesias-Ponte et al. [28] generalized the geometric Hamilton–Jacobi theorem (see Abraham and Marsden [1, Theorem 5.2.4]) to nonholonomic systems, which has been studied further by de León et al. [15], Ohsawa and Bloch [42], Cariñena et al. [9], and Ohsawa et al. [43]. However, to the authors’ knowledge, no work has been done that can deal with both degeneracy and nonholonomic constraints.

2010 *Mathematics Subject Classification.* 37J60, 70H20, 70H45, 70F25.

Key words and phrases. Hamiltonian–Jacobi theory, Lagrange–Dirac systems, Dirac structures, degenerate Lagrangian systems, nonholonomic constraints.

1.3. Applications to Degenerate Lagrangian Systems with Nonholonomic Constraints. We are particularly interested in applications to degenerate Lagrangian systems with nonholonomic constraints. Such systems arise regularly, in practice, as model reductions of multiscale systems: For example, consider a nonholonomic mechanical system consisting of rigid bodies, some of which are significantly lighter than the rest. Then, one can make an assumption that the light parts are massless for the sake of simplicity; this often results in a degenerate Lagrangian. While naïvely making a massless approximation usually leads to unphysical results¹, a certain class of nonholonomic systems seem to allow massless approximations without such inconsistencies. See, for example, the modelling of a bicycle in Getz [20] and Getz and Marsden [21] (see also Koon and Marsden [33] and Example 3.5 of the present paper).

1.4. Outline. We first briefly review Dirac structures and Lagrange–Dirac systems in Section 2. Section 3 introduces a class of degenerate nonholonomic Lagrangian systems with symmetries that reduce to non-degenerate Lagrangian systems after symmetry reduction; we call them *weakly degenerate Chaplygin systems*. Section 4 gives Hamilton–Jacobi theory for Lagrange–Dirac systems, defining the *Dirac–Hamilton–Jacobi equation*, and shows applications to degenerate Lagrangian systems with holonomic and nonholonomic constraints. We then apply the theory to weakly degenerate Chaplygin systems in Section 5; we derive a formula that relates solutions of the Dirac–Hamilton–Jacobi equations with those of the nonholonomic Hamilton–Jacobi equation for the reduced weakly degenerate Chaplygin systems. Appendix A discusses reduction of weakly degenerate Chaplygin systems by a symmetry reduction of the associated Dirac structure.

2. LAGRANGE–DIRAC SYSTEMS

Lagrange–Dirac (or implicit Lagrangian) systems are a generalization of Lagrangian mechanics to systems with (possibly) degenerate Lagrangians and constraints. Given a configuration manifold Q , a Lagrange–Dirac system is defined using a generalized Dirac structure on T^*Q , or more precisely a subbundle D of the Whitney sum $TT^*Q \oplus T^*T^*Q$.

2.1. Dirac Structures. Let us first recall the definition of a (generalized) Dirac structure on a manifold M . Let M be a manifold. Given a subbundle $D \subset TM \oplus T^*M$, the subbundle $D^\perp \subset TM \oplus T^*M$ is defined as follows:

$$D^\perp := \{(X, \alpha) \in TM \oplus T^*M \mid \langle \alpha', X \rangle + \langle \alpha, X' \rangle = 0 \text{ for any } (X', \alpha') \in D\}.$$

Definition 2.1. A subbundle $D \subset TM \oplus T^*M$ is called a *generalized Dirac structure* if $D^\perp = D$.

Note that the notion of Dirac structures, originally introduced in Courant [12], further satisfies an integrability condition, which we have omitted as it is not compatible with our interest in nonintegrable (nonholonomic) constraints. Hereafter, we refer to generalized Dirac structures as simply “Dirac structures.”

2.2. Induced Dirac Structures. Here we consider the induced Dirac structure $D_{\Delta_Q} \subset TT^*Q \oplus T^*T^*Q$ introduced in Yoshimura and Marsden [52]. See Dalsmo and van der Schaft [13] for more general Dirac structures, Bloch and Crouch [4] and van der Schaft [49] for those defined by Kirchhoff current and voltage laws, and van der Schaft [50] for applications of Dirac structures to interconnected systems.

Let Q be a smooth manifold, $\Delta_Q \subset TQ$ a regular distribution on Q , and Ω the canonical symplectic two-form on T^*Q . Denote by Δ_Q° the annihilator of Δ_Q and by $\Omega^b : TT^*Q \rightarrow T^*T^*Q$ the flat map induced by Ω . The distribution $\Delta_Q \subset TQ$ may be lifted to the distribution Δ_{T^*Q} on T^*Q defined as

$$\Delta_{T^*Q} := (T\pi_Q)^{-1}(\Delta_Q) \subset TT^*Q,$$

where $\pi_Q : T^*Q \rightarrow Q$ is the canonical projection and $T\pi_Q : TT^*Q \rightarrow TQ$ is its tangent map. Denote its annihilator by $\Delta_{T^*Q}^\circ \subset T^*T^*Q$.

Definition 2.2 (Yoshimura and Marsden [52, 53]; see also Dalsmo and van der Schaft [13]). The *induced (generalized) Dirac structure* D_{Δ_Q} on T^*Q is defined, for each $z \in T^*Q$, as

$$D_{\Delta_Q}(z) := \left\{ (v_z, \alpha_z) \in T_z T^*Q \oplus T_z^* T^*Q \mid v_z \in \Delta_{T^*Q}(z), \alpha_z - \Omega^b(z)(v_z) \in \Delta_{T^*Q}^\circ(z) \right\}.$$

¹This setting usually gives a singular perturbation problem with the small mass being its parameter. For example, for the Lagrangian $L_\varepsilon = \varepsilon \dot{x}^2/2 - x^2/2$, the Euler–Lagrange equation gives $\varepsilon \ddot{x} + x = 0$; the solution corresponding to the massless Lagrangian L_0 deviates significantly from the original solution.

If we choose local coordinates $q = (q^i)$ on an open subset U of Q and denote by $(q, \dot{q}) = (q^i, \dot{q}^i)$ (respectively, $(q, p) = (q^i, p_i)$), the corresponding local coordinates on TQ (respectively, T^*Q), then a local representation for the Dirac structure is given by

$$D_{\Delta_Q}(q, p) = \left\{ ((q, p, \dot{q}, \dot{p}), (q, p, \alpha_q, \alpha_p)) \in T_{(q,p)}T^*Q \oplus T_{(q,p)}^*T^*Q \mid \dot{q} \in \Delta_Q(q), \alpha_p = \dot{q}, \alpha_q + \dot{p} \in \Delta_Q^\circ(q) \right\}.$$

2.3. Lagrange–Dirac Systems. To define a Lagrange–Dirac system, it is necessary to introduce the Dirac differential of a Lagrangian function. Following Yoshimura and Marsden [52], let us first introduce the following maps, originally due to Tulczyjew [47, 48], between the iterated tangent and cotangent bundles.

$$\begin{array}{ccc} & \xrightarrow{\gamma_Q} & \\ T^*TQ & \xleftarrow{\kappa_Q} TT^*Q \xrightarrow{\Omega^b} & T^*T^*Q & \xleftarrow{\quad} (q, \delta q, \delta p, p) \xrightarrow{\quad} (q, p, \delta q, \delta p) \xrightarrow{\quad} (q, p, -\delta p, \delta q) \\ & & & \end{array} \quad (2.1)$$

Let $L : TQ \rightarrow \mathbb{R}$ be a Lagrangian function and let $\gamma_Q : T^*TQ \rightarrow T^*T^*Q$ be the diffeomorphism defined as $\gamma_Q := \Omega^b \circ \kappa_Q^{-1}$ (see (2.1)). Then, the Dirac differential of L is the map $\mathfrak{D}L : TQ \rightarrow T^*T^*Q$ given by

$$\mathfrak{D}L = \gamma_Q \circ dL.$$

In local coordinates,

$$\mathfrak{D}L(q, v) = \left(q, \frac{\partial L}{\partial v}, -\frac{\partial L}{\partial q}, v \right).$$

Definition 2.3. Let $L : TQ \rightarrow \mathbb{R}$ be a Lagrangian (possibly degenerate) and $\Delta_Q \subset TQ$ be a given regular constraint distribution on the configuration manifold Q . Let

$$P := \mathbb{F}L(\Delta_Q) \subset T^*Q$$

be the image of Δ_Q by the Legendre transformation and X be a (partial) vector field on T^*Q defined at points of P . Then, a *Lagrange–Dirac system* is the triple (L, Δ_Q, X) that satisfies, for each point $z \in P \subset T^*Q$,

$$(X(z), \mathfrak{D}L(u)) \in D_{\Delta_Q}(z), \quad (2.2)$$

where $u \in \Delta_Q$ such that $\mathbb{F}L(u) = z$. In local coordinates, Eq. (2.2) is written as

$$p = \frac{\partial L}{\partial v}(q, v), \quad \dot{q} \in \Delta_Q(q), \quad \dot{q} = v, \quad \dot{p} - \frac{\partial L}{\partial q}(q, v) \in \Delta_Q^\circ(q), \quad (2.3)$$

which we would like to call the *Lagrange–Dirac equations*.

Definition 2.4. A *solution curve* of a Lagrange–Dirac system (L, Δ_Q, X) is an integral curve $(q(t), p(t))$, $t_1 \leq t \leq t_2$, of X in $P \subset T^*Q$.

2.4. Lagrange–Dirac Systems on the Pontryagin Bundle $TQ \oplus T^*Q$. We may also define a Lagrange–Dirac system on $TQ \oplus T^*Q$ as well. We will use the submanifold \mathcal{K} of the Pontryagin bundle introduced in Yoshimura and Marsden [52] and the (partial) vector field \tilde{X} on $TQ \oplus T^*Q$, associated with a (partial) vector field X on T^*Q , defined in Yoshimura and Marsden [53]. Let us recall the definition of these two objects.

Given a Lagrangian $L : TQ \rightarrow \mathbb{R}$, the *generalized energy*, $\mathcal{E} : TQ \oplus T^*Q \rightarrow \mathbb{R}$, is given by

$$\mathcal{E}(q, v, p) = p \cdot v - L(q, v).$$

The submanifold \mathcal{K} is defined as the set of stationary points of $\mathcal{E}(q, v, p)$ with respect to v , with $v \in \Delta_Q(q)$. So, \mathcal{K} is represented by

$$\mathcal{K} = \left\{ (q, v, p) \in TQ \oplus T^*Q \mid v \in \Delta_Q(q), p = \frac{\partial L}{\partial v}(q, v) \right\} \quad (2.4)$$

This submanifold can also be described as the graph of the Legendre transformation restricted to the constraint distribution Δ_Q . We can also obtain the submanifold \mathcal{K} as follows. Let $\text{pr}_{TQ} : TQ \oplus T^*Q \rightarrow TQ$ be the projection to the first factor and $\pi_{TQ} : T^*TQ \rightarrow TQ$ be the cotangent bundle projection. Consider the map $\rho_{T^*TQ} : T^*TQ \rightarrow TQ \oplus T^*Q$ (see [52, Section 4.10]) which has the property that $\text{pr}_{TQ} \circ \rho_{T^*TQ} = \pi_{TQ}$; this map is defined intrinsically to be the direct sum of $\pi_{TQ} : T^*TQ \rightarrow TQ$ and $\tau_{T^*Q} \circ \kappa_Q^{-1} : T^*TQ \rightarrow T^*Q$

(see Section 4.10 in [52]), where $\tau_{T^*Q} : TT^*Q \rightarrow T^*Q$ is the tangent bundle projection. Then, we can consider the map

$$\rho_{T^*TQ} \circ dL : TQ \rightarrow TQ \oplus T^*Q,$$

whose local expression is

$$\rho_{T^*TQ} \circ dL(q, v) = \left(q, v, \frac{\partial L}{\partial v}(q, v) \right).$$

Therefore, we have

$$\mathcal{K} = \rho_{T^*TQ} \circ dL(\Delta_Q).$$

Now, given a (partial) vector field X on T^*Q defined at points of P , one can construct a (partial) vector field \tilde{X} on $TQ \oplus T^*Q$ defined at points of \mathcal{K} as follows (see Section 3.8 in [53]). For $(q, v, p) \in \mathcal{K}$, $\tilde{X}(q, v, p)$ is tangent to a curve $(q(t), v(t), p(t))$ in $TQ \oplus T^*Q$ such that $(q(0), v(0), p(0)) = (q, v, p)$ and $X(q, p)$ is tangent to the curve $(q(t), p(t))$ in T^*Q . This (partial) vector field \tilde{X} is not unique; however it has the property that, for each $x \in \mathcal{K} \subset TQ \oplus T^*Q$,

$$T\text{pr}_{T^*Q}(\tilde{X}(x)) = X(\text{pr}_{T^*Q}(x)),$$

where $\text{pr}_{T^*Q} : TQ \oplus T^*Q \rightarrow T^*Q$ is the projection to the second factor.

On the other hand, from the distribution Δ_Q on Q , we can define a distribution $\Delta_{TQ \oplus T^*Q}$ on $TQ \oplus T^*Q$ by

$$\Delta_{TQ \oplus T^*Q} = (T\text{pr}_Q)^{-1}(\Delta_Q),$$

where $\text{pr}_Q : TQ \oplus T^*Q \rightarrow Q$. Note that $\Delta_{TQ \oplus T^*Q} = (T\text{pr}_{T^*Q})^{-1}(\Delta_{T^*Q})$, since $\text{pr}_Q = \pi_Q \circ \text{pr}_{T^*Q}$. Then, as $\text{pr}_{T^*Q}^* \Omega$ is a skew-symmetric two-form on $TQ \oplus T^*Q$, we can consider the following induced (generalized) Dirac structure on $TQ \oplus T^*Q$:

$$D_{TQ \oplus T^*Q}(x) := \left\{ (\tilde{v}_x, \tilde{\alpha}_x) \in T_x(TQ \oplus T^*Q) \times T_x^*(TQ \oplus T^*Q) \mid \right. \\ \left. \tilde{v}_x \in \Delta_{TQ \oplus T^*Q}(x) \text{ and } \tilde{\alpha}_x - (\text{pr}_{T^*Q}^* \Omega)^\flat(x)(\tilde{v}_x) \in \Delta_{TQ \oplus T^*Q}^\circ(x) \right\},$$

for $x \in TQ \oplus T^*Q$. A local representation for the Dirac structure $D_{TQ \oplus T^*Q}$ is

$$D_{TQ \oplus T^*Q}(q, v, p) = \left\{ ((q, v, p, \dot{q}, \dot{v}, \dot{p}), (q, v, p, \tilde{\alpha}_q, \tilde{\alpha}_v, \tilde{\alpha}_p)) \mid \right. \\ \left. \dot{q} \in \Delta_Q(q), \tilde{\alpha}_p = \dot{q}, \tilde{\alpha}_v = 0, \tilde{\alpha}_q + \dot{p} \in \Delta_Q^\circ(q) \right\}.$$

Then, we have the following result.

Theorem 2.5. *For every $u \in \Delta_Q$, define $z := \mathbb{F}L(u) \in P$ and $x := \rho_{T^*TQ} \circ dL(u) \in \mathcal{K}$ so that $\text{pr}_{T^*Q}(x) = z$. Then, we have*

$$(X(z), \mathfrak{D}L(u)) \in D_{\Delta_Q}(z) \iff (\tilde{X}(x), d\mathcal{E}(x)) \in D_{TQ \oplus T^*Q}(x).$$

Proof. It is not difficult to prove that the condition $(\tilde{X}(x), d\mathcal{E}(x)) \in D_{TQ \oplus T^*Q}(x)$ locally reads

$$p = \frac{\partial L}{\partial v}(q, v), \quad \dot{q} \in \Delta_Q(q), \quad \dot{q} = v, \quad \dot{p} - \frac{\partial L}{\partial q}(q, v) \in \Delta_Q^\circ(q),$$

that is, the Lagrange–Dirac equations (2.3). So, we have the equivalence. \square

As a consequence, we obtain the following result which was obtained by Yoshimura and Marsden (see Theorem 3.8. in [52]).

Corollary 2.6. *If $(q(t), p(t)) = \mathbb{F}L(q(t), v(t))$, $t_1 \leq t \leq t_2$, is an integral curve of the vector field X on P , then $\rho_{T^*TQ} \circ dL(q(t), v(t))$ is an integral curve of \tilde{X} on \mathcal{K} . Conversely, if $(q(t), v(t), p(t))$, $t_1 \leq t \leq t_2$, is an integral curve of \tilde{X} on \mathcal{K} , then $\text{pr}_{T^*Q}(q(t), v(t), p(t))$ is an integral curve of X .*

Therefore, a Lagrange–Dirac system on the Pontryagin bundle is given by a triple $(\mathcal{E}, \mathcal{K}, \tilde{X})$ satisfying the condition

$$(\tilde{X}(x), d\mathcal{E}(x)) \in D_{TQ \oplus T^*Q}(x),$$

for all $x \in \mathcal{K}$.

3. DEGENERATE LAGRANGIAN SYSTEMS WITH NONHOLONOMIC CONSTRAINTS

If one accurately models a mechanical system, then one usually obtains a non-degenerate Lagrangian, since the kinetic energy of the system is usually written as a positive-definite quadratic form in their velocity components. However, for a complex mechanical system consisting of many moving parts, one can often ignore the masses and/or moments of inertia of relatively light parts of the system in order to simplify the analysis. This turns out to be an effective way of modeling complex systems; for example, one usually models the strings of a puppet as massless moving parts (see, e.g., Johnson and Murphey [29] and Murphey and Egerstedt [40]). With such an approximation, the Lagrangian often turns out to be degenerate, and thus the Euler–Lagrange or Lagrange–d’Alembert equations do not give the dynamics of the massless parts directly; instead, it is determined by mechanical constraints. In other words, the system may be considered as a hybrid of dynamics and kinematics.

We are particularly interested in systems with degenerate Lagrangians and *nonholonomic* constraints, because they possess the two very features that Lagrange–Dirac systems can (and are designed to) incorporate but the standard Lagrangian or Hamiltonian formulation cannot.

In this section, we introduce a class of mechanical systems with degenerate Lagrangians and nonholonomic constraints with symmetry that yield non-degenerate almost Hamiltonian systems² on the reduced space when symmetry reduction is performed.

3.1. Chaplygin Systems. Let us start from the following definition of a well-known class of nonholonomic systems:

Definition 3.1 (Chaplygin Systems; see, e.g., Koiller [32] and Hochgerner and García-Naranjo [27]). A nonholonomic system with Lagrangian L and distribution Δ_Q is called a *Chaplygin system* if there exists a Lie group G with a free and proper action on Q , i.e., $\Phi : G \times Q \rightarrow Q$ or $\Phi_g : Q \rightarrow Q$ for any $g \in G$, such that

- (i) the Lagrangian L and the distribution Δ_Q are invariant under the tangent lift of the G -action, i.e., $L \circ T\Phi_g = L$ and $T\Phi_g(\Delta_Q(q)) = \Delta_Q(gq)$;
- (ii) for each $q \in Q$, the tangent space T_qQ is the direct sum of the constraint distribution and the tangent space to the orbit of the group action, i.e.,

$$T_qQ = \Delta_Q(q) \oplus T_q\mathcal{O}_q,$$

where \mathcal{O}_q is the orbit through q of the G -action on Q , i.e.,

$$\mathcal{O}_q := \{\Phi_g(q) \in Q \mid g \in G\}.$$

This setup gives rise to the principal bundle

$$\pi : Q \rightarrow Q/G =: \bar{Q}$$

and the connection

$$\mathcal{A} : TQ \rightarrow \mathfrak{g}, \tag{3.1}$$

with \mathfrak{g} being the Lie algebra of G such that $\ker \mathcal{A} = \Delta_Q$, i.e., the horizontal space of \mathcal{A} is Δ_Q . Furthermore, for any $q \in Q$ and $\bar{q} := \pi(q) \in \bar{Q}$, the map $T_q\pi|_{\Delta_Q(q)} : \Delta_Q(q) \rightarrow T_{\bar{q}}\bar{Q}$ is a linear isomorphism, and hence we have the horizontal lift

$$\text{hl}_q^\Delta : T_{\bar{q}}\bar{Q} \rightarrow \Delta_Q(q); \quad v_{\bar{q}} \mapsto (T_q\pi|_{\Delta_Q(q)})^{-1}(v_{\bar{q}}).$$

We will occasionally use the following shorthand notation for horizontal lifts:

$$v_q^h := \text{hl}_q^\Delta(v_{\bar{q}}).$$

Then, any vector $W_q \in T_qQ$ can be decomposed into the horizontal and vertical parts as follows:

$$W_q = \text{hor}(W_q) + \text{ver}(W_q),$$

with

$$\text{hor}(W_q) = \text{hl}_q^\Delta(w_{\bar{q}}), \quad \text{ver}(W_q) = (\mathcal{A}_q(W_q))_Q(q),$$

where $w_{\bar{q}} := T_q\pi(W_q)$ and $\xi_Q \in \mathfrak{X}(Q)$ is the infinitesimal generator of $\xi \in \mathfrak{g}$.

²An *almost Hamiltonian system* is a generalized Hamiltonian system defined with a non-closed two-form as opposed to a symplectic form (which is closed by definition) [2, 27].

Suppose that the Lagrangian $L : TQ \rightarrow \mathbb{R}$ is of the form

$$L(v_q) = \frac{1}{2}g_q(v_q, v_q) - V(q), \quad (3.2)$$

where g is a possibly degenerate metric on Q . We may then define the reduced Lagrangian

$$\bar{L} := L \circ \text{hl}^\Delta,$$

or more explicitly,

$$\bar{L} : T\bar{Q} \rightarrow \mathbb{R}; \quad v_{\bar{q}} \mapsto \frac{1}{2}\bar{g}_{\bar{q}}(v_{\bar{q}}, v_{\bar{q}}) - \bar{V}(\bar{q}),$$

where \bar{g} is the metric on the reduced space \bar{Q} induced by g as follows:

$$\bar{g}_{\bar{q}}(v_{\bar{q}}, w_{\bar{q}}) := g_q \left(\text{hl}_q^\Delta(v_{\bar{q}}), \text{hl}_q^\Delta(w_{\bar{q}}) \right) = g_q(v_q^h, w_q^h),$$

and the reduced potential $\bar{V} : \bar{Q} \rightarrow \mathbb{R}$ is defined such that $V = \bar{V} \circ \pi$.

3.2. Weakly Degenerate Chaplygin Systems. The following special class of Chaplygin systems is of particular interest in this paper:

Definition 3.2 (Weakly Degenerate Chaplygin Systems). A Chaplygin system is said to be *weakly degenerate* if the Lagrangian $L : TQ \rightarrow \mathbb{R}$ is degenerate but the reduced Lagrangian $\bar{L} : T\bar{Q} \rightarrow \mathbb{R}$ is non-degenerate; more precisely, the metric g is degenerate on TQ but positive-definite (hence non-degenerate) when restricted to $\Delta_Q \subset TQ$, i.e., the triple (Q, Δ_Q, g) defines a sub-Riemannian manifold (see, e.g., Montgomery [39]), and the induced metric \bar{g} on \bar{Q} is positive-definite and hence Riemannian.

Remark 3.3. This is a mathematical description of the hybrid of dynamics and kinematics mentioned above: The dynamics is essentially dropped to the reduced configuration manifold $\bar{Q} := Q/G$, and the rest is reconstructed by the horizontal lift hl^Δ , which is the kinematic part defined by the (nonholonomic) constraints.

We will look into the geometry associated with weakly degenerate Chaplygin systems in Section 5.1.

Example 3.4 (Simplified Roller Racer; see Tsakiris [46] and Krishnaprasad and Tsakiris [34] and Bloch [3, Section 1.10]). The roller racer, shown in Fig. 1, consists of two (main and second) planar coupled rigid bodies, each of which has a pair of wheels attached at its center of mass. We assume that *the mass of the second body is negligible, and hence so are its kinetic and rotational energies*³. Let (x, y) be the coordinates of the center of mass of the main body, θ the angle of the line passing through the center of mass measured from the x -axis, ϕ the angle between the two bodies; d_1 and d_2 are the distances from centers of mass to the joint, m_1 and I_1 the mass and inertia of the main body.

The configuration space is $Q = SE(2) \times \mathbb{S}^1 = \{(x, y, \theta, \phi)\}$, and the Lagrangian $L : TQ \rightarrow \mathbb{R}$ is given by

$$L = \frac{1}{2}m_1(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I_1\dot{\theta}^2,$$

which is degenerate because of the massless approximation of the second body.

The constraints are given by

$$\dot{x} = \cos \theta \csc \phi \left[(d_1 \cos \phi + d_2) \dot{\theta} + d_2 \dot{\phi} \right], \quad \dot{y} = \sin \theta \csc \phi \left[(d_1 \cos \phi + d_2) \dot{\theta} + d_2 \dot{\phi} \right]. \quad (3.3)$$

Defining the constraint one-forms

$$\omega^1 := dx - \cos \theta \csc \phi [(d_1 \cos \phi + d_2)d\theta + d_2 d\phi], \quad \omega^2 := dy - \sin \theta \csc \phi [(d_1 \cos \phi + d_2)d\theta + d_2 d\phi], \quad (3.4)$$

we can write the constraint distribution $\Delta_Q \subset TQ$ as

$$\Delta_Q = \left\{ \dot{q} = (\dot{x}, \dot{y}, \dot{\theta}, \dot{\phi}) \in TQ \mid \omega^a(\dot{q}) = 0, a = 1, 2 \right\}.$$

³We note that, in the original model in [34, 46], only the kinetic energy of the second body is ignored, and its rotational energy is taken into account; one obtains a *non-degenerate* Lagrangian with such an approximation.

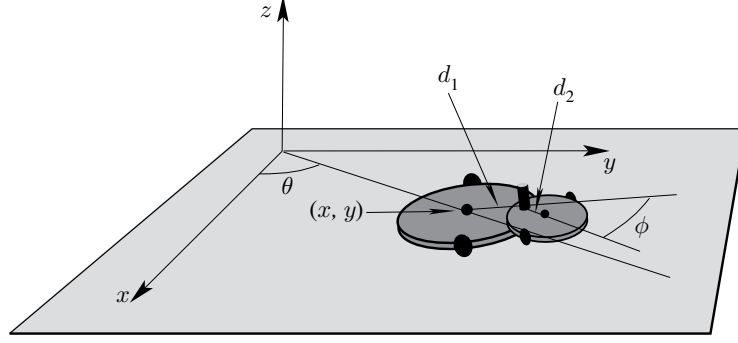


FIGURE 1. Roller Racer (taken from Bloch [3] with permission from the author). The mass of the second body is assumed to be negligible.

The Lagrange–Dirac equations (2.3) give

$$\begin{aligned}
 p_x &= m_1 v_x, & p_y &= m_1 v_y, & p_\theta &= I_1 v_\theta, & p_\phi &= 0, \\
 \dot{x} &= \cos \theta \csc \phi \left[(d_1 \cos \phi + d_2) \dot{\theta} + d_2 \dot{\phi} \right], & \dot{y} &= \sin \theta \csc \phi \left[(d_1 \cos \phi + d_2) \dot{\theta} + d_2 \dot{\phi} \right], \\
 \dot{x} &= v_x, & \dot{y} &= v_y, & \dot{\theta} &= v_\theta, & \dot{\phi} &= v_\phi, \\
 \dot{p}_x &= \lambda \sin \theta, & \dot{p}_y &= -\lambda \cos \theta, & \dot{p}_\theta &= 0, & \dot{p}_\phi &= 0,
 \end{aligned} \tag{3.5}$$

where λ is the Lagrange multiplier.

Let $G = \mathbb{R}^2$ and consider the action of G on Q by translations on the x - y plane, i.e.,

$$G \times Q \rightarrow Q; \quad ((a, b), (x, y, \theta, \phi)) \mapsto (x + a, y + b, \theta, \phi).$$

Then, the tangent space to the group orbit is given by

$$T_q \mathcal{O}(q) = \text{span} \left\{ \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right\},$$

with $q = (x, y, \theta, \phi)$. It is easy to check that this defines a Chaplygin system in the sense of Definition 3.1. The quotient space is $\bar{Q} := Q/G = \{(\theta, \phi)\}$, and the horizontal lift hl^Δ is

$$\text{hl}_q^\Delta(\dot{\theta}, \dot{\phi}) = \left(\cos \theta \csc \phi \left[(d_1 \cos \phi + d_2) \dot{\theta} + d_2 \dot{\phi} \right], \sin \theta \csc \phi \left[(d_1 \cos \phi + d_2) \dot{\theta} + d_2 \dot{\phi} \right], \dot{\theta}, \dot{\phi} \right).$$

Hence, the reduced Lagrangian $\bar{L} : T\bar{Q} \rightarrow \mathbb{R}$ is given by

$$\bar{L} = \frac{1}{2} m_1 \left(d_1 \dot{\theta} \cos \phi + d_2 (\dot{\theta} + \dot{\phi}) \right)^2 \csc^2 \phi + \frac{1}{2} I_1 \dot{\theta}^2, \tag{3.6}$$

which is non-degenerate; hence the simplified roller racer is a weakly degenerate Chaplygin system.

Therefore, the dynamics of the variables θ and ϕ are specified by the equations of motion, which together with the (nonholonomic) constraints, Eq. (3.3), determine the time evolution of the variables x and y .

Example 3.5 (Bicycle; see Getz [20], Getz and Marsden [21], and Koon and Marsden [33]). Consider the simplified model of a bicycle shown in Fig. 2: For the sake of simplicity, the wheels are assumed to be massless, and the mass m of the bicycle is considered to be concentrated at a single point; however we take into account the moment of inertia of the steering wheel.

The configuration space is $Q = SE(2) \times \mathbb{S}^1 \times \mathbb{S}^1 = \{(x, y, \theta, \phi, \psi)\}$; the variables x , y , θ , and ψ are defined as in Fig. 2 and $\phi := \tan \sigma / b$; also let $J(\phi, \psi)$ be the moment of inertia associated with the steering action. The Lagrangian $L : TQ \rightarrow \mathbb{R}$ is given by

$$\begin{aligned}
 L = \frac{m}{2} \left[(\cos \theta \dot{x} + \sin \theta \dot{y} + a \sin \psi \dot{\theta})^2 + (\sin \theta \dot{x} - \cos \theta \dot{y} + a \cos \psi \dot{\theta} - c \dot{\theta})^2 \right. \\
 \left. + a^2 \sin \psi \dot{\psi}^2 \right] + \frac{J(\phi, \psi)}{2} \dot{\phi}^2 - m g a \cos \psi,
 \end{aligned}$$

which is degenerate. The constraints are given by

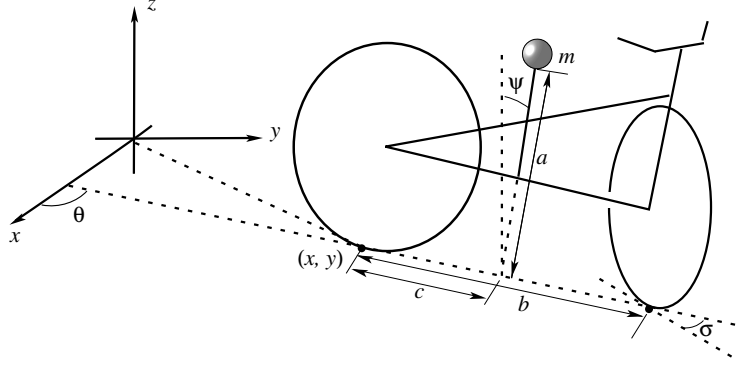


FIGURE 2. Bicycle (taken from Koon and Marsden [33] with permission from Wang Sang Koon).

$$\dot{\theta} = \phi(\cos \theta \dot{x} + \sin \theta \dot{y}), \quad \sin \theta \dot{x} - \cos \theta \dot{y} = 0.$$

Defining the constraint one-forms

$$\omega^1 := \phi(\cos \theta dx + \sin \theta dy), \quad \omega^2 := \sin \theta dx - \cos \theta dy,$$

we can write the constraint distribution $\Delta_Q \subset TQ$ as

$$\Delta_Q = \left\{ \dot{q} = (\dot{x}, \dot{y}, \dot{\theta}, \dot{\phi}, \dot{\psi}) \in TQ \mid \omega^a(\dot{q}) = 0, a = 1, 2 \right\}.$$

Let $G = \mathbb{R}^2$ and consider the action of G on Q by translations on the x - y plane, i.e.,

$$G \times Q \rightarrow Q; \quad ((a, b), (x, y, \theta, \phi, \psi)) \mapsto (x + a, y + b, \theta, \phi, \psi).$$

Then, the tangent space to the group orbit is given by

$$T_q \mathcal{O}(q) = \text{span} \left\{ \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right\},$$

with $q = (x, y, \theta, \phi, \psi)$. It is easy to check that this defines a Chaplygin system in the sense of Definition 3.1. The quotient space is $\bar{Q} := Q/G = \{(\theta, \phi, \psi)\}$, and the horizontal lift hl^Δ is

$$\text{hl}_q^\Delta(\dot{\theta}, \dot{\phi}, \dot{\psi}) = \left(\frac{\dot{\theta}}{\phi} \cos \theta, \frac{\dot{\theta}}{\phi} \sin \theta, \dot{\theta}, \dot{\phi}, \dot{\psi} \right).$$

Hence, the reduced Lagrangian $\bar{L} : T\bar{Q} \rightarrow \mathbb{R}$ is given by

$$\bar{L} = \frac{m}{2} \left[(c\dot{\theta} - a \cos \psi \dot{\psi})^2 + \frac{(\dot{\theta} + a \sin \psi \dot{\theta})^2}{\phi^2} + a^2 \sin \psi \dot{\psi}^2 \right] + \frac{J(\phi, \psi)}{2} \dot{\phi}^2 - mga \cos \psi,$$

which is non-degenerate, and so this is a weakly degenerate Chaplygin system as well.

4. HAMILTON–JACOBI THEORY FOR LAGRANGE–DIRAC SYSTEMS

4.1. Hamilton–Jacobi Theorem for Lagrange–Dirac systems. We now state the main theorem of this paper, which relates the dynamics of the Lagrange–Dirac system with what we refer to as the *Dirac–Hamilton–Jacobi equation*.

Theorem 4.1 (Dirac–Hamilton–Jacobi Theorem). *Suppose that a Lagrangian $L : TQ \rightarrow \mathbb{R}$ and a distribution $\Delta_Q \subset TQ$ are given. Define $\Upsilon : Q \rightarrow TQ \oplus T^*Q$ by*

$$\Upsilon(q) := \mathcal{X}(q) \oplus \gamma(q),$$

with a vector field $\mathcal{X} : Q \rightarrow TQ$ and a one-form $\gamma : Q \rightarrow T^*Q$, and assume that it satisfies

$$\Upsilon(q) \in \mathcal{K}_q \text{ for any } q \in Q, \tag{4.1}$$

and

$$d\gamma|_{\Delta_Q} = 0, \text{ i.e., } d\gamma(v, w) = 0 \text{ for any } v, w \in \Delta_Q. \tag{4.2}$$

Then, the following are equivalent:

(i) For every integral curve $c(t)$ of \mathcal{X} , i.e., for every curve $c : \mathbb{R} \rightarrow Q$ satisfying

$$\dot{c}(t) = \mathcal{X}(c(t)), \quad (4.3)$$

the curve $t \mapsto \Upsilon \circ c(t) = (\mathcal{X} \oplus \gamma) \circ c(t)$ is an integral curve of the Lagrange–Dirac equations (2.3).

(ii) Υ satisfies the following Dirac–Hamilton–Jacobi equation:

$$d(\mathcal{E} \circ \Upsilon) \in \Delta_Q^\circ, \quad (4.4)$$

or, if Q is connected and Δ_Q is completely nonholonomic⁴,

$$\mathcal{E} \circ \Upsilon = E, \quad (4.5)$$

with a constant E .

Proof. Let us first show that (ii) implies (i). Assume (ii) and let $c(t)$ be an integral curve of \mathcal{X} , and then set

$$v(t) \oplus p(t) := \Upsilon \circ c(t) = (\mathcal{X} \oplus \gamma) \circ c(t).$$

Then, clearly $v(t) = \dot{c}(t) = \mathcal{X}(c(t))$. Also, Eq. (4.1) implies that

$$v(t) \in \Delta_Q(c(t)), \quad p(t) = \frac{\partial L}{\partial v}(q(t), v(t)).$$

So it remains to show $\dot{p} - \partial L / \partial q \in \Delta_Q^\circ$. To that end, first calculate

$$\dot{p}_j(t) = \frac{d}{dt} \gamma_j \circ c(t) = \frac{\partial \gamma_j}{\partial q^i}(c(t)) \dot{c}^i(t) = \frac{\partial \gamma_j}{\partial q^i}(c(t)) \mathcal{X}^i(c(t))$$

and so, for any $w \in \Delta_Q$, we have

$$\dot{p}_j(t) w^j = \frac{\partial \gamma_j}{\partial q^i}(c(t)) \mathcal{X}^i(c(t)) w^j = \frac{\partial \gamma_j}{\partial q^i}(c(t)) \mathcal{X}^i(c(t)) w^j, \quad (4.6)$$

since Eq. (4.2) implies, for any $v, w \in \Delta_Q$,

$$\frac{\partial \gamma_i}{\partial q^j} v^i w^j = \frac{\partial \gamma_j}{\partial q^i} v^i w^j,$$

and also Eq. (4.1) gives $\mathcal{X}(q) \in \Delta_Q(q)$. On the other hand,

$$\begin{aligned} d(\mathcal{E} \circ \Upsilon) &= d(\gamma_i(q) \mathcal{X}^i(q) - L(q, \mathcal{X}(q))) \\ &= \left(\frac{\partial \gamma_i}{\partial q^j} \mathcal{X}^i + \gamma_i \frac{\partial \mathcal{X}^i}{\partial q^j} - \frac{\partial L}{\partial q^j} - \frac{\partial L}{\partial v^i} \frac{\partial \mathcal{X}^i}{\partial q^j} \right) dq^j \\ &= \left(\frac{\partial \gamma_i}{\partial q^j} \mathcal{X}^i - \frac{\partial L}{\partial q^j} \right) dq^j, \end{aligned}$$

where we used the following relation that follows from Eq. (4.1):

$$\gamma(q) = \frac{\partial L}{\partial v}(q, \mathcal{X}(q)).$$

So the Dirac–Hamilton–Jacobi equation (4.4) with Eq. (4.6) implies

$$\begin{aligned} d(\mathcal{E} \circ \Upsilon)(c(t)) \cdot w &= \left(\frac{\partial \gamma_i}{\partial q^j}(c(t)) \mathcal{X}^i(c(t)) - \frac{\partial L}{\partial q^j}(c(t), v(t)) \right) w^j \\ &= \left(\dot{p}_j(t) - \frac{\partial L}{\partial q^j}(c(t), v(t)) \right) w^j = 0. \end{aligned}$$

Since $w \in \Delta_Q$ is arbitrary, this implies

$$\dot{p}(t) - \frac{\partial L}{\partial q}(c(t), v(t)) \in \Delta_Q^\circ.$$

Therefore, (i) is satisfied.

⁴A distribution $\Delta_Q \subset TQ$ is said to be *completely nonholonomic* (or *bracket-generating*) if Δ_Q along with all of its iterated Lie brackets $[\Delta_Q, \Delta_Q], [\Delta_Q, [\Delta_Q, \Delta_Q]], \dots$ spans the tangent bundle TQ . See, e.g., Vershik and Gershkovich [51] and Montgomery [39].

Conversely, assume (i); let $c(t)$ be a curve in Q that satisfies Eq. (4.3) and set $v(t) \oplus p(t) := \Upsilon \circ c(t) = (\mathcal{X} \oplus \gamma) \circ c(t)$. Then, by assumption, $(c(t), v(t), p(t))$ is an integral curve of the Lagrange–Dirac system (2.2), and so

$$\dot{p}(t) - \frac{\partial L}{\partial q}(c(t), v(t)) \in \Delta_Q^\circ(c(t)).$$

Following the same calculations as above we have, for any $w \in \Delta_Q$,

$$d(\mathcal{E} \circ \Upsilon)(c(t)) \cdot w = \left(\dot{p}_j(t) - \frac{\partial L}{\partial q^j}(c(t), v(t)) \right) w^j = 0.$$

For an arbitrary point $q \in Q$, we can consider an integral curve $c(t)$ of X such that $c(0) = q$. Therefore, the above equation implies that $d(\mathcal{E} \circ \Upsilon)(q) \cdot w_q = 0$ for any $q \in Q$ and $w_q \in \Delta_Q(q)$, which gives the Dirac–Hamilton–Jacobi equation (4.4). If Q is connected and Δ_Q is completely nonholonomic, then by the same argument as in the proof of Theorem 3.1 in Ohsawa and Bloch [42], $d(\mathcal{E} \circ \Upsilon) \in \Delta_Q^\circ$ reduces to $\mathcal{E} \circ \Upsilon = E$ for some constant E . \square

Theorem 4.1 can be recast in the context of Section 2.4 as follows:

Corollary 4.2. *Under the same conditions as in Theorem 4.1, the following are equivalent:*

(i) *For every curve $c(t)$ such that*

$$\dot{c}(t) = T\text{pr}_Q \cdot \tilde{X}(\Upsilon \circ c(t)),$$

the curve $t \mapsto \Upsilon \circ c(t)$ is an integral curve of \tilde{X} , and so it is an integral curve of the Lagrange–Dirac equations (2.3).

(ii) *Υ satisfies $(0, d(\mathcal{E} \circ \Upsilon \circ \text{pr}_Q)) \in D_{TQ \oplus T^*Q}$, or equivalently, $d(\mathcal{E} \circ \Upsilon \circ \text{pr}_Q) \in \Delta_{TQ \oplus T^*Q}^\circ$.*

Proof. The equivalence of (i) with that of Theorem 4.1 follows from the relation $T\text{pr}_Q \circ \tilde{X} \circ \Upsilon = \mathcal{X}$, which is easily checked by coordinate calculations.

On the other hand, for (ii), first observe that $\text{pr}_Q^*(\Delta_Q^\circ) = \Delta_{TQ \oplus T^*Q}^\circ$. Then, since $\text{pr}_Q : TQ \oplus T^*Q \rightarrow Q$ is a surjective submersion, it follows that

$$d(\mathcal{E} \circ \Upsilon) \in \Delta_Q^\circ \iff \text{pr}_Q^* d(\mathcal{E} \circ \Upsilon) \in \text{pr}_Q^*(\Delta_Q^\circ) \iff d(\mathcal{E} \circ \Upsilon \circ \text{pr}_Q) \in \Delta_{TQ \oplus T^*Q}^\circ.$$

This proves the equivalence of (ii) with that of Theorem 4.1. \square

4.2. Nonholonomic Hamilton–Jacobi Theory as a Special Case. Let us show that the nonholonomic Hamilton–Jacobi equation of Iglesias-Ponte et al. [28] and Ohsawa and Bloch [42] follows as a special case of the above theorem. Consider the special case where the Lagrangian $L : TQ \rightarrow \mathbb{R}$ is non-degenerate, i.e., the Legendre transformation $\mathbb{F}L : TQ \rightarrow T^*Q$ is invertible. Then, we may rewrite the definition of the submanifold $\mathcal{K} \subset TQ \oplus T^*Q$, Eq. (2.4), by

$$\begin{aligned} \mathcal{K} &= \{v_q \oplus p_q \in TQ \oplus T^*Q \mid v_q \in \Delta_Q(q), p_q = \mathbb{F}L(v_q)\} \\ &= \{v_q \oplus p_q \in TQ \oplus T^*Q \mid p_q \in P_q, v_q = (\mathbb{F}L)^{-1}(p_q)\} \\ &= \Delta_Q \oplus P, \end{aligned}$$

where we recall that $P := \mathbb{F}L(\Delta_Q)$. It implies that if $\Upsilon = \mathcal{X} \oplus \gamma$ then $\mathcal{X} = (\mathbb{F}L)^{-1} \circ \gamma$, and thus

$$\mathcal{E} \circ \Upsilon(q) = \langle \gamma(q), (\mathbb{F}L)^{-1}(\gamma(q)) \rangle - L \circ (\mathbb{F}L)^{-1}(\gamma(q)) = H \circ \gamma(q),$$

with γ taking values in P and the Hamiltonian $H : T^*Q \rightarrow \mathbb{R}$ defined by

$$H(q, p) := \langle p_q, (\mathbb{F}L)^{-1}(p_q) \rangle - L \circ (\mathbb{F}L)^{-1}(p_q).$$

Then, the Lagrange–Dirac equations (2.3) become the nonholonomic Hamilton’s equations

$$\dot{q} = \frac{\partial H}{\partial p}(q, p), \quad \dot{p} + \frac{\partial H}{\partial q}(q, p) \in \Delta_Q^\circ(q). \quad \dot{q} \in \Delta_Q(q),$$

or, in an intrinsic form,

$$i_{X_H^{\text{nh}}} \Omega - dH \in \Delta_{T^*Q}^\circ, \quad T\pi_Q(X_H^{\text{nh}}) \in \Delta_Q$$

for a vector field X_H^{nh} on T^*Q . Furthermore, it is straightforward to show that

$$(\mathbb{F}L)^{-1} = \mathbb{F}H = T\pi_Q \circ X_H,$$

where X_H is the Hamiltonian vector field of the unconstrained system with the same Hamiltonian, i.e., $i_{X_H}\Omega = dH$; hence we obtain

$$\mathcal{X}(q) = (\mathbb{F}L)^{-1} \circ \gamma(q) = T\pi_Q \cdot X_H(\gamma(q)).$$

Therefore, Theorem 4.1 specializes to the nonholonomic Hamilton–Jacobi theorem of Iglesias-Ponte et al. [28] and Ohsawa and Bloch [42]:

Corollary 4.3 (Nonholonomic Hamilton–Jacobi [28, 42]). *Consider a nonholonomic system defined on a configuration manifold Q with a Lagrangian of the form Eq. (3.2) and a nonholonomic constraint distribution $\Delta_Q \subset TQ$. Let $\gamma : Q \rightarrow T^*Q$ be a one-form that satisfies*

$$\gamma(q) \in P_q \text{ for any } q \in Q,$$

and

$$d\gamma|_{\Delta_Q} = 0, \text{ i.e., } d\gamma(v, w) = 0 \text{ for any } v, w \in \Delta_Q.$$

Then, the following are equivalent:

(i) For every curve $c(t)$ in Q satisfying

$$\dot{c}(t) = T\pi_Q \cdot X_H(\gamma \circ c(t)),$$

the curve $t \mapsto \gamma \circ c(t)$ is an integral curve of X_H^{nh} , where X_H is the Hamiltonian vector field of the unconstrained system with the same Hamiltonian, i.e., $i_{X_H}\Omega = dH$.

(ii) The one-form γ satisfies the nonholonomic Hamilton–Jacobi equation:

$$d(H \circ \gamma) \in \Delta_Q^\circ,$$

or, if Q is connected and Δ_Q is completely nonholonomic,

$$H \circ \gamma = E,$$

with a constant E .

4.3. Applications to Degenerate Lagrangian System with Holonomic Constraints. If the constraints are holonomic, then the distribution $\Delta_Q \subset TQ$ is integrable, and so there exists a local submanifold $S \subset Q$ such that $T_s S = \Delta_Q(s)$ for any $s \in S$. Let $\iota_S : S \hookrightarrow Q$ be the inclusion. Then, the Dirac–Hamilton–Jacobi equation (4.4) gives

$$\iota_S^* d(\mathcal{E} \circ \Upsilon) \in (TS)^\circ = 0,$$

and thus

$$d(\mathcal{E} \circ \Upsilon \circ \iota_S) = 0,$$

which implies that we have

$$\mathcal{E} \circ \Upsilon \circ \iota_S = E, \tag{4.7}$$

with a constant E .

On the other hand, the condition (4.2) becomes

$$\iota_S^* d\gamma = d(\gamma \circ \iota_S) = 0, \tag{4.8}$$

and so $\gamma \circ \iota_S = dW$ for some function W defined locally on S .

Example 4.4 (LC circuit; see Yoshimura and Marsden [52, 54]). Consider the LC circuit shown in Fig. 3. The configuration space is the 4-dimensional vector space $Q = \{(q_\ell, q_{c_1}, q_{c_2}, q_{c_3})\}$, which represents charges

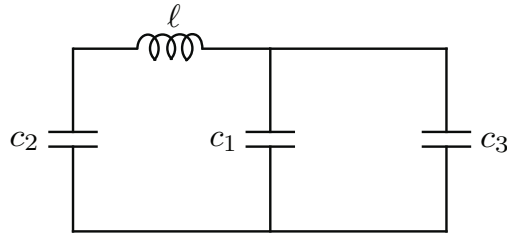


FIGURE 3. LC circuit (see Yoshimura and Marsden [52]).

in the circuit elements. Then $TQ \cong Q \times Q$ and $f_q = (f_\ell, f_{c_1}, f_{c_2}, f_{c_3}) \in T_q Q$ represents the currents in the corresponding circuit elements. The Lagrangian $L : TQ \rightarrow \mathbb{R}$ is given by

$$L(q, f) = \frac{1}{2} \ell f_\ell^2 - \frac{1}{2} \frac{q_{c_1}^2}{c_1} - \frac{1}{2} \frac{q_{c_2}^2}{c_2} - \frac{1}{2} \frac{q_{c_3}^2}{c_3},$$

which is clearly degenerate.

The generalized energy $\mathcal{E} : TQ \oplus T^*Q \rightarrow \mathbb{R}$ is

$$\begin{aligned} \mathcal{E}(q, f, p) &= p \cdot f - L(q, f) \\ &= p_\ell f_\ell + p_{c_1} f_{c_1} + p_{c_2} f_{c_2} + p_{c_3} f_{c_3} - \frac{1}{2} \ell f_\ell^2 + \frac{1}{2} \frac{q_{c_1}^2}{c_1} + \frac{1}{2} \frac{q_{c_2}^2}{c_2} + \frac{1}{2} \frac{q_{c_3}^2}{c_3}. \end{aligned}$$

The Kirchhoff current law gives the constraints $-f_\ell + f_{c_2} = 0$ and $f_{c_1} - f_{c_2} + f_{c_3} = 0$, or in terms of constraint one-forms, $\omega^1 = -dq_\ell + dq_{c_2}$ and $\omega^2 = dq_{c_1} - dq_{c_2} + dq_{c_3}$. Thus, the constraint distribution $\Delta_Q \subset TQ$ is given by

$$\Delta_Q = \{f \in TQ \mid \omega^a(f) = 0, a = 1, 2\}.$$

So the submanifold $\mathcal{K} \subset TQ \oplus T^*Q$ is

$$\mathcal{K} = \{(q, f, p) \in TQ \oplus T^*Q \mid f_\ell = f_{c_2}, f_{c_2} = f_{c_1} + f_{c_3}, p_\ell = \ell f_\ell, p_{c_1} = p_{c_2} = p_{c_3} = 0\}.$$

Hence, the generalized energy constrained to \mathcal{K} is

$$\mathcal{E}|_{\mathcal{K}} = \frac{1}{2} \ell f_\ell^2 + \frac{1}{2} \frac{q_{c_1}^2}{c_1} + \frac{1}{2} \frac{q_{c_2}^2}{c_2} + \frac{1}{2} \frac{q_{c_3}^2}{c_3}.$$

Notice that the constraints are holonomic, i.e., the constraints can be integrated to give

$$q_\ell - q_{c_2} = a_0, \quad q_{c_2} - q_{c_1} - q_{c_3} = a_1,$$

with some constants a_0 and a_1 . So we define a submanifold $S \subset Q$ by

$$S := \{(q_\ell, q_{c_1}, q_{c_2}, q_{c_3}) \in Q \mid q_{c_2} = q_\ell - a_0, q_{c_3} = q_{c_2} - q_{c_1} - a_1\} = \{(q_\ell, q_{c_1})\},$$

and the inclusion

$$\iota_S : S \rightarrow Q; \quad (q_\ell, q_{c_1}) \mapsto (q_\ell, q_{c_1}, q_\ell - a_0, q_{c_2} - q_{c_1} - a_1).$$

Now, the Dirac–Hamilton–Jacobi equation for holonomic systems, Eq. (4.7), gives

$$\mathcal{E} \circ \Upsilon \circ \iota_S = E, \tag{4.9}$$

with some constant E , where $\Upsilon \circ \iota_S : S \rightarrow TQ \oplus T^*Q$ is

$$\Upsilon \circ \iota_S(q_\ell, q_{c_1}) = \left(q_\ell, q_{c_1}, \tilde{\mathcal{X}}(q_\ell, q_{c_1}), \tilde{\gamma}(q_\ell, q_{c_1}) \right)$$

with $\tilde{\mathcal{X}} := \mathcal{X} \circ \iota_S : S \rightarrow TQ$ and $\tilde{\gamma} := \gamma \circ \iota_S : S \rightarrow T^*Q$ given by

$$\begin{aligned} \tilde{\mathcal{X}}(q_\ell, q_{c_1}) &= \left(\tilde{\mathcal{X}}_\ell(q_\ell, q_{c_1}), \tilde{\mathcal{X}}_{c_1}(q_\ell, q_{c_1}), \tilde{\mathcal{X}}_{c_2}(q_\ell, q_{c_1}), \tilde{\mathcal{X}}_{c_3}(q_\ell, q_{c_1}) \right), \\ \tilde{\gamma}(q_\ell, q_{c_1}) &= (\tilde{\gamma}_\ell(q_\ell, q_{c_1}), \tilde{\gamma}_{c_1}(q_\ell, q_{c_1}), \tilde{\gamma}_{c_2}(q_\ell, q_{c_1}), \tilde{\gamma}_{c_3}(q_\ell, q_{c_1})). \end{aligned}$$

The condition $\Upsilon \circ \iota_S(q_\ell, q_{c_1}) \in \mathcal{K}$ implies

$$\tilde{\mathcal{X}}_\ell = \tilde{\mathcal{X}}_{c_2}, \quad \tilde{\mathcal{X}}_{c_2} = \tilde{\mathcal{X}}_{c_1} + \tilde{\mathcal{X}}_{c_3}, \quad \tilde{\gamma}_\ell = \ell \tilde{\mathcal{X}}_\ell, \quad \tilde{\gamma}_{c_1} = \tilde{\gamma}_{c_2} = \tilde{\gamma}_{c_3} = 0.$$

Then,

$$\tilde{\gamma} = \gamma \circ \iota_S = \ell \tilde{\mathcal{X}}_\ell(q_\ell, q_{c_1}) dq_\ell,$$

and thus condition (4.8) gives

$$\frac{\partial \tilde{\mathcal{X}}_\ell}{\partial q_{c_1}} = 0,$$

and hence $\tilde{\mathcal{X}}_\ell(q_\ell, q_{c_1}) = \tilde{\mathcal{X}}_\ell(q_\ell)$. The Dirac–Hamilton–Jacobi equation (4.9) then becomes

$$\frac{1}{2} \ell \tilde{\mathcal{X}}_\ell(q_\ell)^2 + \frac{1}{2} \frac{q_{c_1}^2}{c_1} + \frac{1}{2} \frac{(q_\ell - a_0)^2}{c_2} + \frac{1}{2} \frac{(q_\ell - q_{c_1} - a_0 - a_1)^2}{c_3} = E. \tag{4.10}$$

We impose the condition that $\mathcal{X}_\ell = 0$ when $q_\ell = q_{c_1} = 0$ and $E = 0$, which corresponds to the case where nothing is happening in the circuit. Then, we have

$$\frac{a_0^2}{c_2} + \frac{(a_0 + a_1)^2}{c_3} = 0,$$

which gives $a_0 = a_1 = 0$, since c_2 and c_3 are both positive. Therefore, Eq. (4.10) becomes

$$\frac{1}{2} \ell \tilde{\mathcal{X}}_\ell(q_\ell)^2 + \frac{1}{2} \frac{q_{c_1}^2}{c_1} + \frac{1}{2} \frac{q_\ell^2}{c_2} + \frac{1}{2} \frac{(q_\ell - q_{c_1})^2}{c_3} = E. \quad (4.11)$$

Taking the derivative with respect to q_{c_1} of both sides and solving for q_{c_1} , we have

$$q_{c_1} = \frac{c_1}{c_1 + c_3} q_\ell.$$

Substituting this into Eq. (4.11) gives

$$\frac{1}{2} \left(\ell \tilde{\mathcal{X}}_\ell(q_\ell)^2 + \frac{c_1 + c_2 + c_3}{c_2(c_1 + c_3)} q_\ell^2 \right) = E.$$

Solving for $\tilde{\mathcal{X}}_\ell(q_\ell)$, we obtain

$$\mathcal{X}_\ell(q) = \tilde{\mathcal{X}}_\ell(q_\ell) = \pm \sqrt{\frac{1}{\ell} \left(2E - \frac{c_1 + c_2 + c_3}{c_2(c_1 + c_3)} q_\ell^2 \right)}.$$

Taking the positive root, Eq. (4.3) for q_ℓ gives

$$\dot{q}_\ell = \sqrt{\frac{1}{\ell} \left(2E - \frac{c_1 + c_2 + c_3}{c_2(c_1 + c_3)} q_\ell^2 \right)},$$

which can be solved easily:

$$q_\ell(t) = \sqrt{\frac{2E}{\ell \nu^2}} \sin(\nu t + \alpha),$$

where

$$\nu := \sqrt{\frac{c_1 + c_2 + c_3}{c_2(c_1 + c_3)} \ell}$$

and α is a phase constant to be determined by the initial condition.

Remark 4.5. In the conventional LC circuit theory, one often simplifies problems by “combining” capacitors. Using this technique, the above example simplifies to an LC circuit with an inductor with inductance ℓ and a single capacitance C , that satisfies the following equation:

$$\frac{1}{C} = \frac{1}{c_2} + \frac{1}{c_1 + c_3},$$

which gives

$$C = \frac{c_2(c_1 + c_3)}{c_1 + c_2 + c_3}.$$

Then, the equation for the current $i_\ell := \dot{q}_\ell$ is given by

$$\ell \frac{d^2 i_\ell}{dt^2} + \frac{i_\ell}{C} = 0,$$

or

$$\frac{d^2 i_\ell}{dt^2} + \nu i_\ell = 0,$$

with

$$\nu = \frac{1}{\sqrt{\ell C}} = \sqrt{\frac{c_1 + c_2 + c_3}{c_2(c_1 + c_3)} \ell},$$

which coincides the one defined above. The general solution of the above ODE is

$$i_\ell(t) = \dot{q}_\ell(t) = A \sin(\nu t + \alpha)$$

for some constants A and α . Therefore, our solution is consistent with the conventional theory.

4.4. Applications to Degenerate Lagrangian System with Nonholonomic Constraints.

Example 4.6 (Simplified Roller Racer; see Example 3.4). The submanifold $\mathcal{K} \subset TQ \oplus T^*Q$ is given by

$$\mathcal{K} = \left\{ (q, v, p) \in TQ \oplus T^*Q \mid v_x = \cos \theta \csc \phi [(d_1 \cos \phi + d_2) v_\theta + d_2 v_\phi], \right. \\ \left. v_y = \sin \theta \csc \phi [(d_1 \cos \phi + d_2) v_\theta + d_2 v_\phi], p_x = m_1 v_x, p_y = m_1 v_y, p_\theta = I_1 v_\theta, p_\phi = 0 \right\},$$

and the generalized energy constrained to \mathcal{K} is

$$\mathcal{E}|_{\mathcal{K}} = \frac{1}{2} m_1 \csc^2 \phi [(d_1 \cos \phi + d_2) v_\theta + d_2 v_\phi]^2 + \frac{1}{2} I_1 v_\theta^2.$$

The distribution Δ_Q is easily shown to be completely nonholonomic, and thus we may use the Dirac–Hamilton–Jacobi equation (4.5), which gives

$$\frac{1}{2} m_1 \csc^2 \phi [(d_1 \cos \phi + d_2) \mathcal{X}_\theta(q) + d_2 \mathcal{X}_\phi(q)]^2 + \frac{1}{2} I_1 \mathcal{X}_\theta(q)^2 = E. \quad (4.12)$$

Now, we assume the following ansatz⁵:

$$\mathcal{X}_\theta(x, y, \theta, \phi) = \mathcal{X}_\theta(\theta, \phi), \quad \mathcal{X}_\phi(x, y, \theta, \phi) = \mathcal{X}_\phi(\phi). \quad (4.13)$$

However, substituting them into Eq. (4.12) and solving for \mathcal{X}_θ shows that \mathcal{X}_θ does not depend on θ either; hence we set $\mathcal{X}_\theta(\theta, \phi) = \mathcal{X}_\theta(\phi)$. Then, solving Eq. (4.12) for \mathcal{X}_ϕ , we have

$$\mathcal{X}_\phi(\phi) = \frac{-(d_1 \cos \phi + d_2) \mathcal{X}_\theta(\phi) \pm \sin \phi \sqrt{2E - I_1 \mathcal{X}_\theta(\phi)^2}}{\sqrt{m_1} d_2}. \quad (4.14)$$

Substituting the first solution into condition (4.2), we have

$$\left[(d_1 \cos \phi + d_2) \mathcal{X}_\theta(\phi) - \sin \phi \sqrt{2E - I_1 \mathcal{X}_\theta(\phi)^2} \right] \mathcal{X}'_\theta(\phi) = 0.$$

We choose $\mathcal{X}'_\theta(\phi) = 0$ and hence

$$\mathcal{X}_\theta(\phi) = v_\theta,$$

where v_θ is the initial angular velocity in the θ -direction. This is consistent with the Lagrange–Dirac equations (3.5), which give $\ddot{\theta} = 0$. Substituting this into the first case of Eq. (4.14), we obtain

$$\mathcal{X}_\phi(\phi) = -v_\theta \left(1 + \frac{d_1}{d_2} \cos \phi \right) + \frac{v_r}{d_2} \sin \phi,$$

where $v_r := \sqrt{(2E - I_1 v_\theta^2)/m_1}$.

Then, the condition $\mathcal{X}(q) \in \Delta_Q(q)$ gives the other components of the vector field \mathcal{X} , and hence Eq. (4.3) gives

$$\dot{x} = v_r \cos \theta, \quad \dot{y} = v_r \sin \theta, \\ \dot{\theta} = 0, \quad \dot{\phi} = -v_\theta \left(1 + \frac{d_1}{d_2} \cos \phi \right) + \frac{v_r}{d_2} \sin \phi.$$

We can solve the last equation by separation of variables, and the rest is explicitly solvable.

4.5. Lagrangians that are Linear in Velocity. There are some physical systems, such as point vortices (see, e.g., Chapman [11] and Newton [41]), which are described by Lagrangians that are linear in velocity, i.e.,

$$L(q, \dot{q}) = \langle \alpha(q), \dot{q} \rangle - h(q), \quad (4.15)$$

where α is a one-form on Q . The Lagrangian is clearly degenerate and Lagrange–Dirac equation (2.3) gives the following equations of motion (see Rowley and Marsden [45] and Yoshimura and Marsden [54]):

$$-i_{\mathcal{X}} d\alpha = dh, \quad (4.16)$$

where \mathcal{X} is a vector field on Q ; hence the Lagrange–Dirac equation (2.3) reduces to the first-order dynamics $\dot{q} = \mathcal{X}(q)$ defined on Q .

⁵The (x, y) -dependence is eliminated because we expect that the vector field \mathcal{X} to be \mathbb{R}^2 -translational invariant since the system possesses \mathbb{R}^2 -symmetry.

Now, the assumption in (4.1) of Theorem 4.1 implies $\gamma(q) = \alpha(q)$ and thus

$$\mathcal{E} \circ \Upsilon(q) = h(q);$$

so the Dirac–Hamilton–Jacobi equation (4.4) gives

$$h(q) = E,$$

which simply defines a constant energy surface of the dynamics on Q , i.e., the Dirac–Hamilton–Jacobi equation (4.5) does not give any information on the dynamics on Q . This is because the original dynamics, which is naturally defined on Q with the one-form α and the function h , is somewhat artificially lifted to the tangent bundle TQ through the linear Lagrangian (4.15). In fact, for point vortices on the plane, one has $Q = \mathbb{R}^2$, and the two-form $-d\alpha$ is a symplectic form; hence $Q = \mathbb{R}^2$ is a symplectic manifold and Eq. (4.16) defines a Hamiltonian system on Q with the Hamiltonian h . Therefore, for such systems, the Hamilton–Jacobi equation should be naturally formulated on Q , not on the Pontryagin bundle $TQ \oplus T^*Q$.

5. HAMILTON–JACOBI THEORY FOR WEAKLY DEGENERATE CHAPLYGIN SYSTEMS

In this section, we first show that a weakly Chaplygin system introduced in Section 3.2 reduces to an almost Hamiltonian system on $T^*\bar{Q}$ with a reduced Hamiltonian $\bar{H} : T^*\bar{Q} \rightarrow \mathbb{R}$. Accordingly, we may consider a variant of the nonholonomic Hamilton–Jacobi equation [28, 42] for the reduced system, which we would like to call the *reduced Dirac–Hamilton–Jacobi equation*. We then show an explicit formula that maps solutions of the reduced Dirac–Hamilton–Jacobi equation to those of the original one. Thus, one may solve the reduced Dirac–Hamilton–Jacobi equation, which is simpler than the original one, and then construct solutions of the original Dirac–Hamilton–Jacobi equation by the formula.

5.1. The Geometry of Weakly Degenerate Chaplygin Systems. For weakly degenerate Chaplygin systems, the geometric structure introduced in Section 3.1 is carried over to the Hamiltonian side. Specifically, we define the horizontal lift $\text{hl}_q^P : T_{\bar{q}}^* \bar{Q} \rightarrow P_q$ by (see Ehlers et al. [19])

$$\text{hl}_q^P := \mathbb{F}L_q \circ \text{hl}_q^\Delta \circ (\mathbb{F}\bar{L})_{\bar{q}}^{-1},$$

or by requiring that the diagram below commutes.

$$\begin{array}{ccc} \Delta_Q(q) & \xrightarrow{\mathbb{F}L_q} & P_q \\ \text{hl}_q^\Delta \uparrow & & \uparrow \text{hl}_q^P \\ T_{\bar{q}}^* \bar{Q} & \xleftarrow{(\mathbb{F}\bar{L})_{\bar{q}}^{-1}} & T_{\bar{q}}^* \bar{Q} \end{array}$$

It is easy to show that the following equality holds for the pairing between the two horizontal lifts (see Lemma A.1 in Ohsawa et al. [43]): For any $\alpha_{\bar{q}} \in T_{\bar{q}}^* \bar{Q}$ and $v_{\bar{q}} \in T_{\bar{q}} \bar{Q}$,

$$\langle \text{hl}_q^P(\alpha_{\bar{q}}), \text{hl}_q^\Delta(v_{\bar{q}}) \rangle = \langle \alpha_{\bar{q}}, v_{\bar{q}} \rangle. \quad (5.1)$$

We also define a map $\text{hl}_q^{\mathcal{K}} : T_{\bar{q}}^* \bar{Q} \rightarrow \mathcal{K}_q \subset T_q Q \oplus T_q^* Q$ by

$$\text{hl}_q^{\mathcal{K}} := \left(\text{hl}_q^\Delta \circ (\mathbb{F}\bar{L})_{\bar{q}}^{-1} \right) \oplus \text{hl}_q^P.$$

Since the reduced Lagrangian \bar{L} is non-degenerate, we can also define the reduced Hamiltonian⁶ $\bar{H} : T^*\bar{Q} \rightarrow \mathbb{R}$ as follows:

$$\bar{H}(p_{\bar{q}}) := \langle p_{\bar{q}}, v_{\bar{q}} \rangle - \bar{L}(v_{\bar{q}}), \quad (5.2)$$

with $v_{\bar{q}} = (\mathbb{F}\bar{L})_{\bar{q}}^{-1}(p_{\bar{q}})$.

Lemma 5.1. *The generalized energy $\mathcal{E} : TQ \oplus T^*Q \rightarrow \mathbb{R}$ and the reduced Hamiltonian \bar{H} are related as follows:*

$$\mathcal{E} \circ \text{hl}^{\mathcal{K}} = \bar{H}.$$

⁶Recall that we cannot define a Hamiltonian $H : T^*Q \rightarrow \mathbb{R}$ for the original system because the original Lagrangian $L : TQ \rightarrow \mathbb{R}$ is degenerate.

Proof. Follows from the following simple calculation: For an arbitrary $\alpha_{\bar{q}} \in T_{\bar{q}}^* \bar{Q}$, let $q \in Q$ be a point such that $\pi(q) = \bar{q}$. Then, we obtain

$$\begin{aligned} \mathcal{E} \circ \text{hl}_q^{\mathcal{K}}(\alpha_{\bar{q}}) &= \left\langle \text{hl}_q^P(\alpha_{\bar{q}}), \text{hl}_q^\Delta \circ (\mathbb{F}\bar{L})_{\bar{q}}^{-1}(\alpha_{\bar{q}}) \right\rangle - L \circ \text{hl}_q^\Delta \circ (\mathbb{F}\bar{L})_{\bar{q}}^{-1}(\alpha_{\bar{q}}) \\ &= \left\langle \alpha_{\bar{q}}, (\mathbb{F}\bar{L})_{\bar{q}}^{-1}(\alpha_{\bar{q}}) \right\rangle - \bar{L} \circ (\mathbb{F}\bar{L})_{\bar{q}}^{-1}(\alpha_{\bar{q}}) \\ &= \bar{H}(\alpha_{\bar{q}}), \end{aligned}$$

where we used Eq. (5.1) and the definition of \bar{H} in Eq. (5.2). \square

Furthermore, as shown in Theorem A.4 of the Appendix A (see also Koiller [32], Bates and Sniatycki [2], Cantrijn et al. [7], Hochgerner and García-Naranjo [27]), we have the reduced system

$$i_{\bar{X}} \bar{\Omega}^{\text{nh}} = d\bar{H} \quad (5.3)$$

on $T^* \bar{Q}$ defined with the reduced Hamiltonian \bar{H} and the almost symplectic form

$$\bar{\Omega}^{\text{nh}} := \bar{\Omega} - \Xi, \quad (5.4)$$

where Ξ is the non-closed two-form on $T^* \bar{Q}$ defined in Eq. (A.9).

5.2. Hamilton–Jacobi Theorem for Weakly Degenerate Chaplygin Systems. The previous subsection showed that a weakly degenerate Chaplygin system reduces to a non-degenerate Lagrangian and hence an almost Hamiltonian system (5.3). Moreover, Lemma 5.1 shows how the generalized energy \mathcal{E} is related to the reduced Hamiltonian \bar{H} ; see also the upper half of the diagram (5.5) below. The lower half of the diagram suggests the relationship between the reduced and original Dirac–Hamilton–Jacobi equations alluded above. Specifically, $\bar{\gamma}$ is a one-form on $\bar{Q} := Q/G$ and is a solution of the reduced Dirac–Hamilton–Jacobi equation (5.6) defined below, and the diagram suggests how to define the map $\Upsilon : Q \rightarrow \mathcal{K}$ so that it is a solution of the original Dirac–Hamilton–Jacobi equation (4.4).

$$\begin{array}{ccc} & \mathbb{R} & \\ & \nearrow \varepsilon & \nwarrow \bar{H} \\ \mathcal{K} & \longleftarrow \text{hl}^{\mathcal{K}} & T^* \bar{Q} \\ \Upsilon \uparrow & & \uparrow \bar{\gamma} \\ Q & \xrightarrow{\pi} & \bar{Q} \end{array} \quad (5.5)$$

The whole diagram (5.5) leads us to the following main result of this section:

Theorem 5.2 (Reduced Dirac–Hamilton–Jacobi Equation). *Consider a weakly degenerate Chaplygin system on a connected configuration space Q and assume that the distribution Δ_Q is completely nonholonomic. Let $\bar{\gamma} : \bar{Q} \rightarrow T^* \bar{Q}$ be a one-form on \bar{Q} that satisfies the reduced Dirac–Hamilton–Jacobi equation*

$$\bar{H} \circ \bar{\gamma}(\bar{q}) = E, \quad (5.6)$$

with a constant E , as well as

$$d\bar{\gamma} + \bar{\gamma}^* \Xi = 0, \quad (5.7)$$

where Ξ is the two-form on $T^* \bar{Q}$ that appeared in the definition of the almost symplectic form $\bar{\Omega}^{\text{nh}}$ in Eq. (5.4) (see also Eq. (A.9)). Define $\Upsilon = \mathcal{X} \oplus \gamma : Q \rightarrow \mathcal{K}$ by (see the diagram (5.5))

$$\Upsilon(q) := \text{hl}_q^{\mathcal{K}} \circ \bar{\gamma} \circ \pi(q) = \text{hl}_q^{\mathcal{K}}(\bar{\gamma}(\bar{q})), \quad (5.8)$$

where $\bar{q} := \pi(q)$, i.e.,

$$\mathcal{X}(q) := \text{hl}_q^\Delta \circ (\mathbb{F}\bar{L})_{\bar{q}}^{-1}(\bar{\gamma}(\bar{q})), \quad \gamma(q) := \text{hl}_q^P(\bar{\gamma}(\bar{q})).$$

Then, $\Upsilon = \mathcal{X} \oplus \gamma$ satisfies the Dirac–Hamilton–Jacobi equation (4.5) as well as condition (4.2).

Proof. This proof is very similar to that of Theorem 4.1 in Ohsawa et al. [43].

The diagram (5.5) shows that if the one-form $\bar{\gamma}$ satisfies Eq. (5.6) then the map Υ defined by Eq. (5.8) satisfies the Dirac–Hamilton–Jacobi equation (4.5).

To show that it also satisfies the condition (4.2), we perform the following calculations: Let $Y^h, Z^h \in \mathfrak{X}(Q)$ be arbitrary horizontal vector fields, i.e., $Y_q^h, Z_q^h \in \Delta_Q(q)$ for any $q \in Q$. We start from the following identity:

$$d\gamma(Y^h, Z^h) = Y^h[\gamma(Z^h)] - Z^h[\gamma(Y^h)] - \gamma([Y^h, Z^h]). \quad (5.9)$$

The goal is to show that the right-hand side vanishes. Let us first evaluate the first two terms on the right-hand side of the above identity at an arbitrary point $q \in Q$: Let $Z_{\bar{q}} := T_q\pi_Q(Z_q^h) \in T_{\bar{q}}\bar{Q}$, then $Z_q^h = \text{hl}_q^\Delta(Z_{\bar{q}})$. Thus, we have

$$\begin{aligned} \gamma(Z^h)(q) &= \left\langle \text{hl}_q^P \circ \bar{\gamma}(\bar{q}), \text{hl}_q^\Delta(Z_{\bar{q}}) \right\rangle \\ &= \langle \bar{\gamma}(\bar{q}), Z_{\bar{q}} \rangle \\ &= \bar{\gamma}(Z)(\bar{q}). \end{aligned}$$

Hence, writing $\gamma_Z = \bar{\gamma}(Z)$ for short, we have $\gamma(Z^h) = \gamma_Z \circ \pi$. Therefore, defining $Y_{\bar{q}} := T_q\pi(Y_q^h)$, i.e., $Y_q^h = \text{hl}_q^\Delta(Y_{\bar{q}})$,

$$\begin{aligned} Y^h[\gamma(Z^h)](q) &= Y^h[\gamma_Z \circ \pi](q) \\ &= \langle d(\gamma_Z \circ \pi)_q, Y_q^h \rangle \\ &= \langle (\pi^* d\gamma_Z)_q, Y_q^h \rangle \\ &= \langle d\gamma_Z(\bar{q}), T_q\pi(Y_q^h) \rangle \\ &= \langle d\gamma_Z(\bar{q}), Y_{\bar{q}} \rangle \\ &= Y[\gamma_Z](\bar{q}) \\ &= Y[\bar{\gamma}(Z)](\bar{q}). \end{aligned}$$

Hence, we have

$$Y^h[\gamma(Z^h)] - Z^h[\gamma(Y^h)] = Y[\bar{\gamma}(Z)] - Z[\bar{\gamma}(Y)], \quad (5.10)$$

where we have omitted q and \bar{q} for simplicity.

Now, let us evaluate the last term on the right-hand side of Eq. (5.9): First we would like to decompose $[Y^h, Z^h]_q$ into the horizontal and vertical parts. Since both Y^h and Z^h are horizontal, we have⁷

$$\text{hor}([Y^h, Z^h]_q) = \text{hl}_q^\Delta([Y, Z]_{\bar{q}}),$$

whereas the vertical part is

$$\text{ver}([Y^h, Z^h]_q) = (\mathcal{A}_q([Y^h, Z^h]_q))_Q(q) = -(\mathcal{B}_q(Y_q^h, Z_q^h))_Q(q),$$

where we used the following relation between the connection \mathcal{A} and its curvature \mathcal{B} that holds for horizontal vector fields Y^h and Z^h :

$$\begin{aligned} \mathcal{B}_q(Y_q^h, Z_q^h) &= d\mathcal{A}_q(Y_q^h, Z_q^h) \\ &= Y^h[\mathcal{A}(Z^h)](q) - Z^h[\mathcal{A}(Y^h)](q) - \mathcal{A}([Y^h, Z^h])(q) \\ &= -\mathcal{A}([Y^h, Z^h])(q). \end{aligned}$$

As a result, we have the decomposition

$$[Y^h, Z^h]_q = \text{hl}_q^\Delta([Y, Z]_{\bar{q}}) - (\mathcal{B}_q(Y_q^h, Z_q^h))_Q(q).$$

Therefore,

$$\begin{aligned} \gamma([Y^h, Z^h])(q) &= \left\langle \text{hl}_q^P \circ \bar{\gamma} \circ \pi(q), \text{hl}_q^\Delta([Y, Z]_{\bar{q}}) \right\rangle - \left\langle \text{hl}_q^P \circ \bar{\gamma} \circ \pi(q), (\mathcal{B}_q(Y_q^h, Z_q^h))_Q(q) \right\rangle \\ &= \langle \bar{\gamma}(\bar{q}), [Y, Z]_{\bar{q}} \rangle - \left\langle \mathbf{J} \left(\text{hl}_q^P \circ \bar{\gamma}(\bar{q}) \right), \mathcal{B}_q(Y_q^h, Z_q^h) \right\rangle \\ &= \bar{\gamma}([Y, Z])(\bar{q}) - \bar{\gamma}^* \Xi(Y, Z)(\bar{q}), \end{aligned} \quad (5.11)$$

⁷See, e.g., Kobayashi and Nomizu [31, Proposition 1.3 (3), p. 65].

where the second equality follows from Eq. (5.1) and the definition of the momentum map \mathbf{J} ; the last equality follows from the definition of Ξ in Eq. (A.9): Let $\pi_{\bar{Q}} : T^*\bar{Q} \rightarrow \bar{Q}$ be the cotangent bundle projection; since $\pi_{\bar{Q}} \circ \bar{\gamma} = \text{id}_{\bar{Q}}$ and thus $T\pi_{\bar{Q}} \circ T\bar{\gamma} = \text{id}_{T\bar{Q}}$, we have

$$\begin{aligned} \bar{\gamma}^* \Xi(Y, Z)(\bar{q}) &= \Xi_{\bar{\gamma}(\bar{q})}(T\bar{\gamma}(Y_{\bar{q}}), T\bar{\gamma}(Z_{\bar{q}})) \\ &= \left\langle \mathbf{J} \circ \text{hl}_q^P(\bar{\gamma}(\bar{q})), \mathcal{B}_q \left(\text{hl}_q^\Delta(Y_{\bar{q}}), \text{hl}_q^\Delta(Z_{\bar{q}}) \right) \right\rangle. \end{aligned}$$

Substituting Eqs. (5.10) and (5.11) into Eq. (5.9), we obtain

$$\begin{aligned} d\gamma(Y^h, Z^h) &= Y[\bar{\gamma}(Z)] - Z[\bar{\gamma}(Y)] - \bar{\gamma}([Y, Z])(\bar{q}) + \bar{\gamma}^* \Xi(Y, Z) \\ &= d\bar{\gamma}(Y, Z) + \bar{\gamma}^* \Xi(Y, Z) \\ &= (d\bar{\gamma} + \bar{\gamma}^* \Xi)(Y, Z) = 0. \end{aligned} \quad \square$$

Example 5.3 (Simplified Roller Racer; see Examples 3.4 and 4.6). The Lie algebra \mathfrak{g} of $G = \mathbb{R}^2$ is identified with \mathbb{R}^2 ; let us use (ξ, η) as the coordinates for \mathfrak{g} . Then, we may write the connection $\mathcal{A} : TQ \rightarrow \mathfrak{g}$ as

$$\mathcal{A} = \omega^1 \otimes \frac{\partial}{\partial \xi} + \omega^2 \otimes \frac{\partial}{\partial \eta},$$

where ω^1 and ω^2 are the constraint one-forms defined in Eq. (3.4); hence its curvature is given by

$$\mathcal{B} = -\csc^2 \phi [d_1 \cos \theta + d_2 \cos(\theta + \phi)] d\theta \wedge d\phi \otimes d\xi - \csc^2 \phi [d_1 \sin \theta + d_2 \sin(\theta + \phi)] d\theta \wedge d\phi \otimes d\eta.$$

Furthermore, the momentum map $\mathbf{J} : T^*Q \rightarrow \mathfrak{g}^*$ is given by

$$\mathbf{J}(p_q) = p_x d\xi + p_y d\eta.$$

Therefore, we have

$$\Xi = -p_\phi \left(\frac{d_1}{d_2} + \cos \phi \right) \csc \phi d\theta \wedge d\phi.$$

Since the reduced Lagrangian \bar{L} (see Eq. (3.6)) is non-degenerate, we have the reduced Hamiltonian $\bar{H} : T^*\bar{Q} \rightarrow \mathbb{R}$ given by

$$\bar{H} = \frac{1}{2I_1} \left[p_\theta - \left(1 + \frac{d_1}{d_2} \cos \phi \right) p_\phi \right]^2 + \frac{\sin^2 \phi}{2m_1 d_2^2} p_\phi^2.$$

We assume the ansatz

$$\bar{\gamma}_\theta(\theta, \phi) = \bar{\gamma}_\phi(\phi).$$

Then, the reduced Dirac–Hamilton–Jacobi equation (5.6) gives

$$\frac{1}{2I_1} \left[\bar{\gamma}_\theta(\theta, \phi) - \left(1 + \frac{d_1}{d_2} \cos \phi \right) \bar{\gamma}_\phi(\phi) \right]^2 + \frac{\sin^2 \phi}{2m_1 d_2^2} \bar{\gamma}_\phi(\phi)^2 = E,$$

which implies that $\bar{\gamma}_\theta(\theta, \phi) = \bar{\gamma}_\phi(\phi)$. Solving this for $\bar{\gamma}_\theta(\phi)$ gives

$$\bar{\gamma}_\theta(\phi) = \left(1 + \frac{d_1}{d_2} \cos \phi \right) \bar{\gamma}_\phi(\phi) \pm \sqrt{I_1 \left(2E - \frac{\sin^2 \phi}{m_1 d_2^2} \bar{\gamma}_\phi(\phi)^2 \right)}.$$

Substituting the first case into Eq. (5.7), we obtain

$$\bar{\gamma}'_\phi(\phi) = -\cot \phi \bar{\gamma}_\phi(\phi),$$

which gives

$$\bar{\gamma}_\phi(\phi) = C \csc \phi$$

for some constant C . Therefore,

$$\bar{\gamma}_\theta(\phi) = C \left(1 + \frac{d_1}{d_2} \cos \phi \right) \csc \phi + \sqrt{I_1 \left(2E - \frac{C^2}{m_1 d_2^2} \right)}.$$

It is straightforward to check that, with the choice

$$C = d_2 \sqrt{m_1 (2E - I_1 v_\theta^2)},$$

Eq. (5.8) gives the solution obtained in Example 4.6.

Remark 5.4. Notice that the ansatz we used here is less elaborate compared to the one, Eq. (4.13), used for the Dirac–Hamilton–Jacobi equation without the reduction. Specifically, accounting for the \mathbb{R}^2 -symmetry is not necessary here, since the reduced Dirac–Hamilton–Jacobi equation is defined for the \mathbb{R}^2 -reduced system.

6. CONCLUSION AND FUTURE WORK

Conclusion. We developed Hamilton–Jacobi theory for degenerate Lagrangian systems with holonomic and nonholonomic constraints. In particular, we illustrated, through a few examples, that solutions of the Dirac–Hamilton–Jacobi equation can be used to obtain exact solutions of the equations of motion. Also, motivated by those degenerate Lagrangian systems that appear as simplified models of nonholonomic mechanical systems, we defined a class of degenerate nonholonomic Lagrangian systems that reduce to non-degenerate almost Hamiltonian systems. We then showed that the Dirac–Hamilton–Jacobi equation reduces to the nonholonomic Hamilton–Jacobi equation for the reduced non-degenerate system.

Future Work.

- *Relationship with discrete variational Dirac mechanics.* Hamilton–Jacobi theory has been an important ingredient in discrete mechanics and symplectic integrators from both the theoretical and implementation points of view (see Marsden and West [38, Sections 1.7, 1.8, 4.7, and 4.8] and Channell and Scovel [10]). It is interesting to see if the Dirac–Hamilton–Jacobi equation plays the same role in discrete variational Dirac mechanics of Leok and Ohsawa [36, 37].
- *Hamilton–Jacobi theory for systems with Lagrangians linear in velocity.* As briefly mentioned in Section 4.5, the Dirac–Hamilton–Jacobi equation is not appropriate for those systems with Lagrangians that are linear in velocity. However, Rothe and Scholtz [44, Example 4] illustrate that their formulation of the Hamilton–Jacobi equation can be applied to such systems. We are interested in a possible generalization of our formulation to deal with such systems, and also a link with their formulation.
- *Asymptotic analysis of massless approximation.* Massless approximations for some nonholonomic systems seem to give good approximations to the full formulation. It seems that the nonholonomic constraints “regularize” the otherwise singular perturbation problem, and hence makes the massless approximations viable. We expect that asymptotic analysis will reveal how the perturbation problem becomes regular, particularly for those cases where massless approximations lead to weakly degenerate Chaplygin systems.
- *Hamilton–Jacobi theory for general systems on the Pontryagin bundle.* Section 2.4 naturally leads us to consider systems on the Pontryagin bundle described by an arbitrary Dirac structure. We are interested in this generalization, its corresponding Hamilton–Jacobi theory, and its applications.

ACKNOWLEDGEMENTS

We would like to thank Anthony Bloch, Henry Jacobs, Jerrold Marsden, Joris Vankerschaver, and Hiroaki Yoshimura for their helpful comments, and also Anthony Bloch and Wang Sang Koon for their permission to use their figures. This material is based upon work supported by the National Science Foundation under the applied mathematics grant DMS-0726263, the Faculty Early Career Development (CAREER) award DMS-1010687, MICINN (Spain) grants MTM2009-13383 and MTM2009-08166-E, and the projects of the Canary government SOLSUBC200801000238 and ProID20100210.

APPENDIX A. REDUCTION OF WEAKLY DEGENERATE CHAPLYGIN SYSTEMS

A.1. Constrained Dirac Structure. We may restrict the Dirac structure D_{Δ_Q} to $P \subset T^*Q$ as follows (see Yoshimura and Marsden [53, Section 5.6] and references therein): Let us define a distribution $\mathcal{H} \subset TP$ on P by

$$\mathcal{H} := TP \cap \Delta_{T^*Q}, \quad (\text{A.1})$$

and also, using the inclusion $\iota_P : P \hookrightarrow T^*Q$, define the two-form $\Omega_P := \iota_P^* \Omega$ on P . Then, define the *constrained Dirac structure* $D_P \subset TP \oplus T^*P$, for each $z \in P$, by

$$D_P(z) := \left\{ (v_z, \alpha_z) \in T_z P \oplus T_z^* P \mid v_z \in \mathcal{H}_z, \alpha_z - \Omega_P^b(z)(v_z) \in \mathcal{H}_z^\circ \right\},$$

where $\Omega_P^b : TP \rightarrow T^*P$ is the flat map induced by Ω_P . Then, we have the *constrained Lagrange–Dirac system* defined by

$$(X_P, \mathfrak{D}L_c) \in D_P, \quad (\text{A.2})$$

where X_P is a vector field on P , $L_c := L|_{\Delta_Q}$ the constrained Lagrangian, and $\mathfrak{D}L_c(u) := \mathfrak{D}L(u)|_{TP}$ for any $u \in \Delta_Q$.

If the constrained Lagrangian L_c is non-degenerate, i.e., the partial Legendre transformation $\mathbb{F}L|_{\Delta_Q} : \Delta_Q \rightarrow P$ is invertible, then we may define the constrained Hamiltonian $H_P : P \rightarrow \mathbb{R}$ [54] by

$$H_P(p_q) := \langle p_q, v_q \rangle - L_c(v_q),$$

where $v_q := (\mathbb{F}L|_{\Delta_Q})^{-1}(p_q)$. Then, the constrained Lagrange–Dirac system (A.2), is equivalent to the *constrained implicit Hamiltonian system* defined by

$$(X_P, dH_P) \in D_P. \quad (\text{A.3})$$

Remark A.1. Let

$$\Omega_{\mathcal{H}} := \Omega_P|_{\mathcal{H}} \quad (\text{A.4})$$

be the restriction of Ω_P to $\mathcal{H} \subset TP$ and hence a skew-symmetric bilinear form in \mathcal{H} . If $\Omega_{\mathcal{H}}$ is non-degenerate, then Eq. (A.3) gives

$$i_{X_P}\Omega_{\mathcal{H}} = dH_P|_{\mathcal{H}},$$

which is nonholonomic Hamilton’s equations of Bates and Sniatycki [2] (see also Koon and Marsden [33]).

A.2. Reduction of Constrained Dirac Structure. Let us now show how to reduce the constrained Dirac structure D_P to a Dirac structure on $T^*\bar{Q}$. This special case of Dirac reduction to follow gives a Dirac point of view on the nonholonomic reduction of Koiller [32], and hence provides a natural framework for the reduction of weakly degenerate Chaplygin systems. See Yoshimura and Marsden [55] for reduction of Dirac structures without constraints, Jotz and Ratiu [30] for the relationship between Dirac and nonholonomic reduction of Bates and Sniatycki [2]; see also Cantrijn et al. [6, 7] for a theory of reducing degenerate Lagrangian systems to non-degenerate ones.

Let $\Phi : G \times Q \rightarrow Q$ be the action of the Lie group G given in Definition 3.1 and $T^*\Phi_{g^{-1}} : T^*Q \rightarrow T^*Q$ be its cotangent lift defined by

$$\langle T^*\Phi_{g^{-1}}(\alpha), v \rangle = \langle \alpha, T\Phi_{g^{-1}}(v) \rangle.$$

It is easy to show that the G -symmetries of the Lagrangian L and the distribution Δ_Q imply that the submanifold $P \subset T^*Q$ is invariant under the action of the cotangent lift. Hence, we may restrict the action to P and define $\Phi^P : G \times P \rightarrow P$, i.e., $\Phi_g^P : P \rightarrow P$ by $\Phi_g^P := T^*\Phi_{g^{-1}}|_P$ for any $g \in G$. This gives rise to the principal bundle

$$\pi_G^P : P \rightarrow P/G.$$

The geometric structure of weakly degenerate Chaplygin systems summarized in Section 5.1 gives rise to a diffeomorphism $\varphi : T^*\bar{Q} \rightarrow P/G$; this then induces the map $\rho : P \rightarrow T^*\bar{Q}$ so that the diagram below commutes (see Hochgerner and García-Naranjo [27]).

$$\begin{array}{ccc} P & & \\ \pi_G^P \downarrow & \searrow \rho & \\ P/G & \xrightarrow{\varphi^{-1}} & T^*\bar{Q} \end{array} \quad (\text{A.5})$$

Furthermore, the principal connection $\mathcal{A} : TQ \rightarrow \mathfrak{g}$ defined in Eq. (3.1) induces the principal connection $\mathcal{A}_P : TP \rightarrow \mathfrak{g}$ defined by

$$\mathcal{A}_P := \mathcal{A} \circ T\pi_Q \circ T\iota_P,$$

and the horizontal space for this principal connection is $\mathcal{H} \subset TP$ defined in Eq. (A.1), i.e., $\mathcal{H} = \ker \mathcal{A}_P$ [27]. Therefore, writing $[z] := \pi_G^P(z) \in P/G$, we have the horizontal lift

$$\text{hl}_z^{\mathcal{H}} : T_{[z]}(P/G) \rightarrow \mathcal{H}_z; \quad v_{[z]} \mapsto (T_z\pi_G^P|_{\mathcal{H}_z})^{-1}(v_{[z]}).$$

Then, clearly the following diagram commutes:

$$\begin{array}{ccc}
 & T_z P & \\
 \text{hl}_{\bar{z}}^{\mathcal{H}} \uparrow & & \searrow T_z \rho \\
 T_{[z]}(P/G) & \xrightarrow{T_{[z]}\varphi^{-1}} & T_{\bar{z}} T^* \bar{Q}
 \end{array} \tag{A.6}$$

where $\bar{z} := \varphi^{-1}([z]) \in T^* \bar{Q}$.

Lemma A.2. *The two-form Ω_P is invariant under the G -action, i.e., for any $g \in G$,*

$$(\Phi_g^P)^* \Omega_P = \Omega_P. \tag{A.7}$$

Proof. Using the relation $T^* \Phi_{g^{-1}} \circ \iota_P = \iota_P \circ \Phi_g^P$, we have

$$\begin{aligned}
 (\Phi_g^P)^* \Omega_P &= (\Phi_g^P)^* \iota_P^* \Omega \\
 &= (\iota_P \circ \Phi_g^P)^* \Omega \\
 &= (T^* \Phi_{g^{-1}} \circ \iota_P)^* \Omega \\
 &= (\iota_P)^* \circ (T^* \Phi_{g^{-1}})^* \Omega \\
 &= \iota_P^* \Omega \\
 &= \Omega_P,
 \end{aligned}$$

where we used the fact that the cotangent lift $T^* \Phi_{g^{-1}}$ is symplectic. \square

Now, consider the action of G on the Whitney sum $TP \oplus T^*P$ defined by

$$\Psi : G \times (TP \oplus T^*P) \rightarrow TP \oplus T^*P; \quad (g, (v_z, \alpha_z)) \mapsto \left(T_z \Phi_g^P(v_z), T_{gz}^* \Phi_{g^{-1}}^P(\alpha_z) \right) =: (g \cdot v_z, g \cdot \alpha_z).$$

Then, we have the following:

Proposition A.3. *The constrained Dirac structure D_P is invariant under the action Ψ defined above.*

Proof. Let $z \in P$ be arbitrary and $(v_z, \alpha_z) \in D_P(z)$. Then, $v_z \in \mathcal{H}_z$ and $\alpha_z - \Omega_P^{\flat}(v_z) \in \mathcal{H}_z^{\circ}$. Now, the G -invariance of $\mathcal{H} = \ker \mathcal{A}_P$ implies $T\Phi_g(v_z) \in \mathcal{H}_{gz}$. Also, for any $w_{gz} = T_z \Phi_g^P(w_z) \in \mathcal{H}_{gz}$ with $w_z \in \mathcal{H}_z$, we have

$$\begin{aligned}
 \left\langle T_{gz}^* \Phi_{g^{-1}}^P(\alpha_z) - \Omega_P^{\flat}(T_z \Phi_g^P(v_z)), w_{gz} \right\rangle &= \left\langle \alpha_z, T_{gz} \Phi_{g^{-1}}^P(w_{gz}) \right\rangle - \Omega_P(T_z \Phi_g^P(v_z), T_z \Phi_g^P(w_z)) \\
 &= \langle \alpha_z, w_z \rangle - \Omega_P(T_z \Phi_g^P(v_z), T_z \Phi_g^P(w_z)) \\
 &= \langle \alpha_z, w_z \rangle - (\Phi_g^P)^* \Omega_P(v_z, w_z) \\
 &= \langle \alpha_z, w_z \rangle - \Omega_P(v_z, w_z) \\
 &= \left\langle \alpha_z - \Omega_P^{\flat}(v_z), w_z \right\rangle \\
 &= 0,
 \end{aligned}$$

where the fourth line follows from Eq. (A.7). Hence

$$(g \cdot v_z, g \cdot \alpha_z) = \left(T_z \Phi_g^P(v_z), T_{gz}^* \Phi_{g^{-1}}^P(\alpha_z) \right) \in D_P(gz),$$

and thus the claim follows. \square

Now, the main result in this section is the following:

Theorem A.4. *The reduced constrained Dirac structure $[D_P]_G := D_P/G$ is identified with the Dirac structure \bar{D} on $T^* \bar{Q}$ defined, for any $\bar{z} \in T^* \bar{Q}$, by*

$$\bar{D}(\bar{z}) := \left\{ (v_{\bar{z}}, \alpha_{\bar{z}}) \in T_{\bar{z}} T^* \bar{Q} \oplus T_{\bar{z}}^* T^* \bar{Q} \mid \alpha_{\bar{z}} = (\bar{\Omega}^{\text{nh}})^{\flat}(v_{\bar{z}}) \right\}, \tag{A.8}$$

where $\bar{\Omega}^{\text{nh}} = \bar{\Omega} - \Xi$ with $\bar{\Omega}$ being the standard symplectic form on $T^*\bar{Q}$, and the two-form Ξ on $T^*\bar{Q}$ is defined as follows: For any $\alpha_{\bar{q}} \in T_{\bar{q}}^*\bar{Q}$ and $\mathcal{Y}_{\alpha_{\bar{q}}}, \mathcal{Z}_{\alpha_{\bar{q}}} \in T_{\alpha_{\bar{q}}}T^*\bar{Q}$, let $Y_{\bar{q}} := T\pi_{\bar{Q}}(\mathcal{Y}_{\alpha_{\bar{q}}})$ and $Z_{\bar{q}} := T\pi_{\bar{Q}}(\mathcal{Z}_{\alpha_{\bar{q}}})$ where $\pi_{\bar{Q}} : T^*\bar{Q} \rightarrow \bar{Q}$ is the cotangent bundle projection, and then set

$$\Xi_{\alpha_{\bar{q}}}(\mathcal{Y}_{\alpha_{\bar{q}}}, \mathcal{Z}_{\alpha_{\bar{q}}}) := \left\langle \mathbf{J} \circ \text{hl}_q^P(\alpha_{\bar{q}}), \mathcal{B}_q\left(\text{hl}_q^\Delta(Y_{\bar{q}}), \text{hl}_q^\Delta(Z_{\bar{q}})\right) \right\rangle, \quad (\text{A.9})$$

where $\mathbf{J} : T^*Q \rightarrow \mathfrak{g}^*$ is the momentum map corresponding to the G -action, and \mathcal{B} is the curvature two-form of the connection \mathcal{A} .

Lemma A.5. Define, for any $z \in P$,

$$f_z : T_z P \oplus T_z^* P \rightarrow T_z T^*\bar{Q} \oplus T_z^* T^*\bar{Q}; \quad f_z(v_z, \alpha_z) = \left(T_z \rho(v_z), T_z^* \varphi \circ (\text{hl}_z^{\mathcal{H}})^*(\alpha_z) \right),$$

where $(\text{hl}_z^{\mathcal{H}})^* : T_z^* P \rightarrow T_{[z]}^*(P/G)$ is the adjoint map of $\text{hl}_z^{\mathcal{H}}$. Then, f is G -invariant, i.e., $f \circ \Psi_g = f$ for any $g \in G$.

Remark A.6. The map $f_z|_{D_P(z)}$, i.e., f_z defined above restricted to $D_P(z) \subset T_z P \oplus T_z^* P$, is the backward Dirac map (see Bursztyn and Radko [5]) of

$$\phi_z := \text{hl}_z^{\mathcal{H}} \circ T_z \varphi : T_z T^*\bar{Q} \rightarrow T_z P,$$

that is, $f_z = \mathcal{B}\phi_z$ using the notation in [5]; hence the image $f(D_P) \subset TT^*\bar{Q} \oplus T^*T^*\bar{Q}$ is a Dirac structure.

Proof of Lemma A.5. Let $(v_z, \alpha_z) \in T_z P \oplus T_z^* P$ and $(\tilde{v}_{gz}, \tilde{\alpha}_{gz}) := \Psi_g(v_z, \alpha_z)$ for $g \in G$, i.e.,

$$\tilde{v}_{gz} = T_z \Phi_g^P(v_z), \quad \tilde{\alpha}_{gz} = T_{gz}^* \Phi_{g^{-1}}^P(\alpha_z).$$

Using the identities $\rho = \varphi^{-1} \circ \pi_G^P$ (see diagram (A.5)) and $\pi_G^P \circ \Phi_g^P = \pi_G^P$, we have

$$\begin{aligned} T_{gz} \rho(\tilde{v}_{gz}) &= T_{gz} \rho \circ T_z \Phi_g^P(v_z) \\ &= T_{[z]} \varphi^{-1} \circ T_{gz} \pi_G^P \circ T_z \Phi_g^P(v_z) \\ &= T_{[z]} \varphi^{-1} \circ T_z (\pi_G^P \circ \Phi_g^P)(v_z) \\ &= T_{[z]} \varphi^{-1} \circ T_z \pi_G^P(v_z) \\ &= T_z \rho(v_z). \end{aligned}$$

On the other hand, for any $w_{[z]} \in T_{[z]}(P/G)$,

$$\begin{aligned} (\text{hl}_{gz}^{\mathcal{H}})^*(\tilde{\alpha}_{gz}) &= (\text{hl}_{gz}^{\mathcal{H}})^* \circ T_{gz}^* \Phi_{g^{-1}}^P(\alpha_z) \\ &= \left(T_{gz} \Phi_{g^{-1}}^P \circ \text{hl}_{gz}^{\mathcal{H}} \right)^*(\alpha_z) \\ &= (\text{hl}_z^{\mathcal{H}})^*(\alpha_z), \end{aligned}$$

because of the invariance property of the horizontal lift $\text{hl}^{\mathcal{H}}$, i.e., $T_{gz} \Phi_{g^{-1}}^P \circ \text{hl}_{gz}^{\mathcal{H}} = \text{hl}_z^{\mathcal{H}}$. Hence it follows that $f_{gz} \circ \Psi_g(v_z, \alpha_z) = f_{gz}(\tilde{v}_{gz}, \tilde{\alpha}_{gz}) = f_z(v_z, \alpha_z)$. \square

Proof of Theorem A.4. Lemma A.5 implies that the map $f|_{D_P}$ defined above induces the following well-defined map:

$$\bar{f} : [D_P]_G \rightarrow TT^*\bar{Q} \oplus T^*T^*\bar{Q}; \quad [(v_z, \alpha_z)]_G \mapsto \left(T_z \rho(v_z), T_z^* \varphi \circ (\text{hl}_z^{\mathcal{H}})^*(\alpha_z) \right),$$

i.e., the diagram below commutes.

$$\begin{array}{ccc} D_P & & \\ \downarrow /G & \searrow f|_{D_P} & \\ [D_P]_G & \xrightarrow{\bar{f}} & \bar{D} \end{array}$$

Let us look into the image $\bar{D} := \bar{f}([D_P]_G)$. Notice first that

$$T_z \rho(\mathcal{H}_z) = T_{[z]} \varphi^{-1} \circ T_z \pi_G^P(\mathcal{H}_z) = T_z T^*\bar{Q},$$

since $T_z \pi_G^P(\mathcal{H}_z) = T_{[z]}(P/G)$ and $T_{[z]}\varphi^{-1}$ is surjective.

On the other hand, notice that $w_z^h := \text{hl}_z^{\mathcal{H}} \circ T_{\bar{z}}\varphi(w_{\bar{z}})$ is in \mathcal{H}_z for any $w_{\bar{z}} \in T_{\bar{z}}T^*\bar{Q}$, whereas $\alpha_z - \Omega_P^b(v_z) \in \mathcal{H}_z^c$. So we have

$$\begin{aligned} 0 &= \left\langle \alpha_z - \Omega_P^b(v_z), \text{hl}_z^{\mathcal{H}} \circ T_{\bar{z}}\varphi(w_{\bar{z}}) \right\rangle \\ &= \left\langle T_{\bar{z}}^*\varphi \circ (\text{hl}_z^{\mathcal{H}})^* \alpha_z - T_{\bar{z}}^*\varphi \circ (\text{hl}_z^{\mathcal{H}})^* \Omega_P^b(v_z), w_{\bar{z}} \right\rangle. \end{aligned}$$

Therefore,

$$T_{\bar{z}}^*\varphi \circ (\text{hl}_z^{\mathcal{H}})^* \alpha_z = T_{\bar{z}}^*\varphi \circ (\text{hl}_z^{\mathcal{H}})^* \Omega_P^b(v_z).$$

However, for an arbitrary $w_{\bar{z}} \in T_{\bar{z}}T^*\bar{Q}$,

$$\begin{aligned} \left\langle T_{\bar{z}}^*\varphi \circ (\text{hl}_z^{\mathcal{H}})^* \Omega_P^b(v_z), w_{\bar{z}} \right\rangle &= \Omega_P \left(v_z, \text{hl}_z^{\mathcal{H}} \circ T_{\bar{z}}\varphi(w_{\bar{z}}) \right) \\ &= \Omega_{\mathcal{H}} \left(v_z, \text{hl}_z^{\mathcal{H}} \circ T_{\bar{z}}\varphi(w_{\bar{z}}) \right) \\ &= \rho^* \bar{\Omega}^{\text{nh}} \left(v_z, \text{hl}_z^{\mathcal{H}} \circ T_{\bar{z}}\varphi(w_{\bar{z}}) \right) \\ &= \bar{\Omega}^{\text{nh}} \left(T_z \rho(v_z), T_z \rho \circ \text{hl}_z^{\mathcal{H}} \circ T_{\bar{z}}\varphi(w_{\bar{z}}) \right) \\ &= \bar{\Omega}^{\text{nh}}(T_z \rho(v_z), w_{\bar{z}}) \\ &= \left\langle (\bar{\Omega}^{\text{nh}})^b \circ T_z \rho(v_z), w_{\bar{z}} \right\rangle, \end{aligned}$$

where the second line follows from the definition of $\Omega_{\mathcal{H}}$, Eq. (A.4), since $(v_z, \alpha_z) \in D_P(z)$ implies $v_z \in \mathcal{H}_z$; the third line follows from $\rho^* \bar{\Omega}^{\text{nh}}|_{\mathcal{H}} = \Omega_{\mathcal{H}}$ (see [27, Proposition 2.2]); the fifth from diagram (A.6). As a result, we have

$$T_{\bar{z}}^*\varphi \circ (\text{hl}_z^{\mathcal{H}})^* \alpha_z = (\bar{\Omega}^{\text{nh}})^b \circ T_z \rho(v_z),$$

and thus

$$\bar{f}([(v_z, \alpha_z)]_G) = f(v_z, \alpha_z) = \left(T_z \rho(v_z), (\bar{\Omega}^{\text{nh}})^b \circ T_z \rho(v_z) \right).$$

Since $T_z \rho(\mathcal{H}_z) = T_{\bar{z}}T^*\bar{Q}$, the image $\bar{D} = \bar{f}([D_P]_G) = f(D_P)$ is given by Eq. (A.8). \square

A.3. Reduction of Weakly Degenerate Chaplygin Systems. Reduced dynamics of the constrained implicit Hamiltonian system, Eq. (A.3), for weakly Chaplygin systems follows easily from Theorem A.4: For weakly Chaplygin systems, it is straightforward to show that the constrained Hamiltonian H_P is related to the reduced Hamiltonian defined in Eq. (5.2) as follows:

$$\bar{H} = H_P \circ \text{hl}^P, \quad H_P = \bar{H} \circ \rho, \quad (\text{A.10})$$

and also that if $(X_P, dH_P) \in D_P$ then, defining $\bar{X}(\bar{z}) := T_z \rho \cdot X_P(z)$, we have

$$f(X_P(z), dH_P(z)) = (\bar{X}(\bar{z}), d\bar{H}(\bar{z})),$$

because, using $\text{hl}_z^{\mathcal{H}} \circ T_{\bar{z}}\varphi = (T_z \rho|_{\mathcal{H}_z})^{-1}$ (see diagram (A.6)) and Eq. (A.10), for any $v_{\bar{z}} \in T_{\bar{z}}T^*\bar{Q}$,

$$\begin{aligned} \left\langle T_{\bar{z}}^*\varphi \circ (\text{hl}_z^{\mathcal{H}})^* dH_P(z), v_{\bar{z}} \right\rangle &= \left\langle dH_P(z), \text{hl}_z^{\mathcal{H}} \circ T_{\bar{z}}\varphi(v_{\bar{z}}) \right\rangle \\ &= \left\langle dH_P(z), (T_z \rho|_{\mathcal{H}_z})^{-1}(v_{\bar{z}}) \right\rangle \\ &= \left\langle \rho^* d\bar{H}(z), (T_z \rho|_{\mathcal{H}_z})^{-1}(v_{\bar{z}}) \right\rangle \\ &= \left\langle d\bar{H}(\bar{z}), T_z \rho \circ (T_z \rho|_{\mathcal{H}_z})^{-1}(v_{\bar{z}}) \right\rangle \\ &= \left\langle d\bar{H}(\bar{z}), v_{\bar{z}} \right\rangle. \end{aligned}$$

Therefore, the constrained implicit Hamiltonian system, Eq. (A.3), reduces to

$$(\bar{X}, d\bar{H}) \in \bar{D},$$

or

$$i_{\bar{X}} \bar{\Omega}^{\text{nh}} = d\bar{H}.$$

Remark A.7. Again, this result is essentially a restatement of the nonholonomic reduction of Koiller [32] (see also Bates and Sniatycki [2], Cantrijn et al. [7], and Hochgerner and García-Naranjo [27]) in the language of Dirac structures and implicit Hamiltonian systems.

REFERENCES

- [1] R. Abraham and J. E. Marsden. *Foundations of Mechanics*. Addison–Wesley, 2nd edition, 1978.
- [2] L. Bates and J. Sniatycki. Nonholonomic reduction. *Reports on Mathematical Physics*, 32(1):99–115, 1993.
- [3] A. M. Bloch. *Nonholonomic Mechanics and Control*. Springer, 2003.
- [4] A. M. Bloch and P. E. Crouch. Representations of Dirac structures on vector spaces and nonlinear L-C circuits. In *Differential Geometry and Control Theory*, pages 103–117. American Mathematical Society, 1997.
- [5] H. Bursztyn and O. Radko. Gauge equivalence of Dirac structures and symplectic groupoids. *Annales de l’institut Fourier*, 53(1):309–337, 2003.
- [6] F. Cantrijn, J. F. Cariñena, M. Crampin, and L. A. Ibort. Reduction of degenerate Lagrangian systems. *Journal of Geometry and Physics*, 3(3):353–400, 1986.
- [7] F. Cantrijn, M. de León, J. C. Marrero, and D. Martín de Diego. Reduction of constrained systems with symmetries. *Journal of Mathematical Physics*, 40(2):795–820, 1999.
- [8] J. F. Cariñena, X. Gracia, G. Marmo, E. Martínez, M. Muñoz Lecanda, and N. Román-Roy. Geometric Hamilton–Jacobi theory. *International Journal of Geometric Methods in Modern Physics*, 3(7):1417–1458, 2006.
- [9] J. F. Cariñena, X. Gracia, G. Marmo, E. Martínez, M. C. Muñoz Lecanda, and N. Román-Roy. Geometric Hamilton–Jacobi theory for nonholonomic dynamical systems. *International Journal of Geometric Methods in Modern Physics*, 7(3):431–454, 2010.
- [10] P. J. Channell and C. Scovel. Symplectic integration of Hamiltonian systems. *Nonlinearity*, 3(2):231–259, 1990.
- [11] D. M. F. Chapman. Ideal vortex motion in two dimensions: Symmetries and conservation laws. *Journal of Mathematical Physics*, 19(9):1988–1992, 1978.
- [12] T. Courant. Dirac manifolds. *Transactions of the American Mathematical Society*, 319(2):631–661, 1990.
- [13] M. Dalsmo and A. J. van der Schaft. On representations and integrability of mathematical structures in energy-conserving physical systems. *SIAM Journal on Control and Optimization*, 37(1):54–91, 1998.
- [14] M. de León and D. Martín de Diego. A constraint algorithm for singular Lagrangians subjected to nonholonomic constraints. *Journal of Mathematical Physics*, 38(6):3055–3062, 1997.
- [15] M. de León, J. C. Marrero, and D. Martín de Diego. Linear almost Poisson structures and Hamilton–Jacobi equation. Applications to nonholonomic mechanics. *Journal of Geometric Mechanics*, 2(2):159–198, 2010.
- [16] P. A. M. Dirac. Generalized Hamiltonian dynamics. *Canad. J. Math.*, 2:129–148, 1950.
- [17] P. A. M. Dirac. Generalized Hamiltonian dynamics. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 246(1246):326–332, 1958.
- [18] P. A. M. Dirac. *Lectures on quantum mechanics*. Belfer Graduate School of Science, Yeshiva University, New York, 1964.
- [19] K. M. Ehlers, J. Koiller, R. Montgomery, and P. M. Rios. Nonholonomic systems via moving frames: Cartan equivalence and Chaplygin Hamiltonization. In *The Breadth of Symplectic and Poisson Geometry*, pages 75–120. Birkhäuser, 2004.
- [20] N. Getz. Control of balance for a nonlinear nonholonomic non-minimum phase model of a bicycle. *Proceedings of the American Control Conference*, 1:148–151, 1994.
- [21] N. H. Getz and J. E. Marsden. Control for an autonomous bicycle. *Robotics and Automation, 1995. Proceedings., 1995 IEEE International Conference on*, 2:1397–1402 vol.2, 1995.
- [22] M. J. Gotay and J. M. Nester. Presymplectic Hamilton and Lagrange systems, gauge transformations and the Dirac theory of constraints. In Wolf Beiglböck, Arno Böhm, E. Takasugi, Mark Gotay, and James Nester, editors, *Group Theoretical Methods in Physics*, volume 94, pages 272–279. Springer, 1979.
- [23] M. J. Gotay and J. M. Nester. Presymplectic Lagrangian systems. I: the constraint algorithm and the equivalence theorem. *Annales de l’institut Henri Poincaré (A)*, 30(2):129–142, 1979.

- [24] M. J. Gotay and J. M. Nester. Presymplectic Lagrangian systems. II: the second-order equation problem. *Annales de l'institut Henri Poincaré (A)*, 32(1):1–13, 1980.
- [25] M. J. Gotay, J. M. Nester, and G. Hinds. Presymplectic manifolds and the Dirac–Bergmann theory of constraints. *Journal of Mathematical Physics*, 19(11):2388–2399, 1978.
- [26] M. Henneaux and C. Teitelboim. *Quantization of gauge systems*. Princeton University Press, Princeton, N.J., 1992.
- [27] S. Hochgerner and L. García-Naranjo. G -Chaplygin systems with internal symmetries, truncation, and an (almost) symplectic view of Chaplygin’s ball. *Journal of Geometric Mechanics*, 1(1):35–53, 2009.
- [28] D. Iglesias-Ponte, M. de León, and D. Martín de Diego. Towards a Hamilton–Jacobi theory for non-holonomic mechanical systems. *Journal of Physics A: Mathematical and Theoretical*, 41(1), 2008.
- [29] E. Johnson and T. D. Murphey. Dynamic modeling and motion planning for marionettes: Rigid bodies articulated by massless strings. *Robotics and Automation, 2007 IEEE International Conference on*, pages 330–335, 2007.
- [30] M. Jotz and T. S. Ratiu. Dirac structures, nonholonomic systems and reduction. *Preprint*, 2011.
- [31] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*. Interscience, New York, 1963.
- [32] J. Koiller. Reduction of some classical non-holonomic systems with symmetry. *Archive for Rational Mechanics and Analysis*, 118(2):113–148, 1992.
- [33] W. S. Koon and J. E. Marsden. The Hamiltonian and Lagrangian approaches to the dynamics of nonholonomic systems. *Reports on Mathematical Physics*, 40(1):21–62, 1997.
- [34] P. S. Krishnaprasad and D. P. Tsakiris. Oscillations, SE(2)-snakes and motion control: a study of the roller racer. *Dynamical Systems*, 16(4):347–397, 2001.
- [35] H. P. Künzle. Degenerate Lagrangean systems. *Annales de l'institut Henri Poincaré (A)*, 11(4):393–414, 1969.
- [36] M. Leok and T. Ohsawa. Discrete Dirac structures and implicit discrete Lagrangian and Hamiltonian systems. In M. Asorey, J. Clemente-Gallardo, E. Martinez, and J. F. Carinena, editors, *XVIII International Fall Workshop on Geometry and Physics*, volume 1260, pages 91–102. AIP, 2010.
- [37] M. Leok and T. Ohsawa. Variational and geometric structures of discrete Dirac mechanics. *Foundations of Computational Mathematics*, published online, 2011.
- [38] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numerica*, pages 357–514, 2001.
- [39] R. Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Applications*. American Mathematical Society, 2002.
- [40] T. D. Murphey and M. Egerstedt. Choreography for marionettes: Imitation, planning, and control. In *IEEE Int. Conf. on Intelligent and Robotic Systems Workshop on Art and Robotics*, 2007.
- [41] P. K. Newton. *The N -vortex problem*. Springer, New York, 2001.
- [42] T. Ohsawa and A. M. Bloch. Nonholonomic Hamilton–Jacobi equation and integrability. *Journal of Geometric Mechanics*, 1(4):461–481, 2009.
- [43] T. Ohsawa, O. E. Fernandez, A. M. Bloch, and D. V. Zenkov. Nonholonomic Hamilton–Jacobi theory via Chaplygin Hamiltonization. *Journal of Geometry and Physics*, 61(8):1263–1291, 2011.
- [44] K. D. Rothe and F. G. Scholtz. On the Hamilton–Jacobi equation for second-class constrained systems. *Annals of Physics*, 308(2):639–651, 2003.
- [45] C. W. Rowley and J. E. Marsden. Variational integrators for degenerate Lagrangians, with application to point vortices. In *Proceedings of the 41st IEEE CDC*, 2002.
- [46] D. P. Tsakiris. *Motion Control and Planning for Nonholonomic Kinematic Chains*. PhD thesis, University of Maryland, College Park, 1995.
- [47] W. M. Tulczyjew. Les sous-variétés lagrangiennes et la dynamique hamiltonienne. *C. R. Acad. Sc. Paris*, 283:15–18, 1976.
- [48] W. M. Tulczyjew. Les sous-variétés lagrangiennes et la dynamique lagrangienne. *C. R. Acad. Sc. Paris*, 283:675–678, 1976.
- [49] A. J. van der Schaft. Implicit Hamiltonian systems with symmetry. *Reports on Mathematical Physics*, 41(2):203–221, 1998.
- [50] A. J. van der Schaft. Port-Hamiltonian systems: an introductory survey. In *Proceedings of the International Congress of Mathematicians*, volume 3, pages 1339–1365, 2006.

- [51] A. M. Vershik and V. Ya. Gershkovich. Nonholonomic problems and the theory of distributions. *Acta Applicandae Mathematicae: An International Survey Journal on Applying Mathematics and Mathematical Applications*, 12(2):181–209, 1988.
- [52] H. Yoshimura and J. E. Marsden. Dirac structures in Lagrangian mechanics Part I: Implicit Lagrangian systems. *Journal of Geometry and Physics*, 57(1):133–156, 2006.
- [53] H. Yoshimura and J. E. Marsden. Dirac structures in Lagrangian mechanics Part II: Variational structures. *Journal of Geometry and Physics*, 57(1):209–250, 2006.
- [54] H. Yoshimura and J. E. Marsden. Dirac structures and the Legendre transformation for implicit Lagrangian and Hamiltonian systems. In *Lagrangian and Hamiltonian Methods for Nonlinear Control 2006*, pages 233–247, 2007.
- [55] H. Yoshimura and J. E. Marsden. Dirac cotangent bundle reduction. *Journal of Geometric Mechanics*, 1(1):87–158, 2009.

M. LEOK & T. OHSAWA: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SAN DIEGO, 9500 GILMAN DRIVE, LA JOLLA, CALIFORNIA, USA

E-mail address: mleok@math.ucsd.edu, tohsawa@ucsd.edu

D. SOSA: DEPARTAMENTO DE ECONOMÍA APLICADA Y UNIDAD ASOCIADA ULL-CSIC GEOMETRÍA DIFERENCIAL Y MECÁNICA GEOMÉTRICA, FACULTAD DE CC. EE. Y EMPRESARIALES, UNIVERSIDAD DE LA LAGUNA, LA LAGUNA, TENERIFE, CANARY ISLANDS, SPAIN

E-mail address: dnsosa@ull.es