Vol. 69 (2012) *REPORTS ON MATHEMATICAL PHYSICS* No. 1

DIRAC STRUCTURES, NONHOLONOMIC SYSTEMS AND REDUCTION

MADELEINE JOTZ

Mathematisches Institut Georg August, Universitt Göttingen,
Bunsenstrasse 3-5 37073 Göttingen, Germany

and

T. S. RATIU

Section de Mathématiques et Bernoulli Center, Ecole Polytechnique Fédérale de Lausanne,
1015 Lausanne, Switzerland
(e-mails: mjotz@uni-math.gwdg.de, tudor.ratiu@epfl.ch)

(*Received May 7, 2011*)

The reduction of nonholonomic systems is formulated in terms of Dirac reduction. An optimal reduction method for a class of nonholonomic systems is formulated. Several examples are studied in detail.

AMS Classification: 53D20, 37J15, 70F25, 70H45, 37J60, 53D17

Keywords: Dirac structure, symmetry reduction, nonholonomic mechanics, momentum map.

1. Introduction

The equations of motion of nonholonomic mechanical systems and those in circuit theory can be geometrically described using a Dirac structure (introduced by [12, 13]) in taking either a Hamiltonian or Lagrangian point of view (see e.g. [3–5, 35–37]). A Dirac structure D on a manifold M is a subbundle of the Pontryagin bundle $TM \oplus T^*M$ which is Lagrangian relative to the canonical symmetric pairing on it. Dirac structures were introduced by [12, 13] to provide a geometric framework for the study of constrained mechanical systems. An easy example of a Dirac structure is the graph of a 2-form $\omega \in \Omega^2(M)$. Integrable Dirac structures have an additional integrability condition. They have been more intensively studied because they generalize, in a certain sense, Poisson structures. For example, if the Dirac structure is the graph of $\omega \in \Omega^2(M)$, then it is integrable if and only if $\mathbf{d}\omega = 0$. Other examples of integrable Dirac structures include various foliated

M.J. was partially supported by Swiss NSF grant 200021-121512 and by a Dorothea Schlözer Fellowship of the University of Göttingen. T.S.R. was partially supported by Swiss NSF grant 200021-121512 and by the government grant of the Russian Federation for support of research projects implemented by leading scientists, Lomonosov Moscow State University, under the agreement No. 11.G34.31.0054.

manifolds. In general, an integrable Dirac structure determines a singular foliation on M whose leaves carry a natural induced presymplectic structure.

Dirac structures simultaneously generalize symplectic and Poisson structures and also form the correct setting for the description of implicit Hamiltonian and Lagrangian systems usually appearing as systems of algebraic-differential equations. In symplectic and Poisson geometry, as well as geometric mechanics, a major role is played by the reduction method since it creates, under suitable hypotheses or in categories weaker than smooth manifolds, new spaces with the same type of motion equations on them. Briefly put, it is a method that eliminates variables and hence yields systems on smaller dimensional manifolds. Due to the spectacular array of applications, reduction has been extensively studied in various settings, including that of Dirac manifolds. The present paper connects Dirac and nonholonomic reduction, introduces an "optimal" reduction for special nonholonomic systems with symmetries and presents several classical examples.

First we recall the necessary background on Dirac geometry in Section 2, as well as descriptions of Dirac reduction by symmetry groups. It is known that under certain assumptions beyond the usual ones, the quotient manifold carries a natural Dirac structure. These hypotheses are formulated in the literature in two different manners: using sections (see [5]) or appealing to the theory of fiber bundles (see [7]). We show in Section 3 that Dirac reduction as presented in Subsection 2.3, coincides with the method of reduction for nonholonomic systems due to [2]. This is achieved by reformulating their Hamiltonian approach to nonholonomic systems in the context of Dirac structures. We give several standard examples that illustrate this.

We reformulate in Section 4.1 the nonholonomic Noether theorem (see [2], §6, [14], Theorem 2, and [6] [1]) on the Hamiltonian side and show that, under the *dimension assumption*, it is equivalent to the nonholonomic Noether theorem in [6]. We study in Section 4.2 a distribution where the fundamental vector fields have to lie to yield constants of motion. This gives an explanation for certain constants of motion that sometimes appear as a consequence of the nonholonomic Noether theorem (see [15]).

Under certain integrability assumptions imposed on a distribution associated to a Dirac structure modeling these, it is possible to extend the ideas in Marsden–Weinstein reduction to nonholonomic systems. This is achieved in Section 4. These integrability conditions are certainly strong since they imply that the nonholonomic Noether 1-forms that descend to the quotient are exact. This is not true in general but holds in the case of certain systems such as the vertical rolling disk or the constrained particle. We discuss some of these examples at the end of the paper.

CONVENTIONS. Throughout the paper M is a *paracompact* manifold, that is, it is Hausdorff and every open covering admits a locally finite refinement. The orientation preserving rotation group SO(2) of the plane $\mathbb{R}^2$ is also denoted by $\mathbb{S}^1$

[1] A somewhat restricted version of the momentum equation was given in [20]; see also [1].

and consists of matrices of the form

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}, \qquad \alpha \in \mathbb{R}.$$

If $E \to M$ is a smooth fiber bundle over a manifold M, the spaces of smooth global and local sections are denoted by $\Gamma_{\text{global}}(E)$ and $\Gamma(E)$, respectively. For example, $\mathfrak{X}(M) := \Gamma(TM)$ denotes the Lie algebra of smooth local vector fields endowed with the usual Jacobi–Lie bracket $[X, Y](f) = X[Y[f]] - Y[X[f]]$, where $X, Y \in \mathfrak{X}(M)$, f is a smooth (possibly only locally defined) function on M, and $X[f] := \pounds_X f = \mathbf{d}f(X)$ denotes the Lie derivative of f in the direction X. If $\wedge^k(M) \to M$ denotes the vector bundle of exterior k-forms on M then $\Omega^k(M) := \Gamma(\wedge^k(M))$ is the space of local k-forms on the manifold M.

Recall that a subset $N \subset M$ is an *initial* submanifold of M if N carries a manifold structure such that the inclusion $\iota : N \hookrightarrow M$ is a smooth immersion and satisfies the following condition: for any smooth manifold P an arbitrary map $g : P \to N$ is smooth if and only if $\iota \circ g : P \to M$ is smooth. The notion of initial submanifold lies strictly between those of injectively immersed and embedded submanifolds.

2. Background and Dirac structures

This section summarizes the key facts from the theory of foliations and Dirac manifolds needed in the rest of the paper. It also establishes notation, terminology, and conventions, since these are not uniform in the literature. The proofs of the statements below can be found in [4, 5, 7, 12].

2.1. Distributions and foliations

We will need a few standard facts from the theory of generalized distributions on a smooth manifold M (see [29–31], [33] for the original articles and [21, 24, 25, 34], for a quick review of this theory).

A *generalized distribution* Δ on M is a subset of the tangent bundle TM such that $\Delta(m) := \Delta \cap T_mM$ is a vector subspace of T_mM. The number $\dim \Delta(m)$ is called the *rank* of Δ at $m \in M$. A local *differentiable section* of Δ is a smooth vector field $X \in \mathfrak{X}(M)$ defined on some open subset $U \subset M$ such that $X(u) \in \Delta(u)$ for each $u \in U$. A generalized distribution is said to be *differentiable* or *smooth* if for every point $m \in M$ and every vector $v \in \Delta(m)$, there is a differentiable section $X \in \Gamma(\Delta)$ defined on an open neighbourhood U of m such that $X(m) = v$.

The term *distribution* is usually synonymous to that of a vector subbundle of TM. Since we shall work mostly with generalized distributions, we shall call below all generalized distributions simply distributions. If the generalized distribution happens to be a vector subbundle we shall always state this fact explicitly.

In all that follows, Δ is a smooth distribution. An *integral manifold* of Δ is an injectively immersed connected manifold $\iota_L : L \hookrightarrow M$ satisfying the condition

$T_m\iota_L(T_mL) \subset \Delta(m)$ for every $m \in L$. The integral manifold L is of *maximal dimension* at $m \in L$ if $T_m\iota_L(T_mL) = \Delta(m)$. The distribution Δ is *completely integrable* if for every $m \in M$ there is an integral manifold L of Δ, $m \in L$, everywhere of maximal dimension. The distribution Δ is *involutive* if it is invariant under the (local) flows associated to differentiable sections of Δ. The distribution Δ is *algebraically involutive* if for any two smooth vector fields defined on an open set of M which take values in Δ, their bracket also takes values in Δ. Clearly involutive distributions are algebraically involutive and the converse is true if the distribution is a subbundle. The analog of the Frobenius theorem (which deals only with vector subbundles of TM) for distributions is known as the Stefan–Sussmann theorem. Its statement is the same except that one needs the distribution to be involutive and not just algebraically involutive: *Δ is completely integrable if and only if Δ is involutive*.

Recall that the Frobenius theorem states that a vector subbundle of TM is (algebraically) involutive if and only if it is the tangent bundle of a foliation on M. The same is true for distributions: *A smooth distribution is involutive if and only if it coincides with the set of vectors tangent to a generalized foliation.* To give content to this statement and elaborate on it, we need to quickly review the concept and main properties of generalized foliations.

A *generalized foliation* on M is a partition $\mathfrak{F} := \{\mathcal{L}_\alpha\}_{\alpha\in A}$ of M into disjoint connected sets, called *leaves*, such that each point $m \in M$ has a *generalized foliated chart* $(U, \varphi : U \to V \in \mathbb{R}^{\dim M})$, $m \in U$. This means that there is some natural number $p_\alpha \leq \dim M$, called the *dimension* of the leaf $\mathcal{L}_\alpha$, and a subset $S_\alpha \subset \mathbb{R}^{\dim M - p_\alpha}$ such that $\varphi(U \cap \mathcal{L}_\alpha) = \{(x^1, \ldots, x^{\dim M}) \in V \mid (x^{p_\alpha+1}, \ldots, x^{\dim M}) \in S_\alpha\}$ and each $(x_\circ^{p_\alpha+1}, \ldots, x_\circ^{\dim M}) \in S_\alpha$ determines a connected component $(U \cap \mathcal{L}_\alpha)_\circ$ of $U \cap \mathcal{L}_\alpha$, that is, $\varphi((U \cap \mathcal{L}_\alpha)_\circ) = \{(x^1, \ldots, x^{p_\alpha}, x_\circ^{p_\alpha+1}, \ldots, x_\circ^{\dim M}) \in V\}$. The key difference with the concept of foliation is that the number p_α can change from leaf to leaf. The generalized foliated charts induce on each leaf a smooth manifold structure that makes them into initial submanifolds of M.

A leaf $\mathcal{L}_\alpha$ is called *regular* if it has an open neighbourhood that intersects only leaves whose dimension equals $\dim \mathcal{L}_\alpha$. If such a neighbourhood does not exist, then $\mathcal{L}_\alpha$ is called a *singular* leaf. A point is called *regular* (*singular*) if it is contained in a regular (singular) leaf. The set of vectors tangent to the leaves of $\mathfrak{F}$ is defined by

$$T(M, \mathfrak{F}) := \bigcup_{\alpha\in A} \bigcup_{m\in\mathcal{L}_\alpha} T_m\mathcal{L}_\alpha \subset TM.$$

Under mild topological conditions on M a generalized foliation has very useful properties. Assume that M is second countable. Then for each p_α-dimensional leaf $\mathcal{L}_\alpha$ and any generalized foliated chart $(U, \varphi : U \to V \in \mathbb{R}^{\dim M})$ that intersects it, the corresponding set S_α is countable. The set of regular points is open and dense in M. Finally, any closed leaf is embedded in M. Note that this last property is specific to (generalized) foliations since an injectively immersed submanifold whose range is closed is not necessarily embedded.

Let us return now to the relationship between distributions and generalized foliations. As already mentioned, given an involutive (and hence a completely integrable) distribution Δ, each point $m \in M$ belongs to exactly one connected integral manifold $\mathcal{L}_m$ that is maximal relative to inclusion. It turns out that $\mathcal{L}_m$ is an initial submanifold and that it is also the *accessible* set of m, that is, $\mathcal{L}_m$ equals the subset of points in M that can be reached by applying to m a finite number of composition of flows of elements of $\Gamma(\Delta)$. The collection of all maximal integral submanifolds of Δ forms a generalized foliation $\mathfrak{F}_\Delta$ such that $\Delta = T(M, \mathfrak{F}_\Delta)$. Conversely, given a generalized foliation $\mathfrak{F}$ on M, the subset $T(M, \mathfrak{F}) \subset TM$ is a smooth completely integrable (and hence involutive) distribution whose collection of maximal integral submanifolds coincides with $\mathfrak{F}$. These two statements expand the Stefan–Sussmann theorem cited above.

In the study of Dirac manifolds we will also need the concept of codistribution. A *generalized codistribution* Ξ on M is a subset of the cotangent bundle T^*M such that $\Xi(m) := \Xi \cap T_m^*M$ is a vector subspace of T_m^*M. The notions of rank, differentiable section, and smooth codistribution are completely analogous to those for distributions.

If $\Delta \subset TM$ is a smooth distribution on M, its (*smooth*) *annihilator* Δ° is defined by

$$\Delta^\circ(m) := \left\{ \alpha(m) \;\middle|\; \begin{array}{l} \alpha \in \Omega^1(M), \qquad \langle \alpha, X \rangle = 0 \text{ for all } X \in \mathfrak{X}(U), \\ m \in U \text{ open such that } X(u) \in \Delta(u) \text{ for all } u \in U \end{array} \right\}.$$

We have, in general strict, inclusion $\Delta \subset \Delta^{\circ\circ}$. A similar definition holds for smooth codistributions. Note that the annihilators are smooth by construction. If a distribution (codistribution) is a vector subbundle of TM (respectively of T^*M), then its annihilator is also a vector subbundle of T^*M (respectively of TM). If Δ is a subbundle then $\Delta = \Delta^{\circ\circ}$ and similarly for codistributions.

2.2. Dirac structures

For a smooth manifold M denote by $\langle \cdot, \cdot \rangle$ the duality pairing between the cotangent bundle T^*M and the tangent bundle TM (or $\Omega^1(M)$ and $\mathfrak{X}(M)$). The *Pontryagin bundle* $TM \oplus T^*M$ is endowed with a nondegenerate symmetric fiberwise bilinear form of signature $(\dim M, \dim M)$ given by

$$\langle\langle (u_m, \alpha_m), (v_m, \beta_m) \rangle\rangle := \langle \beta_m, u_m \rangle + \langle \alpha_m, v_m \rangle \tag{1}$$

for all $u_m, v_m \in T_mM$ and $\alpha_m, \beta_m \in T_m^*M$. A *Dirac structure* (see [12]) on M is a Lagrangian subbundle $D \subset TM \oplus T^*M$, that is, D coincides with its orthogonal relative to (1) and so its fibers are necessarily $\dim M$-dimensional.

The space $\Gamma(TM \oplus T^*M)$ of local sections of the Pontryagin bundle is endowed with an $\mathbb{R}$-bilinear skew-symmetric bracket (which does not satisfy the Jacobi identity) given by

$$[(X,\alpha),(Y,\beta)] := \left([X,Y], \boldsymbol{\pounds}_X\beta - \boldsymbol{\pounds}_Y\alpha + \frac{1}{2}\mathbf{d}\,(\alpha(Y) - \beta(X)) \right)$$
$$= \left([X,Y], \boldsymbol{\pounds}_X\beta - \mathbf{i}_Y\mathbf{d}\alpha - \frac{1}{2}\mathbf{d}\,\langle (X,\alpha),(Y,\beta)\rangle \right) \tag{2}$$

(see [12]). The Dirac structure is *integrable* if $[\Gamma(D),\Gamma(D)] \subset \Gamma(D)$. Since $\langle (X,\alpha),(Y,\beta)\rangle = 0$ if $(X,\alpha),(Y,\beta) \in \Gamma(D)$, integrability of the Dirac structure is often expressed in the literature relative to a non-skew-symmetric bracket that differs from (2) by eliminating in the second line the third term of the second component. This truncated expression which satisfies the Jacobi identity but is no longer skew-symmetric is called the *Courant–Dorfman bracket* (see [7–9, 22, 28]).

A Dirac structure defines two smooth distributions $\mathsf{G}_0, \mathsf{G}_1 \subset TM$ and two smooth codistributions $\mathsf{P}_0, \mathsf{P}_1 \subset T^*M$:

$$\mathsf{G}_0(m) := \{X(m) \in T_mM \mid X \in \mathfrak{X}(M), (X,0) \in \Gamma(D)\},$$
$$\mathsf{G}_1(m) := \{X(m) \in T_mM \mid X \in \mathfrak{X}(M), \text{ there is an } \alpha \in \Omega^1(M), \text{such that } (X,\alpha) \in \Gamma(D)\}$$

and

$$\mathsf{P}_0(m) := \{\alpha(m) \in T_m^*M \mid \alpha \in \Omega^1(M), (0,\alpha) \in \Gamma(D)\},$$
$$\mathsf{P}_1(m) := \{\alpha(m) \in T_m^*M \mid \alpha \in \Omega^1(M), \text{ there is an } X \in \mathfrak{X}(M), \text{such that } (X,\alpha) \in \Gamma(D)\}.$$

The smoothness of $\mathsf{G}_0, \mathsf{G}_1, \mathsf{P}_0, \mathsf{P}_1$ is obvious since, by definition, they are generated by smooth local sections. In general, these are not vector subbundles of TM and T^*M, respectively. It is also clear that $\mathsf{G}_0 \subset \mathsf{G}_1$ and $\mathsf{P}_0 \subset \mathsf{P}_1$.

The *characteristic equations* of a Dirac structure are:

(i) $\mathsf{G}_0 = \mathsf{P}_1^\circ$, $\mathsf{P}_0 = \mathsf{G}_1^\circ$,

(ii) $\mathsf{P}_1 \subset \mathsf{G}_0^\circ$, $\mathsf{G}_1 \subset \mathsf{P}_0^\circ$,

(iii) If P_1 has constant rank, then $\mathsf{P}_1 = \mathsf{G}_0^\circ$. If G_1 has constant rank, then $\mathsf{G}_1 = \mathsf{P}_0^\circ$.

If D is a Dirac structure on M having the property that $\mathsf{G}_1 \subset TM$ is a constant rank distribution on M, then there exists a skew-symmetric vector bundle map $\flat : \mathsf{G}_1 \to \mathsf{G}_1^*$ such that D is given by

$$D(m) := \{(X(m),\alpha(m)) \in T_mM \oplus T_m^*M \mid X \text{ a smooth local section of } \mathsf{P}^\circ, \alpha \in \Omega^1(M),\ \alpha|_{\mathsf{P}^\circ} = X^\flat\} \tag{3}$$

with $\mathsf{P} := \mathsf{P}_0 = \mathsf{G}_1^\circ$. Also, $\ker(\flat : \mathsf{G}_1 \to \mathsf{G}_1^*) = \mathsf{G}_0$.

A function $f \in C^\infty(M)$ is called *admissible* if $\mathbf{d}f \in \Gamma(\mathsf{P}_1)$. There is an induced bracket $\{\cdot,\cdot\}_D$ on the admissible functions given by

$$\{f,g\}_D = X_g[f] = -X_f[g], \tag{4}$$

where $X_f \in \mathfrak{X}(M)$ is such that $(X_f, \mathbf{d}f) \in \Gamma(D)$. If the Dirac structure D on M is integrable, this bracket is a Poisson bracket. Note that $X_f \in \mathfrak{X}(M)$ is not uniquely determined by this condition. If the Dirac structure is not integrable, we get with the same definition an almost Poisson structure, that is, the Jacobi-identity does not necessarily hold.

Integrable Dirac structures as Lie algebroids. A *Lie algebroid* $E \to M$ is a smooth vector bundle over M with a vector bundle homomorphism $\rho : E \to TM$, called the *anchor*, and a Lie algebra bracket $[\cdot,\cdot] : \Gamma(E) \times \Gamma(E) \to \Gamma(E)$ satisfying:

1. ρ is a Lie algebra homomorphism,
2. for all $f \in C^\infty(M)$ and $X, Y \in \Gamma(E)$:

$$[X, fY] = f[X, Y] + \rho(X)[f]Y.$$

It is shown in [12] that for an arbitrary Lie algebroid $E \to M$, the smooth distribution $\rho(E)$ is completely integrable.

Assume that D is an integrable Dirac structure. Then, relative to the Courant bracket (2) and the anchor $\pi_1 : D \to TM$ given by the projection on the first factor, D becomes a Lie algebroid over M. The smooth distribution $\pi_1(D) \subset TM$ coincides with G_1. Hence, G_1 is completely integrable and Theorem 2.3.6 in [12] states the following result.

THEOREM 1. *An integrable Dirac structure has a generalized foliation by presymplectic leaves.*

The presymplectic form ω_N on a leaf N of the generalized foliation by leaves of G_1 is given by

$$\omega_N(\tilde{X}, \tilde{Y})(p) = \alpha(Y)(p) = -\beta(X)(p) \tag{5}$$

for all $p \in N$ and $\tilde{X}, \tilde{Y} \in \mathfrak{X}(N)$, where $i_N : N \hookrightarrow M$ is the inclusion and $X, Y \in \Gamma(\mathsf{G}_1)$ are i_N-related to $\tilde{X}, \tilde{Y}$, respectively; we shall denote i_N-relatedness by $\tilde{X} \sim_{i_N} X$ and $\tilde{Y} \sim_{i_N} Y$. The 1-forms $\alpha, \beta \in \Omega^1(M)$ are such that $(X, \alpha), (Y, \beta) \in \Gamma(D)$. Formula (5) is independent of all the choices involved.

Implicit Hamiltonian systems. Let D be a Dirac structure on M and $H \in C^\infty(M)$. The *implicit Hamiltonian system* (M, D, H) is defined as the set of C^∞ solutions $x(t)$ satisfying the condition

$$(\dot{x}, \mathbf{d}H(x(t))) \in D(x(t)), \qquad \text{for all } t. \tag{6}$$

In this general situation, *conservation of energy* is still valid: $\dot{H}(t) = \langle \mathbf{d}H(x(t)), \dot{x}(t)\rangle = 0$, for all t for which the solution exists. In addition, these equations contain algebraic constraints, namely, $\mathbf{d}H(x(t)) \in \mathsf{P}_1(x(t))$, for all t. Note that $\dot{x}(t) \in \mathsf{G}_1(x(t))$, so the set of *admissible flows* have velocities in the distribution G_1. Thus, an implicit Hamiltonian system defines a set of *differential and algebraic equations*.

Note that if G_1 is an involutive subbundle of TM, then there are $\dim M - \operatorname{rank} \mathsf{G}_1$ independent conserved quantities for the Hamiltonian system (6). We want to emphasize that standard existence and uniqueness theorems do not apply to (6), even if all the distributions and codistributions are subbundles. The only general theorems that ensure the local existence and uniqueness of solutions for (6) are for the so-called implicit Hamiltonian systems of index one (see [3, 5]).

Restriction of Dirac structures. Let D be a Dirac structure on M and $N \subset M$ a submanifold of M. Define the map $\sigma(m) : T_mN \times T^*_mM \to T_mN \times T^*_mN$, $m \in N$, by $\sigma(m)(v_m, \alpha_m) = (v_m, \alpha_m|_{T_mN})$. Assume that the dimension of $\mathsf{G}_1(m) \cap T_mN$ is independent of $m \in N$ and that the rank of G_1 is constant on M. Define the vector subbundle $D_N \subset TN \oplus T^*N$ by

$$D_N(m) = \sigma(m)\left(D(m) \cap (T_mN \times T^*_mM)\right), \qquad m \in N.$$

Then D_N is a Lagrangian subbundle in the Pontryagin bundle $TN \oplus T^*N$ and it is thus a Dirac structure on N. Let $\iota : N \hookrightarrow M$ denote the inclusion map and define for all $m \in N$

$$E_s(m) := \left\{ (X(m), \alpha(m)) \in T_mM \times T^*_mM \;\middle|\; \begin{array}{c} \alpha \in \Omega^1(M), \qquad X \in \mathfrak{X}(M) \text{ such that} \\ X(n) \in T_nN \text{ for all } n \in N \text{ for which } X \text{ is defined} \end{array} \right\}$$

(where the subscript s stands for submanifold). This defines a smooth bundle $E_s = \cup_{m\in N} E_s(m)$ on N. G. Blankenstein and A. J. van der Schaft in [5] show that under the assumption that the fibers of $E_s \cap D$ have constant dimension on M, there is another way to give the induced Dirac structure, namely, $(\widetilde{X}, \widetilde{\alpha})$ is a local section of D_N if and only if there exists a local section (X, α) of D such that $\widetilde{X} \sim_\iota X$ and $\widetilde{\alpha} = \iota^*\alpha$. Otherwise stated,

$$\Gamma(D_N) = \left\{ (\widetilde{X}, \widetilde{\alpha}) \in \mathfrak{X}(N) \oplus \Omega^1(N) \;\middle|\; \begin{array}{c} \text{there is } (X, \alpha) \in \Gamma(D) \text{ such that} \\ \widetilde{X} \sim_\iota X \text{ and } \widetilde{\alpha} = \iota^*\alpha \end{array} \right\}. \tag{7}$$

Furthermore, if D is integrable, then D_N is also integrable. As stated in [5], if G_1 is constant dimensional, the assumptions for both methods of restriction are equivalent.

Second, we recall the restriction construction for implicit Hamiltonian systems. Given is the implicit Hamiltonian system (M, D, H) and $N \subset M$ an invariant submanifold under the integral curves of (M, D, H) (if they exist). Define $H_N := H|_N = H \circ \iota$. Then every solution $x(t)$ of (M, D, H) which leaves N invariant (that is, $x(t) \in N$ for all t) is a solution of (N, D_N, H_N). The converse statement is not true, in general.

Symmetries of Dirac manifolds. Let G be a Lie group and $\Phi : G \times M \to M$ a smooth left action. Then G is called a *symmetry Lie group of* D if for every $g \in G$ the condition $(X, \alpha) \in \Gamma(D)$ implies that $\left(\Phi^*_g X, \Phi^*_g \alpha\right) \in \Gamma(D)$. We say then that the Lie group G acts *canonically* or *by Dirac actions* on M.

For any admissible $f \in C^\infty(M)$, i.e. a function such that $(X_f, \mathbf{d}f) \in \Gamma(D)$ for some $X_f \in \mathfrak{X}(M)$, this yields $(\Phi_g^* X_f, \Phi_g^* \mathbf{d}f) \in \Gamma(D)$ or $(\Phi_g^* X_f, \mathbf{d}(\Phi_g^* f)) \in \Gamma(D)$. Hence we have simultaneously the facts that $\Phi_g^* f$ is admissible and that $\Phi_g^* X_f - X_{\Phi_g^* f} =: Y \in \Gamma(\mathsf{G}_0)$. This implies for the almost Poisson bracket on admissible functions (see 2.2),

$$\begin{aligned}\Phi_g^*\{f, h\}_D &= -\Phi_g^*(X_f[h]) = -(\Phi_g^* X_f)[\Phi_g^* h] = -(Y + X_{\Phi_g^* f})[\Phi_g^* h] \\ &= -\mathbf{d}(\Phi_g^* h)(Y + X_{\Phi_g^* f}) = -\mathbf{d}(\Phi_g^* h)(X_{\Phi_g^* f}) = \{\Phi_g^* f, \Phi_g^* h\}_D\end{aligned}$$

since $\Phi_g^* h$ is an admissible function (and hence $\mathbf{d}(\Phi_g^* h) \in \Gamma(\mathsf{P}_1) \subset \Gamma\left(\mathsf{G}_0^\circ\right)$ and $Y \in \Gamma(\mathsf{G}_0)$).

The Lie group G is a *symmetry Lie group of the implicit Hamiltonian system* (M, D, H) if, in addition, H is G-invariant, that is, $H \circ \Phi_g = H$ for all $g \in G$.

Let $\mathfrak{g}$ be a Lie algebra and $\xi \in \mathfrak{g} \mapsto \xi_M \in \mathfrak{X}(M)$ be a smooth left Lie algebra action, that is, the map $(x, \xi) \in M \times \mathfrak{g} \mapsto \xi_M(x) \in TM$ is smooth and $\xi \in \mathfrak{g} \mapsto \xi_M \in \mathfrak{X}(M)$ is a Lie algebra anti-homomorphism. The Lie algebra $\mathfrak{g}$ is said to be a *symmetry Lie algebra of* D if for every $\xi \in \mathfrak{g}$ the condition $(X, \alpha) \in \Gamma(D)$ implies that $\left(\pounds_{\xi_M} X, \pounds_{\xi_M} \alpha\right) \in \Gamma(D)$. If, in addition, $\pounds_{\xi_M} H = 0$ for all $\xi \in \mathfrak{g}$, then $\mathfrak{g}$ is a *symmetry Lie algebra of the implicit Hamiltonian system* (M, D, H). Of course, if $\mathfrak{g}$ is the Lie algebra of G and $\xi \mapsto \xi_M$ is the associated infinitesimal generator, then G is a symmetry Lie group of D if and only if $\mathfrak{g}$ is a symmetry Lie algebra of D.

2.3. Regular reduction of Dirac structures

In all that follows we shall assume that (M, D) is a smooth Dirac manifold, G is a symmetry Lie group of the Dirac structure D on M and that the action is free and proper. Thus, the projection on the quotient $\pi : M \to M/G := \bar{M}$ defines a left principal G-bundle. Note that the Dirac structure $D \subset TM \oplus T^*M$ is G-invariant as a subbundle since for all $g \in G$ and $(X, \alpha) \in \Gamma(D)$ we have $(\Phi_g^* X, \Phi_g^* \alpha) \in \Gamma(D)$. Set $\mathfrak{g}_M := \{\xi_M \mid \xi \in \mathfrak{g}\} \subset \mathfrak{X}(M)$ and, for $m \in M$, define the vector subspace $\mathcal{V}(m) := \{\xi_M(m) \mid \xi \in \mathfrak{g}\} \subseteq T_m M$ and the distribution $\mathcal{V} := \cup_{m \in M} \mathcal{V}(m)$. Since the G-action is free, $\mathcal{V}$ is a G-invariant vector subbundle of TM and its annihilator $\mathcal{V}^\circ$ is a G-invariant subbundle of T^*M. It is worth noting that the space of sections $\Gamma(\mathcal{V})$ coincides with the $C^\infty(M)$-module spanned by $\mathfrak{g}_M$.

For all $m \in M$ the map $T_m\pi : T_m M \to T_{\pi(m)}\bar{M}$ is surjective with kernel $\mathcal{V}(m)$. This yields an isomorphism between $T_m M/\mathcal{V}(m)$ and $T_{\pi(m)}\bar{M}$. The Lie group G acts smoothly on the quotient vector bundle $TM/\mathcal{V}$ by $g \cdot \hat{v} := \widehat{T\Phi_g(v)}$, where $\hat{v} \in TM/\mathcal{V}$.

For $X \in \mathfrak{X}(M)$, we will say that the section $\widehat{X} := X(\mathrm{mod}\,\mathcal{V})$ of $TM/\mathcal{V}$ is *G-equivariant*, if there is a representative X^G of $\widehat{X}$ that is G-equivariant, i.e. a smooth section $X^G \in \mathfrak{X}(M)^G$ with $X - X^G \in \Gamma(\mathcal{V})$. This is equivalent to the condition $[X, V] \in \Gamma(\mathcal{V})$ for all representatives X of $\widehat{X}$ and for all $V \in \Gamma(\mathcal{V})$ (see for instance [19]). In what follows we shall use these two equivalent definitions interchangeably.

The representative X^G of $\widehat{X}$ uniquely induces a smooth vector field $\bar{X}$ on $\bar{M}$, where $\bar{X}$ is defined by the condition $X^G \sim_\pi \bar{X}$, that is, $T\pi \circ X^G = \bar{X} \circ \pi$. The map

$$\begin{aligned} \Pi : \Gamma(TM/\mathcal{V})^G &\to \mathfrak{X}(\bar{M}), \\ X(\operatorname{mod} \mathcal{V}) &\mapsto \bar{X}, \end{aligned} \tag{8}$$

is a well-defined homomorphism of $C^\infty(\bar{M})$-modules (note that $C^\infty(\bar{M}) \simeq C^\infty(M)^G$ via $\bar{f} \mapsto \pi^* \bar{f}$). This map (8) is in fact an isomorphism (use, for instance, the results in [27]).

In the same way, for all $\bar{\alpha} \in \Omega^1(\bar{M})$, we have $\pi^*\bar{\alpha} \in \Gamma(\mathcal{V}^\circ)^G$. Note that if $\alpha \in \Gamma(\mathcal{V}^\circ)^G$, then the 1-form $\bar{\alpha} \in \Omega^1(\bar{M})$ defined by $\langle \bar{\alpha}(\pi(m)), T_m\pi(v_m)\rangle := \langle \alpha(m), v_m\rangle$, for all $v_m \in T_mM$, is well defined and satisfies $\pi^*\bar{\alpha} = \alpha$. This shows that the map $\bar{\alpha} \in \Omega^1(\bar{M}) \mapsto \pi^*\bar{\alpha} \in \Gamma(\mathcal{V}^\circ)^G$ is an isomorphism of $C^\infty(\bar{M})$-modules.

We close these preliminary remarks by recording that the G-action on $(TM/\mathcal{V}) \oplus \mathcal{V}^\circ$,

$$g \cdot \left(\hat{v}_m, \alpha_m\right) := \left(\widehat{T_m\Phi_g(v_m)}, T^*_{g\cdot m}\Phi_{g^{-1}}\alpha_m\right)$$

is free and proper.

Consider the vector subbundle $\mathcal{K} := \mathcal{V} \oplus \{0\} \subset TM \oplus T^*M$ of the Pontryagin bundle and its orthogonal $\mathcal{K}^\perp = TM \oplus \mathcal{V}^\circ$. Both vector subbundles are G-invariant and it is easy to show (in agreement with the more general results of [7]) that

$$\frac{\mathcal{K}^\perp}{\mathcal{K}}\Big/ G = \frac{TM \oplus \mathcal{V}^\circ}{\mathcal{V} \oplus \{0\}}\Big/ G = \frac{TM}{\mathcal{V}} \oplus \mathcal{V}^\circ \Big/ G \tag{9}$$

is a Courant algebroid over $\bar{M}$ with the symmetric bilinear 2-form that descends from the one on $\mathcal{K}^\perp/\mathcal{K}$ given by

$$\langle (\widehat{X}, \alpha), (\widehat{Y}, \beta)\rangle_{\mathcal{K}^\perp/\mathcal{K}} = \beta(X) + \alpha(Y) \tag{10}$$

for all α, β in $\Gamma(\mathcal{V}^\circ)$ and X, Y in $\mathfrak{X}(M)$; here $\widehat{X} := X(\operatorname{mod}\mathcal{V})$, $\widehat{Y} := Y(\operatorname{mod}\mathcal{V})$ denote local sections of $TM/\mathcal{V}$ induced by local vector fields on M.

We have used above the following general fact that will be needed also in later arguments. The proof is straightforward.

LEMMA 1. *Let $\pi : E \to M$ be a smooth vector bundle over M. Assume that there are two free proper G-actions on E and M, respectively, such that π is equivariant and the action on E is linear on the fibers. Then the induced map $\pi_G : E/G \to M/G$ defined by the commutative diagram*

$$\begin{array}{ccc} E & \xrightarrow{\pi} & M \\ {\scriptstyle \pi_E}\downarrow & & \downarrow{\scriptstyle \pi_M} \\ E/G & \xrightarrow{\pi_G} & M/G \end{array}$$

is also a smooth vector bundle whose rank is equal to the rank of E.

In fact, with the identifications given above of $\Gamma(\mathcal{V}^\circ)^G$ with $\Omega^1(\bar{M})$ and $\Gamma(TM/\mathcal{V})^G$ with $\mathfrak{X}(\bar{M})$, it is obvious that the G-equivariant sections of (9) are in one-to-one correspondence with those of $T\bar{M} \oplus T^*\bar{M}$. Note that this says that we have a vector bundle isomorphism

$$\frac{\mathcal{K}^\perp}{\mathcal{K}} \Big/ G \simeq T\bar{M} \oplus T^*\bar{M} \tag{11}$$

over $\bar{M} = M/G$. This vector bundle isomorphism preserves the symmetric pairing and it is easy to check that the Courant bracket on $T\bar{M} \oplus T^*\bar{M}$ also descends from the Courant bracket on $TM \oplus T^*M$.

We summarize here the approach for Dirac reduction in [7]. Assuming that $D \cap \mathcal{K}^\perp$ has constant rank, that is, $D \cap \mathcal{K}^\perp$ is a smooth vector subbundle of $TM \oplus T^*M$, it follows that $(D \cap \mathcal{K}^\perp)^\perp = D + \mathcal{K}$ and $D \cap \mathcal{K}$ are vector subbundles of $TM \oplus T^*M$. Form the pointwise quotient

$$\frac{(D \cap \mathcal{K}^\perp) + \mathcal{K}}{\mathcal{K}} = \frac{(D \cap (TM \oplus \mathcal{V}^\circ)) + (\mathcal{V} \oplus \{0\})}{\mathcal{V} \oplus \{0\}} \tag{12}$$

with base M. At each point $m \in M$, one gets a subspace of the vector space $(T_mM/\mathcal{V}(m)) \oplus \mathcal{V}^\circ(m) \simeq \mathcal{K}^\perp(m)/\mathcal{K}(m)$ (see (9)).

PROPOSITION 1. *Relative to the symmetric nondegenerate bilinear form* (10) *on* $\mathcal{K}^\perp/\mathcal{K}$, *the vector subspace*

$$\tilde{D}(m) := \frac{(D(m) \cap \mathcal{K}(m)^\perp) + \mathcal{K}(m)}{\mathcal{K}(m)} \qquad \textit{of} \qquad \frac{\mathcal{K}(m)^\perp}{\mathcal{K}(m)}$$

satisfies $\tilde{D}(m) = \tilde{D}(m)^\perp$.

Proof: Let us prove that $\tilde{D}(m) \subseteq \tilde{D}(m)^\perp$. Let $(\widehat{X}(m), \alpha(m)) \in \tilde{D}(m)$. If $(\widehat{X}, \alpha) \in \Gamma(\tilde{D})$ are local sections about m, then $\alpha \in \Gamma(\mathcal{V}^\circ)$ and there are $X \in \mathfrak{X}(M)$ and $V \in \Gamma(\mathcal{V})$ such that $(X + V, \alpha) \in \Gamma(D)$ and $\widehat{X} = X (\operatorname{mod} \mathcal{V})$. For all $(\widehat{Y}, \beta) \in \Gamma(\tilde{D})$ we have analogously local vector fields $Y \in \mathfrak{X}(M)$ and $W \in \Gamma(\mathcal{V})$ such that $(Y + W, \beta) \in \Gamma(D)$ and $\widehat{Y} = Y (\operatorname{mod} \mathcal{V})$. This yields

$$\langle (\widehat{X}, \alpha), (\widehat{Y}, \beta) \rangle_{\mathcal{K}^\perp/\mathcal{K}} \overset{(10)}{=} \langle (X + V, \alpha), (Y + W, \beta) \rangle = 0,$$

since $(X + V, \alpha), (Y + W, \beta) \in \Gamma(D)$.

To prove the inclusion, $\tilde{D}(m)^\perp \subseteq \tilde{D}(m)$ let $(\widehat{X}(m), \alpha(m)) \in \tilde{D}(m)^\perp$ be such that $(\widehat{X}, \alpha) \in \Gamma(\mathcal{K}^\perp/\mathcal{K})$ and for all $(\widehat{Y}, \beta) \in \Gamma(\tilde{D})$ we have $\langle (\widehat{X}, \alpha), (\widehat{Y}, \beta) \rangle_{\mathcal{K}^\perp/\mathcal{K}} = 0$. Choose $X \in \mathfrak{X}(M)$ such that $\widehat{X} = X (\operatorname{mod} \mathcal{V})$. For all $(Y, \beta) \in \Gamma(D \cap \mathcal{K}^\perp)$, $(\widehat{Y}, \beta)$ lies in $\Gamma(\tilde{D})$ and we get

$$0 = \langle (\widehat{X}, \alpha), (\widehat{Y}, \beta) \rangle_{\mathcal{K}^\perp/\mathcal{K}} = \langle (X, \alpha), (Y, \beta) \rangle = \alpha(Y) + \beta(X).$$

This yields $(X, \alpha) \in \Gamma((D \cap \mathcal{K}^\perp)^\perp)$. We have $(D_q \cap \mathcal{K}_q^\perp)^\perp = D_q^\perp + (\mathcal{K}_q^\perp)^\perp = D_q + \mathcal{K}_q$ for every q in the domain of definition of (X, α). Thus, since D and $\mathcal{K}$ are smooth vector bundles, there exists $X' \in \mathfrak{X}(M)$ and $W \in \Gamma(\mathcal{V})$ such that $(X', \alpha) \in \Gamma(D)$ and $X = X' + W$. Now recall that the 1-form α is in fact in $\Gamma(\mathcal{V}^\circ)$ since $(\widehat{X}, \alpha)$ was an element of $\Gamma(\mathcal{K}^\perp/\mathcal{K})$. The pair (X', α) is consequently in $\Gamma(D \cap \mathcal{K}^\perp)$ and, since $\widehat{X} = (X' + W)(\operatorname{mod} \mathcal{V}) = X'(\operatorname{mod} \mathcal{V})$, our $(\widehat{X}, \alpha)$ is a local section of $\tilde{D}$, as required. □

This proposition immediately implies that $\dim \tilde{D}(m)$ is constant on M and equal to

$$\frac{\dim \mathcal{K}^\perp(m) - \dim \mathcal{K}(m)}{2} = \frac{\dim M + (\dim M - \dim G) - \dim G}{2} = \dim M - \dim G.$$

Thus $\tilde{D}$ is a smooth G-invariant subbundle of $\mathcal{K}^\perp/\mathcal{K}$. Its image by the isomorphism (11) gives a subbundle $D_{\text{red}} = \tilde{D}/G$ of

$$\frac{\mathcal{K}^\perp}{\mathcal{K}} \Big/ G \simeq T\bar{M} \oplus T^*\bar{M},$$

whose rank is $(\dim M - \dim G)$, which is isotropic relative to the symmetric pairing on $T\bar{M} \oplus T^*\bar{M}$. Hence D_{red} is a Dirac structure called the *reduction of D by G*. This discussion and Proposition 1 yield that the sections of D_{red} are in one-to-one correspondence with the G-equivariant sections of the quotient (12) via the isomorphisms $\mathfrak{X}(\bar{M}) \simeq \Gamma(TM/\mathcal{V})^G$ and $\Omega^1(\bar{M}) \simeq \Gamma(\mathcal{V}^\circ)^G$ given at the beginning of this subsection.

It is customary to denote the "quotient" Dirac structure on M/G by

$$D_{\text{red}} = \frac{(D \cap \mathcal{K}^\perp) + \mathcal{K}}{\mathcal{K}} \Big/ G.$$

It is easy to check that if the Dirac structure D is integrable, then the reduced Dirac structure is also integrable (see also [32]).

There is an other formulation in terms of smooth sections for the reduction of Dirac manifolds and implicit Hamiltonian systems, due to [3] and [5] (see [4] for the singular case). This extension of the Poisson reduction was historically the first method to reduce Dirac structures.

In [5], there is the additional assumption that $\mathcal{V} + \mathsf{G}_0$ is constant dimensional on M. It turns out that no hypothesis on G_0 is needed; the proof of the theorem in [5] works with the same assumptions as in [7] if one just uses the results of [18] applied to $D \cap \mathcal{K}^\perp$, assuming only that this intersection is smooth (see [16]).

The formula for the reduced Dirac structure given in [5] is

$$\Gamma(D_{\text{red}}) = \Big\{ \big(\bar{X}, \bar{\alpha}\big) \in \Gamma(T\bar{M} \oplus T^*\bar{M}) \mid \text{there is } X \in \mathfrak{X}(M) \text{ such that } X \sim_\pi \bar{X} \text{ and } (X, \pi^*\bar{\alpha}) \in \Gamma(D) \Big\}. \quad (13)$$

The description of D_{red} shows that the smooth distribution $\mathsf{G}_0/\mathcal{V}$ projects to $\mathsf{G}_0^{\text{red}}$ and that the smooth codistribution $\pi_2(D \cap (\mathcal{V} \oplus \mathcal{V}^\circ))$ projects to $\mathsf{P}_0^{\text{red}}$, where π_2 is the projection $\pi_2 : TM \oplus T^*M \to T^*M$. There is no analogous description as quotients of the distribution $\mathsf{G}_1^{\mathrm{red}}$, and $\mathsf{P}_1^{\mathrm{red}}$; they need to be computed from the definition on a case by case basis.

Depending on the example, one needs to choose which method of Dirac reduction is easier to implement. In the next section, we will present cases where we have global bases of sections for the Dirac structure and in that situation the first method is more convenient.

There is a third method of reduction that is due to [35–37]. It is undergoing a major extension to encompass both the Lagrangian and Hamiltonian version of classical reduction (see [11]). Since this work is still in progress we shall not comment on it here.

3. Reduction of nonholonomic systems

3.1. Summary of the nonholonomic reduction method

The authors of [2] propose a reduction method for constrained Hamiltonian systems. They start with the configuration space Q, a hyperregular Lagrangian $L : TQ \to \mathbb{R}$ taken as the kinetic energy of a Riemannian metric minus a potential, and a *constraint distribution* $\mathcal{D}$ on Q equal to the kernel of smooth 1-forms $\phi^1, \ldots, \phi^k \in \Omega^1(Q)$ satisfying pointwise $\phi^1 \wedge \ldots \wedge \phi^k \neq 0$, that is,

$$\mathcal{D} := \{v \in TQ \mid \phi^j(v) = 0,\ j = 1, \ldots, k\}.$$

The independence of the forms ensures that $\mathcal{D}$ is a smooth vector subbundle of TQ.

Denote by $\langle \cdot, \cdot \rangle : T^*Q \times TQ \to \mathbb{R}$ the duality pairing between 1-forms and tangent vectors. Let $\mathbb{F}L : TQ \to T^*Q$,

$$\langle \mathbb{F}L(v), w \rangle := \left.\frac{d}{dt}\right|_{t=0} L(v + tw), \qquad v, w \in T_qQ,$$

be the Legendre transformation associated to L which is a diffeomorphism since the Lagrangian is hyperregular. If $A(v) := \langle \mathbb{F}L(v), v \rangle$ denotes the action of L, let $H(p) := A((\mathbb{F}L)^{-1}(p)) - L((\mathbb{F}L)^{-1}(p))$, $p \in T^*Q$, be the associated Hamiltonian. The Hamiltonian vector field X determined by H and the constraint forms $\phi^1, \ldots, \phi^k \in \Omega^1(Q)$ is defined classically by

$$\dot{q} = \frac{\partial H}{\partial p}, \qquad \dot{p} = \frac{\partial H}{\partial q} + \lambda_j \phi^j$$

or

$$\mathbf{i}_X \omega_{\mathrm{can}} = \mathbf{d}H + \lambda_j {\pi_{T^*Q}}^* \phi^j \tag{14}$$

where $\pi_{T^*Q} : T^*Q \to Q$ is the cotangent bundle projection and $\lambda_1, \ldots \lambda_k \in C^\infty(Q)$ are the Lagrange multipliers associated to the constraint forms $\phi^1, \ldots, \phi^k$, and the constraint equations

$$\phi^j(\mathbb{F}L^{-1}(p)) = \phi^j(T\pi_{T^*Q}X) = 0, \qquad p \in T^*Q. \tag{15}$$

The counterpart of the constraint distribution $\mathcal{D}$ in phase space is the *constraint manifold*

$$M := \mathbb{F}L(\mathcal{D}) = \left\{p \in T^*Q \mid \phi^j(\pi_{T^*Q}(p))\left((\mathbb{F}L)^{-1}(p)\right) = 0,\ j = 1, \ldots, k\right\} \subset T^*Q. \tag{16}$$

Since we require that the solution is in the constraint submanifold M, it follows that X is tangent to M.

Set $\omega_M := i^*\omega_{\mathrm{can}}$, where $i : M \hookrightarrow T^*Q$ is the inclusion and ω_{can} is the canonical symplectic form on T^*Q. Define

$$\mathcal{F} := \{U \in TT^*Q \mid \pi^*_{T^*Q}\phi^j(U) = 0,\ j = 1, \ldots, k\} \tag{17}$$

and note that $\pi^*_{T^*Q}\phi^1 \wedge \ldots \wedge \pi^*_{T^*Q}\phi^k \neq 0$ on T^*Q. Therefore $\mathcal{F} \to T^*Q$ is a vector subbundle of TT^*Q. The *nonholonomic horizontal distribution* is defined by

$$\mathcal{H} := \mathcal{F} \cap TM \to M. \tag{18}$$

L. Bates and J. Śniatycki in [2] prove that the restriction $\omega_{\mathcal{H}}$ of ω_M to $\mathcal{H} \times \mathcal{H}$ is nondegenerate. (Their proof uses the fact that the Lagrangian is the kinetic energy of a metric minus a potential.) They also show that $\mathcal{H}$ is a vector subbundle of TM. With the condition (15) on X, we get for $j = 1, \ldots, k$,

$$\pi^*_{T^*Q}\phi^j(X) = \phi^j(T\pi_{T^*Q}X) = 0$$

and thus the vector field X is a section of $\mathcal{H}$. Hence it is easy to see that the pull back to M of (14) subject to the constraints (15) is equivalent to $X \in \Gamma(\mathcal{H})$ and $\mathbf{i}_X\omega_{\mathcal{H}} = \mathbf{d}H|_{\mathcal{H}}$.

Let G be a Lie group acting symplectically on T^*Q (not necessarily the lift of an action on Q), that leaves M invariant and preserves the Hamiltonian H. Furthermore, assume that the quotient $\bar{M} = M/G$ is a smooth manifold with projection map $\pi : M \to \bar{M}$, a submersion. Since G is a symmetry group of the nonholonomic system, all intrinsically defined vector fields and distributions push down to $\bar{M}$.

In particular, the vector field X on M pushes down to a vector field $\bar{X}$ with $X \sim_\pi \bar{X}$ and the distribution $\mathcal{H}$ pushes down to a distribution $\mathcal{H}_{\mathrm{red}}$ on $\bar{M}$. However, $\omega_{\mathcal{H}}$ need not push down to a 2-form defined on $\mathcal{H}_{\mathrm{red}}$ on $\bar{M}$, despite the fact that $\omega_{\mathcal{H}}$ is G-invariant. This is because there may be infinitesimal symmetries ξ_M which are horizontal (that is, take values in $\mathcal{H}$), but such that $\mathbf{i}_{\xi_M}\omega_{\mathcal{H}} \neq 0$. Let $\mathcal{V}$ be the distribution on M tangent to the orbits of G, that is, its fibers are $\mathcal{V}(m) := \{\xi_M(m) \mid \xi \in \mathfrak{g}\}$ for all $m \in M \subseteq T^*Q$. Define the *horizontal annihilator* $\mathcal{U}$ of $\mathcal{V}$ by

$$\mathcal{U} = (\mathcal{V} \cap \mathcal{H})^{\omega_M} \cap \mathcal{H} \subseteq TM \subseteq TT^*Q, \tag{19}$$

where the superscript ω_M on a distribution denotes its fiberwise ω_M-orthogonal complement in TM. Clearly, $\mathcal{U}$ and $\mathcal{V}$ are both G-invariant, project down to $\bar{M}$, and the image of $\mathcal{V}$ is $\{0\}$. Define $\bar{\mathcal{H}} := T\pi\,(\mathcal{U}) \subseteq T\bar{M}$ to be the projection of $\mathcal{U}$ to $\bar{M}$. L. Bates and J. Śniatycki in [2] show that X takes values in $\mathcal{U}$ and that the restriction $\omega_{\mathcal{U}}$ of ω_M to $\mathcal{U} \times \mathcal{U}$ pushes down to a nondegenerate form $\omega_{\bar{\mathcal{H}}}$ on

$\bar{\mathcal{H}}$, i.e. $\pi^*\omega_{\bar{\mathcal{H}}} = \omega_{\mathcal{U}}$. In addition, the function $\bar{H} \in C^\infty(\bar{M})$ defined by $\pi^*\bar{H} = H|_M$ and the induced vector field $\bar{X}$ on $\bar{M}$ are related by

$$\mathbf{i}_{\bar{X}}\omega_{\bar{\mathcal{H}}} = \mathbf{d}\bar{H}|_{\bar{\mathcal{H}}} \tag{20}$$

which can be interpreted as the definition of the reduced nonholonomic Hamiltonian vector field $\bar{X}$.

REMARK 1. Note that we have no information about the dimensions of the fibers of $\mathcal{U}$. In general, $\mathcal{U}$ is *not* a vector subbundle of TM. In the following, we will often assume that $\mathcal{V} \cap \mathcal{H}$ has constant rank on the manifold M. In this case, $\mathcal{U} = (\mathcal{V} \cap \mathcal{H})^{\omega_M} \cap \mathcal{H} = (\mathcal{V} \cap \mathcal{H})^{\omega_{\mathcal{H}}}$ also has constant rank on M, since $\omega_{\mathcal{H}}$ is nondegenerate. △

3.2. Link with Dirac reduction

Let M, ω_M, π_{T^*Q}, $\mathcal{H}$, $\bar{M}$, and $\pi : M \to \bar{M}$ be as in the preceding subsection. An easy verification shows that

$$\mathcal{H} = (T(\pi_{T^*Q}|_M))^{-1}(\mathcal{D}) \subseteq TM \subseteq TT^*Q, \tag{21}$$

where

$$\mathcal{D} := \{v \in TQ \mid \langle \phi^j, v\rangle = 0,\ j = 1, \ldots, k\} \subseteq TQ$$

is the constraint distribution on Q.

We introduce the Dirac structure D on M as in [36]:

$$D(m) = \left\{(X(m), \alpha(m)) \in TM \oplus T^*M \mid X \in \Gamma(\mathcal{H}),\ \alpha - \mathbf{i}_X\omega_M \in \Gamma(\mathcal{H}^\circ),\ \alpha \in \Omega^1(M)\right\} \tag{22}$$

for all $m \in M$.

The Lie group G acts on M and leaves $\mathcal{H}$, ω_M, and thus the Dirac structure D invariant. Define $\mathcal{K} := \mathcal{V} \oplus \{0\} \subset TM \oplus T^*M$ and its orthogonal complement $\mathcal{K}^\perp = TM \oplus \mathcal{V}^\circ$ as in Section 2.3. Assume, as in Section 2.3, that $D \cap \mathcal{K}^\perp$ is a vector subbundle of $TM \oplus T^*M$ and consider the reduced Dirac manifold $(\bar{M}, D_{\mathrm{red}})$. The next proposition shows that, if $\bar{\mathcal{H}}$ is constant dimensional, the reduced Dirac structure is given by the formula

$$D_{\mathrm{red}} = \left\{(X, \alpha) \in \Gamma(T\bar{M} \oplus T^*\bar{M}) \mid X \in \Gamma(\bar{\mathcal{H}}),\ \alpha \in \Omega^1(\bar{M}),\ \alpha - \mathbf{i}_X\omega_{\bar{\mathcal{H}}} \in \Gamma\left(\bar{\mathcal{H}}^\circ\right)\right\}$$

where $\bar{\mathcal{H}}$ and $\omega_{\bar{\mathcal{H}}}$ are defined as in the preceding subsection.

PROPOSITION 2. *Let D be as above and assume that $\mathcal{H} \cap \mathcal{V}$ has constant rank on M.*

(i) *The associated generalized distribution G_0 is trivial and the codistribution P_1 is given by $\mathsf{P}_1 = T^*M$.*

(ii) *Let $\mathcal{U} = \mathcal{H} \cap (\mathcal{V} \cap \mathcal{H})^{\omega_M}$ (see (19)). Then*

$$X \in \Gamma(\mathcal{U}) \iff \text{there exists } \alpha \in \Gamma(\mathcal{V}^\circ) \text{ such that } (X, \alpha) \in \Gamma(D \cap \mathcal{K}^\perp). \tag{23}$$

With the additional assumption that $\mathcal{V} + \mathcal{H} = TM$, the section α in (23) is unique.

(iii) *The reduced distributions* $\mathsf{G}_1^{\mathrm{red}}$ *and* $\mathsf{G}_0^{\mathrm{red}}$ *are given by*

$$\mathsf{G}_1^{\mathrm{red}} = \bar{\mathcal{H}} \qquad \text{and} \qquad \mathsf{G}_0^{\mathrm{red}} = \{0\}.$$

(iv) *For each* $\alpha \in \Gamma(\mathcal{V}^\circ)$ *there exists exactly one section* $X \in \Gamma(\mathcal{U})$ *such that* $(X, \alpha) \in \Gamma(D)$. *Hence, we have* $\pi_2(D \cap \mathcal{K}^\perp) = \mathcal{V}^\circ$ *and the reduced codistribution* $\mathsf{P}_1^{\mathrm{red}}$ *is equal to* $T^*(M/G)$.

(v) *Assume that* $\mathsf{G}_1^{\mathrm{red}} = \bar{\mathcal{H}}$ *is constant dimensional. The 2-form defined on* $\mathsf{G}_1^{\mathrm{red}} = \bar{\mathcal{H}}$ *by the Dirac structure* D_{red} *(see* (3)*) is nondegenerate and is equal to* $\omega_{\bar{\mathcal{H}}}$.

Proof: (i) If X is a section of G_0, we have $\mathbf{i}_X\omega_M \in \Gamma(\mathcal{H}^\circ)$ and $X \in \Gamma(\mathcal{H})$. Hence, since $\omega_{\mathcal{H}}$ is nondegenerate, the vector field X has to be the zero section. Thus $\mathsf{G}_0 = \{0\}$. Since the 2-form $\omega_{\mathcal{H}}$ is nondegenerate, an arbitrary $\alpha \in \Omega^1(M)$ determines a unique section X of $\mathcal{H}$ by the equation $\mathbf{i}_X\omega_{\mathcal{H}} = \alpha|_{\mathcal{H}}$. Therefore, $\mathsf{P}_1 = T^*M$.

(ii) If (X, α) is a local section of $D \cap \mathcal{K}^\perp$, then we have $X \in \Gamma(\mathcal{H})$, $\alpha \in \Gamma(\mathcal{V}^\circ)$, and $\alpha = \mathbf{i}_X\omega_M$ on $\mathcal{H}$. Hence, $(\mathbf{i}_X\omega_M)|_{\mathcal{H}\cap\mathcal{V}} = 0$ and thus we have

$$X \in \Gamma(\mathcal{H} \cap (\mathcal{V} \cap \mathcal{H})^{\omega_M}) = \Gamma(\mathcal{U}).$$

Conversely, if $X \in \Gamma(\mathcal{U})$, we have $\mathbf{i}_X\omega_M = 0$ on $\mathcal{V}\cap\mathcal{H}$. Since, by hypothesis, $\mathcal{V}\cap\mathcal{H}$ has constant rank, it is a vector subbundle so we can find a section $\alpha \in \Gamma(\mathcal{V}^\circ)$ such that the restriction of α and $\mathbf{i}_X\omega_M$ to $\mathcal{H}$ are equal.

If, in addition, we make the usual assumption $\mathcal{V} + \mathcal{H} = TM$, we have for each $X \in \Gamma(\mathcal{U})$ exactly one $\alpha \in \Omega^1(M)$ such that $\alpha|_{\mathcal{H}} = \mathbf{i}_X\omega_M$ and $\alpha|_{\mathcal{V}} = 0$.

(iii) By construction, the constraint distribution $\mathsf{G}_1^{\mathrm{red}}$ associated to the Dirac structure D_{red} on $\bar{M}$ is given by

$$\frac{\mathcal{U} + \mathcal{V}}{\mathcal{V}} \Big/ G.$$

This can obviously be identified with

$$\bar{\mathcal{H}} = T\pi(\mathcal{U}).$$

If we have $\bar{X} \in \Gamma(\mathsf{G}_0^{\mathrm{red}})$, then $(\bar{X}, 0) \in \Gamma(D_{\mathrm{red}})$ and there exists $X \in \mathfrak{X}(M)$ with $X \sim_\pi \bar{X}$ and $(X, 0) \in \Gamma(D)$. Hence we have $X \in \Gamma(\mathsf{G}_0)$ and since $\mathsf{G}_0 = \{0\}$, we get $X = 0$. This shows that $\mathsf{G}_0^{\mathrm{red}} = \{0\}$.

(iv) This follows directly from (i) and (ii).

(v) Let $\omega_{D_{\mathrm{red}}}$ be the 2-form defined on $\mathsf{G}_1^{\mathrm{red}} = \bar{\mathcal{H}}$ by the Dirac structure D_{red} (see (3)). If $X \in \Gamma(\bar{\mathcal{H}})$ is such that $\omega_{D_{\mathrm{red}}}(X, Y) = 0$ for all $Y \in \Gamma(\bar{\mathcal{H}})$, then $(X, 0)$ is a section of D_{red} and hence we find by (iii) that $X = 0$. Thus $\omega_{D_{\mathrm{red}}}$ is nondegenerate on $\bar{\mathcal{H}}$. Let $\bar{X}$ and $\bar{Y}$ be sections of $\bar{\mathcal{H}}$. We show now that $\omega_{D_{\mathrm{red}}}(\bar{X}, \bar{Y}) = \omega_{\bar{\mathcal{H}}}(\bar{X}, \bar{Y})$. Indeed, by definition, we have $\omega_{D_{\mathrm{red}}}(\bar{X}, \bar{Y}) = \bar{\alpha}(\bar{Y})$, where $\bar{\alpha}, \bar{\beta} \in \Omega^1(M/G)$ are such that $(\bar{X}, \bar{\alpha}), (\bar{Y}, \bar{\beta}) \in \Gamma(D_{\mathrm{red}})$. Choose $X, Y \in \Gamma(\mathcal{U})$ with $X \sim_\pi \bar{X}$, $Y \sim_\pi \bar{Y}$ and

$(X, \pi^*\bar{\alpha})$, $(Y, \pi^*\bar{\beta}) \in \Gamma(D \cap \mathcal{K}^\perp)$. Then we have

$$\omega_{D_{\text{red}}}(\bar{X}, \bar{Y}) = \bar{\alpha}(\bar{Y}) = (\pi^*\bar{\alpha})(Y) = \omega_{\mathcal{U}}(X, Y) = \omega_{\bar{\mathcal{H}}}(\bar{X}, \bar{Y}),$$

where the last equality follows simply from the definition of $\omega_{\bar{\mathcal{H}}}$. □

We shall use part (ii) of this proposition to simplify certain computations in the examples that follow.

REMARK 2. Note that if $\mathcal{H} \cap \mathcal{V}$ has constant rank on M, we have automatically that $D \cap \mathcal{K}^\perp$ has constant dimensional fibers on M.

This is proved in the following way. Since $\mathcal{H}$, $\mathcal{V}$, $\mathcal{H} \cap \mathcal{V}$ are vector subbundles of TM, $\mathcal{H} + \mathcal{V}$ is also a subbundle of TM. By the nondegeneracy of $\omega_{\mathcal{H}}$, we get that $\mathcal{U} = (\mathcal{H} \cap \mathcal{V})^{\omega_M} \cap \mathcal{H} = (\mathcal{H} \cap \mathcal{V})^{\omega_{\mathcal{H}}}$ has also constant dimensional fibers on M and is in particular a vector subbundle of $\mathcal{H}$. Let u be the dimension of the fibers of $\mathcal{U}$, r the dimension of the fibers of $\mathcal{H}$. Then, if $n = \dim M$, $n - r$ is the rank of the codistribution $\mathcal{H}^\circ$. Let finally l be the rank of the codistribution $\mathcal{H}^\circ \cap \mathcal{V}^\circ = (\mathcal{V} + \mathcal{H})^\circ \subseteq \mathcal{H}^\circ$. Choose local basis vector fields $H_1, \ldots, H_r$ for $\mathcal{H}$ such that $H_1, \ldots, H_u$ are basis vector fields for $\mathcal{U}$. In the same way, choose basis 1-forms $\beta_1, \ldots, \beta_{n-r}$ for $\mathcal{H}^\circ$ such that $\beta_1, \ldots, \beta_l$ are basis 1-forms for $\mathcal{V}^\circ \cap \mathcal{H}^\circ$. Then a local basis of sections of D is

$$\left\{(H_1, \mathbf{i}_{H_1}\omega_M), \ldots, (H_r, \mathbf{i}_{H_r}\omega_M), (0, \beta_1), \ldots, (0, \beta_{n-r})\right\}.$$

The considerations above show that $D \cap \mathcal{K}^\perp$ is then spanned by the sections

$$\left\{\left(H_1, \mathbf{i}_{H_1}\omega_M + \sum_{i=l+1}^{n-r} a_1^i\beta_i\right), \ldots, \left(H_u, \mathbf{i}_{H_u}\omega_M + \sum_{i=l+1}^{n-r} a_u^i\beta_i\right), (0, \beta_1), \ldots, (0, \beta_l)\right\},$$

where a_i^j are smooth functions chosen such that $\mathbf{i}_{H_j}\omega_M + \sum_{i=l+1}^{n-r} a_j^i\beta_i \in \Gamma(\mathcal{V}^\circ)$ for $j = 1, \ldots, u$. Since these sections are linearly independent, they are smooth local basis sections for $D \cap \mathcal{K}^\perp$. △

3.3. Examples

3.3.1. The constrained particle in space

L. Bates and J. Śniatycki in [2] study the motion of the constrained particle in space. The configuration space of this problem is $Q := \mathbb{R}^3$ whose coordinates are denoted by $\mathbf{q} := (x, y, z)$. They take the following concrete constraints on the velocities,

$$\mathcal{D} := \ker(\mathbf{d}z - y\mathbf{d}x) = \left\{ v_x\partial_x + v_y\partial_y + v_z\partial_z \;\middle|\; v_z - yv_x = 0\right\} \subset TQ.$$

The Lagrangian taken to be the kinetic energy of the Euclidean metric, that is, $L(\mathbf{q}, \mathbf{v}) := \frac{1}{2}\|\mathbf{v}\|^2$, and it is hyperregular. Hence the constraint manifold (16) is five-dimensional and given by

$$M := \{(x, y, z, p_x, p_y, p_z) \mid p_z = yp_x\} \subseteq T^*Q,$$

where (x, y, z, p_x, p_y, p_z) are the coordinates of T^*Q. The global coordinates on M are thus (x, y, z, p_x, p_y). The pull back ω_M of the canonical 2-form ω on T^*Q to M has hence the expression

$$\omega_M = \mathbf{d}x \wedge \mathbf{d}p_x + \mathbf{d}y \wedge \mathbf{d}p_y + \mathbf{d}z \wedge (p_x\mathbf{d}y + y\mathbf{d}p_x).$$

The Dirac structure D modeling this problem is given by (22). Formula (21) gives the vector subbundle

$$\mathcal{H} := (T(\pi_{T^*Q}|_M))^{-1}(\mathcal{D}) = \operatorname{span}\{\partial_x + y\partial_z, \partial_y, \partial_{p_x}, \partial_{p_y}\} \subset TM,$$

and consequently

$$\mathcal{H}^\circ = \operatorname{span}\{\mathbf{d}z - y\mathbf{d}x\}.$$

A computation yields

$$\mathbf{i}_{\partial_x + y\partial_z}\omega_M = (1+y^2)\mathbf{d}p_x + yp_x\mathbf{d}y, \qquad \mathbf{i}_{\partial_y}\omega_M = \mathbf{d}p_y - p_x\mathbf{d}z,$$
$$\mathbf{i}_{\partial_{p_y}}\omega_M = -\mathbf{d}y, \qquad \mathbf{i}_{\partial_{p_x}}\omega_M = -y\mathbf{d}z - \mathbf{d}x.$$

Hence

$$\Big\{ \big(\partial_x + y\partial_z, (1+y^2)\mathbf{d}p_x + yp_x\mathbf{d}y\big)\,;\, \big(\partial_y, \mathbf{d}p_y - p_x\mathbf{d}z\big)\,;\, \big(\partial_{p_y}, -\mathbf{d}y\big)\,; \\ \big(\partial_{p_x}, -y\mathbf{d}z - \mathbf{d}x\big)\,;\, (0, \mathbf{d}z - y\mathbf{d}x) \Big\} \qquad (24)$$

is a smooth global basis for D.

We consider the action of the Lie group $G = \mathbb{R}^2$ on M given by

$$\Phi : G \times M \to M, \qquad \Phi((r, s), m) = (x + r, y, z + s, p_x, p_y),$$

where $\mathbf{m} := (x, y, z, p_x, p_y) \in M$. This $\mathbb{R}^2$-action is the restriction to M of the cotangent lift of the action $\phi : G \times Q \to Q$, $\phi((r, s), (x, y, z)) = (x + r, y, z + s)$. It obviously leaves the Hamiltonian $H(\mathbf{m}) = \frac{1}{2}((1 + y^2)p_x^2 + p_y^2)$ on M invariant. Note that if $(X, \alpha) \in \Gamma(D)$ we have

$$(\pounds_{\xi_M} X, \pounds_{\xi_M}\alpha) \in \Gamma(D) \qquad \text{for all} \qquad \xi \in \mathfrak{g} = \mathbb{R}^2.$$

Since the vertical bundle in this example is $\mathcal{V} = \operatorname{span}\{\partial_x, \partial_z\}$, we have

$$\mathcal{K} = \mathcal{V} \oplus \{0\} = \operatorname{span}\{(\partial_x, 0), (\partial_z, 0)\} \subset TM \oplus T^*M$$

and thus

$$\mathcal{K}^\perp = TM \oplus \mathcal{V}^\circ \\ = \operatorname{span}\{(\partial_x, 0), (\partial_y, 0), (\partial_z, 0), (\partial_{p_x}, 0), (\partial_{p_y}, 0), (0, \mathbf{d}y), (0, \mathbf{d}p_x), (0, \mathbf{d}p_y)\}. \qquad (25)$$

A direct computation using (24) and (25) yields

$$D \cap \mathcal{K}^\perp = \operatorname{span}\Big\{\big(\partial_{p_y}, -\mathbf{d}y\big)\,,\, \big(\partial_x + y\partial_z, (1+y^2)\mathbf{d}p_x + yp_x\mathbf{d}y\big)\,, \\ \big((1+y^2)\partial_y - yp_x\partial_{p_x}, (1+y^2)\mathbf{d}p_y\big)\Big\}$$

and

$$(D \cap \mathcal{K}^{\perp}) + \mathcal{K}$$
$$= \operatorname{span}\left\{ \begin{array}{l} \left(\partial_{p_y}, -\mathbf{d}y\right), \quad (\partial_x, 0), \quad (\partial_z, 0), \quad \left(0, (1+y^2)\mathbf{d}p_x + yp_x\mathbf{d}y\right), \\ \left((1+y^2)\partial_y - yp_x\partial_{p_x}, (1+y^2)\mathbf{d}p_y\right) \end{array} \right\}$$

since in this case $(D \cap \mathcal{K}^{\perp}) \cap \mathcal{K} = \{0\}$.

Note that there is an easier way to compute the spanning sections of $D \cap \mathcal{K}^{\perp}$ by using (23). First, one determines spanning sections of $\mathcal{U}$. Second, for each spanning section $X \in \Gamma(\mathcal{U})$ we find $\lambda \in C^{\infty}(M)$ such that

$$\mathbf{i}_X\omega_M + \lambda(\mathbf{d}z - y\mathbf{d}x) \in \Gamma(\mathcal{V}^{\circ}).$$

Third, setting $\alpha := \mathbf{i}_X\omega_M + \lambda(\mathbf{d}z - y\mathbf{d}x)$ we have found a spanning section $(X, \alpha) \in \Gamma(D \cap \mathcal{K}^{\perp})$. In the following examples, we will proceed like this.

We get the reduced Dirac structure

$$D_{\mathrm{red}} = \frac{(D \cap \mathcal{K}^{\perp}) + \mathcal{K}}{\mathcal{K}} \Big/ G = \operatorname{span}\left\{ \begin{array}{l} \left(\partial_{p_y}, -\mathbf{d}y\right), \left(0, (1+y^2)\mathbf{d}p_x + yp_x\mathbf{d}y\right), \\ \left((1+y^2)\partial_y - yp_x\partial_{p_x}, (1+y^2)\mathbf{d}p_y\right) \end{array} \right\}$$

on the three-dimensional manifold $\bar{M} := M/G$ with global coordinates (y, p_y, p_x).

Since $\partial_x + y\partial_z$ is a spanning section of $\mathcal{H} \cap \mathcal{V}$, the distribution $\mathcal{U} \subset TM$ (see (19)) is given by

$$\begin{aligned} \mathcal{U} &= (\mathcal{V} \cap \mathcal{H})^{\omega_M} \cap \mathcal{H} = \ker\{\mathbf{i}_{\partial_x + y\partial_z}\omega_M\} \cap \mathcal{H} = \ker\{(1+y^2)\mathbf{d}p_x + yp_x\mathbf{d}y\} \cap \mathcal{H} \\ &= \operatorname{span}\left\{(1+y^2)\partial_y - yp_x\partial_{p_x}, \partial_x + y\partial_z, \partial_{p_y}\right\}. \end{aligned}$$

Thus

$$\bar{\mathcal{H}} = T\pi(\mathcal{U}) = \operatorname{span}\{\partial_{p_y}, (1+y^2)\partial_y - yp_x\partial_{p_x}\}$$

recovering the result in [2]. Note that, as discussed in Section 3.2, the distribution $\bar{\mathcal{H}} \subset T\bar{M}$ coincides with the projection on the first factor of the reduced Dirac structure (26). As in [2], $\bar{\mathcal{H}}$ is an integrable subbundle of $T\bar{M}$; in fact $[\partial_{p_y}, (1+y^2)\partial_y - yp_x\partial_{p_x}] = 0$. The 2-form $\omega_{\bar{\mathcal{H}}}$ is easily computed to equal

$$\omega_{\bar{\mathcal{H}}}(\partial_{p_y}, (1+y^2)\partial_y - yp_x\partial_{p_x}) = -\mathbf{d}y((1+y^2)\partial_y - yp_x\partial_{p_x}) = -(1+y^2).$$

As predicted by the general theory in Section 3.1, $\omega_{\bar{\mathcal{H}}}$ is nondegenerate.

It is easy to check that the reduced manifold $\bar{M}$ is Poisson relative to the 2-tensor

$$-\partial_y \wedge \partial_{p_y} + \frac{yp_x}{1+y^2}\partial_{p_x} \wedge \partial_{p_y},$$

or with Poisson bracket determined by $\{y, p_y\} = -1$, $\{y, p_x\} = 0$, $\{p_y, p_x\} = yp_x/(1+y^2)$, and that D_{red} given by (26) is the graph of the vector bundle homomorphism $\flat : T^*\bar{M} \to T\bar{M}$ associated to the Poisson structure.

3.3.2. The vertical rolling disk

This example is standard in the theory of nonholonomic mechanical systems; it can be found for example in [6]. Consider a vertical disk of zero width rolling on the xy-plane and free to rotate about its vertical axis. Let x and y denote the position of contact of the disk in the xy-plane. The remaining variables are θ and φ, denoting the orientation of a chosen material point P with respect to the vertical and the "heading angle" of the disk. Thus, the unconstrained configuration space for the vertical rolling disk is $Q := \mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1$. The Lagrangian for the problem is taken to be the kinetic energy

$$L(x, y, \theta, \varphi, \dot{x}, \dot{y}, \dot{\theta}, \dot{\varphi}) = \frac{1}{2}\mu(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I\dot{\theta}^2 + \frac{1}{2}J\dot{\varphi}^2,$$

where μ is the mass of the disk, and I, J are its moments of inertia. Hence, the Hamiltonian of the system is

$$H(x, y, \theta, \varphi, p_x, p_y, p_\theta, p_\varphi) = \frac{1}{2\mu}(p_x^2 + p_y^2) + \frac{1}{2I}p_\theta^2 + \frac{1}{2J}p_\varphi^2.$$

The rolling constraints may be written as $\dot{x} = R\dot{\theta}\cos\varphi$ and $\dot{y} = R\dot{\theta}\sin\varphi$, where R is the radius of the disk, that is,

$$\mathcal{D} := \{(x, y, \theta, \varphi, R\dot{\theta}\cos\varphi, R\dot{\theta}\sin\varphi, \dot{\theta}, \dot{\varphi}) \mid x, y \in \mathbb{R}, \theta, \varphi \in \mathbb{S}^1\} \subset TQ.$$

Note that the 1-forms defining this distribution $\mathcal{D}$ are $\phi_1 := \mathbf{d}x - R\cos\varphi\mathbf{d}\theta$ and $\phi_2 := \mathbf{d}y - R\sin\varphi\mathbf{d}\theta$.

The constraint manifold (16)

$$M := \left\{(x, y, \theta, \varphi, p_x, p_y, p_\theta, p_\varphi) \in T^*Q \;\middle|\; p_x = \frac{\mu R}{I}p_\theta\cos\varphi,\ p_y = \frac{\mu R}{I}p_\theta\sin\varphi\right\} \subseteq T^*Q$$

is in this example a graph over the coordinates $(x, y, \theta, \varphi, p_\theta, p_\varphi)$ and is hence six-dimensional. The induced 2-form $\omega_M = i^*\omega_{\text{can}}$ is given by the formula

$$\begin{aligned}\omega_M = \mathbf{d}x \wedge \left(\frac{\mu R\cos\varphi}{I}\mathbf{d}p_\theta - \frac{\mu R\sin\varphi}{I}p_\theta\mathbf{d}\varphi\right) \\ + \mathbf{d}y \wedge \left(\frac{\mu R\sin\varphi}{I}\mathbf{d}p_\theta + \frac{\mu R\cos\varphi}{I}p_\theta\mathbf{d}\varphi\right) + \mathbf{d}\theta \wedge \mathbf{d}p_\theta + \mathbf{d}\varphi \wedge \mathbf{d}p_\varphi,\end{aligned}$$

and the distribution $\mathcal{H} = \ker\{\mathbf{d}x - R\cos\varphi\mathbf{d}\theta,\ \mathbf{d}y - R\sin\varphi\mathbf{d}\theta\} \subseteq TM$ is in this case

$$\mathcal{H} = \operatorname{span}\{\partial_\varphi, \partial_\theta + R\cos\varphi\,\partial_x + R\sin\varphi\,\partial_y, \partial_{p_\theta}, \partial_{p_\varphi}\} \subset TM. \tag{26}$$

Therefore its annihilator is

$$\mathcal{H}^\circ = \operatorname{span}\{\mathbf{d}x - R\cos\varphi\mathbf{d}\theta, \mathbf{d}y - R\sin\varphi\mathbf{d}\theta\} \subset T^*M.$$

The Dirac structure on M describing the nonholonomic mechanical system is again given by (22). Since

$$\begin{aligned}
\mathbf{i}_{\partial_\varphi}\omega_M &= \mathbf{d}p_\varphi + \frac{\mu R\sin\varphi}{I}p_\theta\mathbf{d}x - \frac{\mu R\cos\varphi}{I}p_\theta\mathbf{d}y,\\
\mathbf{i}_{\partial_{p_\theta}}\omega_M &= -\frac{\mu R\cos\varphi}{I}\mathbf{d}x - \frac{\mu R\sin\varphi}{I}\mathbf{d}y - \mathbf{d}\theta,\\
\mathbf{i}_{\partial_{p_\varphi}}\omega_M &= -\mathbf{d}\varphi,
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{i}_{\partial_\theta+R\cos\varphi\partial_x+R\sin\varphi\partial_y}\omega_M &= \mathbf{d}p_\theta + R\cos\varphi\left(\frac{\mu R\cos\varphi}{I}\mathbf{d}p_\theta - \frac{\mu R\sin\varphi}{I}\mathbf{d}\varphi\right)\\
&\quad + R\sin\varphi\left(\frac{\mu R\sin\varphi}{I}\mathbf{d}p_\theta + \frac{\mu R\cos\varphi}{I}\mathbf{d}\varphi\right)\\
&= \left(1+\frac{\mu R^2}{I}\right)\mathbf{d}p_\theta,
\end{aligned}$$

we get again smooth global spanning sections of D:

$$\begin{gathered}
\left(\partial_\varphi, \mathbf{d}p_\varphi + \frac{\mu R\sin\varphi}{I}p_\theta\mathbf{d}x - \frac{\mu R\cos\varphi}{I}p_\theta\mathbf{d}y\right),\\
\left(\partial_{p_\theta}, -\frac{\mu R\cos\varphi}{I}\mathbf{d}x - \frac{\mu R\sin\varphi}{I}\mathbf{d}y - \mathbf{d}\theta\right),\\
\left(\partial_{p_\varphi}, -\mathbf{d}\varphi\right), \qquad (0, \mathbf{d}x - R\cos\varphi\mathbf{d}\theta), \qquad (0, \mathbf{d}y - R\sin\varphi\mathbf{d}\theta),\\
\left(\partial_\theta + R\cos\varphi\partial_x + R\sin\varphi\partial_y, \left(1+\frac{\mu R^2}{I}\right)\mathbf{d}p_\theta\right).
\end{gathered} \tag{27}$$

In this case, several groups of symmetries are studied in the literature.

1. *The case* $G = \mathbb{R}^2$ ([10]).

The Lie group $\mathbb{R}^2$ acts on M by

$$(r,s)\cdot(x,y,\theta,\varphi,p_\theta,p_\varphi) = (x+r, y+s, \theta, \varphi, p_\theta, p_\varphi)$$

and clearly leaves the Hamiltonian H invariant. The distribution $\mathcal{V}$ on M is in this case $\mathcal{V} = \operatorname{span}\{\partial_x, \partial_y\}$, so that $\mathcal{V}\cap\mathcal{H} = \{0\}$ by (26). Therefore, in this example, $\mathcal{U} = \mathcal{H}$. We have

$$\mathcal{K} = \mathcal{V}\oplus\{0\} = \operatorname{span}\{(\partial_x,0),(\partial_y,0)\} \subset TM\oplus T^*M$$

and

$$\mathcal{K}^\perp = TM\oplus\mathcal{V}^\circ = \operatorname{span}\left\{\begin{matrix}(\partial_x,0), & (\partial_y,0), & (\partial_\theta,0), & (\partial_\varphi,0),\\ (\partial_{p_\theta},0), & (\partial_{p_\varphi},0), & (0,\mathbf{d}p_\varphi), & \\ (0,\mathbf{d}\varphi), & (0,\mathbf{d}p_\theta), & (0,\mathbf{d}\theta) & \end{matrix}\right\}.$$

By (23) and the fact that $\mathcal{V}+\mathcal{H}=TM$, we know that for each spanning section X of $\mathcal{H}$, there exists exactly one $\alpha\in\Gamma(\mathcal{V}^\circ)$ such that the pair (X,α) is a section of $D\cap\mathcal{K}^\perp$. Using (27) and the equalities

$$\mathbf{i}_{\partial_\varphi}\omega_M-\frac{\mu R\sin\varphi}{I}p_\theta(\mathbf{d}x-R\cos\varphi\mathbf{d}\theta)+\frac{\mu R\cos\varphi}{I}(\mathbf{d}y-R\sin\varphi\mathbf{d}\theta)=\mathbf{d}p_\varphi, \tag{28}$$

$$\begin{aligned}\mathbf{i}_{\partial_{p_\theta}}\omega_M&+\frac{\mu R\cos\varphi}{I}p_\theta(\mathbf{d}x-R\cos\varphi\mathbf{d}\theta)\\&+\frac{\mu R\sin\varphi}{I}(\mathbf{d}y-R\sin\varphi\mathbf{d}\theta)=-\left(1+\mu R^2/I\right)\mathbf{d}\theta\end{aligned} \tag{29}$$

we find

$$D\cap\mathcal{K}^\perp=\operatorname{span}\left\{\begin{array}{c}\left(\partial_\varphi,\mathbf{d}p_\varphi\right),\left(\partial_{p_\theta},-\left(1+\mu R^2/I\right)\mathbf{d}\theta\right),\left(\partial_{p_\varphi},-\mathbf{d}\varphi\right),\\ \left(\partial_\theta+R\cos\varphi\partial_x+R\sin\varphi\partial_y,\left(1+\mu R^2/I\right)\mathbf{d}p_\theta\right)\end{array}\right\}.$$

Hence

$$(D\cap\mathcal{K}^\perp)+\mathcal{K}=\operatorname{span}\left\{\begin{array}{c}\left(\partial_\varphi,\mathbf{d}p_\varphi\right),\left(\partial_{p_\theta},-\left(1+\mu R^2/I\right)\mathbf{d}\theta\right),\left(\partial_{p_\varphi},-\mathbf{d}\varphi\right),\\ \left(\partial_\theta,\left(1+\mu R^2/I\right)\mathbf{d}p_\theta\right),(\partial_x,0),\left(\partial_y,0\right)\end{array}\right\}$$

and finally we get the reduced Dirac structure

$$\begin{aligned}D_{\mathrm{red}}&=\frac{(D\cap\mathcal{K}^\perp)+\mathcal{K}}{\mathcal{K}}\Big/G\\&=\operatorname{span}\left\{\begin{array}{l}\left(\partial_\varphi,\mathbf{d}p_\varphi\right),\left(\partial_{p_\theta},-(1+\mu R^2/I)\mathbf{d}\theta\right),\left(\partial_{p_\varphi},-\mathbf{d}\varphi\right),\\ \left(\partial_\theta,(1+\mu R^2/I)\mathbf{d}p_\theta\right)\end{array}\right\}\end{aligned} \tag{30}$$

on the four-dimensional manifold $\bar{M}=M/G$ with coordinates $(\varphi,\theta,p_\varphi,p_\theta)$. Thus, D_{red} is the graph of the symplectic form on $\bar{M}$ given by $\omega_{\mathrm{red}}=\mathbf{d}\varphi\wedge\mathbf{d}p_\varphi+(1+\mu R^2/I)\mathbf{d}\theta\wedge\mathbf{d}p_\theta$.

As already mentioned, in this example, $\mathcal{U}=\mathcal{H}$ and hence $\bar{\mathcal{H}}=T\pi(\mathcal{H})=\operatorname{span}\{\partial_\varphi,\partial_{p_\varphi},\partial_\theta,\partial_{p_\theta}\}$ by (26) which coincides with the projection on the first factor of the reduced Dirac structure (30). In this case $\bar{\mathcal{H}}=T\bar{M}$ and so $\omega_{\bar{\mathcal{H}}}=\omega_{\mathrm{red}}$ is of course nondegenerate.

2. *The case* $G=\mathrm{SE}(2)$ ([6]).

The Lie group $\mathrm{SE}(2):=\mathbb{S}^1 \circledS \mathbb{R}^2$ is the semidirect product of the circle $\mathbb{S}^1$ identified with matrices of the form

$$\begin{bmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{bmatrix}$$

and acting on $\mathbb{R}^2$ by usual matrix multiplication. Denote elements of SE(2) by

(α, r, s) where $\alpha \in \mathbb{S}^1$ and $r, s \in \mathbb{R}$. Define the action of the Lie group SE(2) on M by

$$(\alpha, r, s)\cdot(x, y, \theta, \varphi, p_\theta, p_\varphi) = (x\cos\alpha - y\sin\alpha + r, x\sin\alpha + y\cos\alpha + s, \theta, \varphi+\alpha, p_\theta, p_\varphi)$$

and note that the Hamiltonian H is invariant by this action. The distribution $\mathcal{V}$ on M is in this case $\mathcal{V} = \operatorname{span}\{\partial_x, \partial_y, \partial_\varphi\}$ and we get

$$\mathcal{K} = \mathcal{V} \oplus \{0\} = \operatorname{span}\{(\partial_x, 0), (\partial_y, 0), (\partial_\varphi, 0)\}.$$

Thus

$$\mathcal{K}^\perp = TM \oplus \mathcal{V}^\circ = \operatorname{span}\left\{(\partial_x, 0), (\partial_y, 0), (\partial_\theta, 0), (\partial_\varphi, 0), (\partial_{p_\theta}, 0), (\partial_{p_\varphi}, 0),\right.$$
$$\left.(0, \mathbf{d}p_\varphi), (0, \mathbf{d}p_\theta), (0, \mathbf{d}\theta)\right\}.$$

We have $\mathcal{V} \cap \mathcal{H} = \operatorname{span}\{\partial_\varphi\}$ (see (26)) and hence

$$(\mathcal{V} \cap \mathcal{H})^{\omega_M} = \ker\left(\mathbf{d}p_\varphi + \frac{\mu R\sin\varphi}{I} p_\theta \mathbf{d}x - \frac{\mu R\cos\varphi}{I} p_\theta \mathbf{d}y\right)$$

so that

$$\begin{aligned}\mathcal{U} &= \mathcal{H} \cap \ker\left(\mathbf{d}p_\varphi + \frac{\mu R\sin\varphi}{I} p_\theta \mathbf{d}x - \frac{\mu R\cos\varphi}{I} p_\theta \mathbf{d}y\right)\\ &= \operatorname{span}\{\partial_\varphi, \partial_\theta + R\cos\varphi\, \partial_x + R\sin\varphi\, \partial_y, \partial_{p_\theta}\}.\end{aligned}$$

Using (27), (28), and (29), we get

$$D \cap \mathcal{K}^\perp = \operatorname{span}\left\{\begin{array}{c}\left(\partial_\varphi, \mathbf{d}p_\varphi\right), \left(\partial_{p_\theta}, -\left(1 + \mu R^2/I\right)\mathbf{d}\theta\right)\\ \left(\partial_\theta + R\cos\varphi\partial_x + R\sin\varphi\partial_y, \left(1 + \mu R^2/I\right)\mathbf{d}p_\theta\right)\end{array}\right\}.$$

Thus,

$$D_{\text{red}} = \frac{(D \cap \mathcal{K}^\perp) + \mathcal{K}}{\mathcal{K}} \Big/ G = \operatorname{span}\left\{\begin{array}{l}\left(0, \mathbf{d}p_\varphi\right), \left(\partial_{p_\theta}, -(1 + \mu R^2/I)\mathbf{d}\theta\right),\\ \left(\partial_\theta, (1 + \mu R^2/I)\mathbf{d}p_\theta\right)\end{array}\right\}$$

is the graph of the Poisson tensor

$$\frac{I}{\mu R^2 + I}\partial_{p_\theta} \wedge \partial_\theta$$

defined on the manifold $\bar{M} := M/G$ with coordinates $(\theta, p_\theta, p_\varphi)$.

In addition,

$$\bar{\mathcal{H}} = T\pi(\mathcal{U}) = \operatorname{span}\{\partial_\theta, \partial_{p_\theta}\}$$

is an integrable subbundle of $T\bar{M}$ (since $[\partial_\theta, \partial_{p_\theta}] = 0$). Note that the projection on the first factor of D_{red} equals $\bar{\mathcal{H}}$. Finally, the 2-form $\omega_{\bar{\mathcal{H}}}$ is easily computed to be

$$\omega_{\bar{\mathcal{H}}}\left(\partial_\theta, \partial_{p_\theta}\right) = 1 + \frac{\mu R^2}{I}$$

and, as predicted by the general theory, it is nondegenerate on $\bar{\mathcal{H}}$.

3. *The case* $G = \mathbb{S}^1 \times \mathbb{R}^2$ ([6]). The direct product Lie group $\mathbb{S}^1 \times \mathbb{R}^2$ acts on M by

$$(\alpha, r, s) \cdot (x, y, \theta, \varphi, p_\theta, p_\varphi) = (x + r, y + s, \theta + \alpha, \varphi, p_\theta, p_\varphi).$$

The distribution $\mathcal{V}$ on M is in this case $\mathcal{V} = \operatorname{span}\{\partial_x, \partial_y, \partial_\theta\}$,

$$\mathcal{K} = \mathcal{V} \oplus \{0\} = \operatorname{span}\{(\partial_x, 0), (\partial_y, 0), (\partial_\theta, 0)\},$$

and thus

$$\begin{aligned}\mathcal{K}^\perp &= TM \oplus \mathcal{V}^\circ \\ &= \operatorname{span}\left\{(\partial_x, 0), (\partial_y, 0), (\partial_\theta, 0), (\partial_\varphi, 0), (\partial_{p_\theta}, 0), (\partial_{p_\varphi}, 0), (0, \mathbf{d}p_\varphi), (0, \mathbf{d}p_\theta), (0, \mathbf{d}\varphi)\right\}.\end{aligned}$$

Using (26) we get $\mathcal{V} \cap \mathcal{H} = \operatorname{span}\{\partial_\theta + R\cos\varphi\partial_x + R\sin\varphi\partial_y\}$ and hence

$$(\mathcal{V} \cap \mathcal{H})^{\omega_M} = \ker\left\{\left(1 + \frac{\mu R^2}{I}\right)\mathbf{d}p_\theta\right\}.$$

Therefore, again by (26) we conclude

$$\begin{aligned}\mathcal{U} = \mathcal{H} \cap (\mathcal{V} \cap \mathcal{H})^{\omega_M} &= \mathcal{H} \cap \ker\left\{\left(1 + \frac{\mu R^2}{I}\right)\mathbf{d}p_\theta\right\} \\ &= \operatorname{span}\{\partial_\varphi, \partial_\theta + R\cos\varphi\,\partial_x + R\sin\varphi\,\partial_y, \partial_{p_\varphi}\}.\end{aligned}$$

Using (27) and (28), we obtain

$$\begin{aligned}&D \cap \mathcal{K}^\perp \\ &\quad= \operatorname{span}\left\{\left(\partial_\varphi, \mathbf{d}p_\varphi\right), \left(\partial_{p_\varphi}, -\mathbf{d}\varphi\right), \left(\partial_\theta + R\cos\varphi\partial_x + R\sin\varphi\partial_y, \left(1 + \frac{\mu R^2}{I}\right)\mathbf{d}p_\theta\right)\right\}\end{aligned}$$

and hence

$$D_{\text{red}} = \frac{(D \cap \mathcal{K}^\perp) + \mathcal{K}}{\mathcal{K}} \Big/ G = \operatorname{span}\left\{\left(\partial_\varphi, \mathbf{d}p_\varphi\right), \left(\partial_{p_\varphi}, -\mathbf{d}\varphi\right), (0, \mathbf{d}p_\theta)\right\},$$

which is the graph of the Poisson tensor

$$\partial_{p_\varphi} \wedge \partial_\varphi$$

on the three-dimensional reduced manifold $\bar{M} = M/G$ with coordinates $(\varphi, p_\varphi, p_\theta)$. We have

$$\bar{\mathcal{H}} = T\pi(\mathcal{U}) = \operatorname{span}\{\partial_\varphi, \partial_{p_\varphi}\}$$

which is an integrable subbundle of $T\bar{M}$ (since $[\partial_\varphi, \partial_{p_\varphi}] = 0$). As before, the projection on the first factor of D_{red} equals $\bar{\mathcal{H}}$. The 2-form $\omega_{\bar{\mathcal{H}}}$ has the expression

$$\omega_{\bar{\mathcal{H}}}\left(\partial_\varphi, \partial_{p_\varphi}\right) = 1$$

and, as the general theory states, it is nondegenerate on $\bar{\mathcal{H}}$.

3.3.3. The Chaplygin skate

The standard Chaplygin skate. This example can be found in [26]. It describes the motion of a hatchet on a hatchet planimeter, that behaves like a curved knife edge. It is now commonly known under the name of "Chaplygin skate". Let the contact point of the knife edge have the coordinates $x, y \in \mathbb{R}^2$, let its direction relative to the positive x-axis be θ, and let its center of mass be at distance s from the contact point. Denote the total mass of the knife edge by m. Thus the moment of inertia about an axis through the contact point normal to the xy plane is $I = ms^2$. The configuration space of this problem is the semidirect product $Q := \mathrm{SE}(2) = \mathbb{S}^1 \circledS \mathbb{R}^2$ whose coordinates are denoted by $\mathbf{q} := (\theta, x, y)$. We have the following concrete constraints on the velocities:

$$\mathcal{D} := \ker(\sin\theta \mathbf{d}x - \cos\theta \mathbf{d}y) = \operatorname{span}\left\{\cos\theta\partial_x + \sin\theta\partial_y, \partial_\theta\right\} \subset TQ.$$

The Lagrangian is hyperregular and taken to be the kinetic energy of the knife edge, namely,

$$\begin{aligned} L(\theta, x, y, \dot{\theta}, \dot{x}, \dot{y}) &= \frac{1}{2}m(\dot{x} - s\dot{\theta}\sin\theta)^2 + \frac{1}{2}m(\dot{y} + s\dot{\theta}\cos\theta)^2 \\ &= \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}ms^2\dot{\theta}^2 + ms\dot{\theta}(\dot{y}\cos\theta - \dot{x}\sin\theta), \end{aligned}$$

where we have used that the x and y components of the velocity of the center of mass are, respectively,

$$\dot{x} - s\dot{\theta}\sin\theta \qquad \text{and} \qquad \dot{y} + s\dot{\theta}\cos\theta.$$

Compute

$$\begin{aligned} p_x &= \frac{\partial L}{\partial \dot{x}} = m\dot{x} - ms\dot{\theta}\sin\theta, \\ p_y &= \frac{\partial L}{\partial \dot{y}} = m\dot{y} + ms\dot{\theta}\cos\theta, \\ p_\theta &= \frac{\partial L}{\partial \dot{\theta}} = ms^2\dot{\theta} + ms(\dot{y}\cos\theta - \dot{x}\sin\theta). \end{aligned}$$

In $\mathcal{D}$ we have $\dot{y}\cos\theta - \dot{x}\sin\theta = 0$ and hence we get for $(\theta, x, y, p_\theta, p_x, p_y)$ in the constraint submanifold $M \subseteq T^*Q$:

$$\begin{aligned} p_\theta = ms^2\dot{\theta} \qquad \text{and} \qquad p_x\sin\theta &= m\dot{x}\sin\theta - ms\dot{\theta}\sin^2\theta \\ &= m\dot{y}\cos\theta - ms\dot{\theta}(1 - \cos^2\theta) \\ &= m\dot{y}\cos\theta + ms\dot{\theta}\cos^2\theta - ms\dot{\theta} = p_y\cos\theta - \frac{1}{s}p_\theta. \end{aligned}$$

Hence the constraint manifold M is five-dimensional and given by

$$M := \{(\theta, x, y, p_\theta, p_x, p_y) \mid p_\theta = sp_y\cos\theta - sp_x\sin\theta\} \subseteq T^*Q.$$

The global coordinates on M are thus (θ, x, y, p_x, p_y). The pull back ω_M of the canonical 2-form ω on T^*Q to M has hence the expression

$$\begin{aligned}\omega_M &= \mathbf{d}x \wedge \mathbf{d}p_x + \mathbf{d}y \wedge \mathbf{d}p_y + \mathbf{d}\theta \wedge \mathbf{d}(sp_y \cos\theta - sp_x \sin\theta)\\ &= \mathbf{d}x \wedge \mathbf{d}p_x + \mathbf{d}y \wedge \mathbf{d}p_y + s\cos\theta \mathbf{d}\theta \wedge \mathbf{d}p_y - s\sin\theta \mathbf{d}\theta \wedge \mathbf{d}p_x.\end{aligned}$$

The Dirac structure D modeling this problem is given by (22). Formula (21) gives the vector subbundle

$$\mathcal{H} := (T(\pi_{T^*Q}|_M))^{-1}(\mathcal{D}) = \operatorname{span}\{\cos\theta\partial_x + \sin\theta\partial_y, \partial_\theta, \partial_{p_x}, \partial_{p_y}\} \subset TM,$$

or equivalently

$$\mathcal{H}^\circ = \operatorname{span}\{\sin\theta \mathbf{d}x - \cos\theta \mathbf{d}y\}.$$

A computation yields

$$\begin{aligned}\mathbf{i}_{\cos\theta\partial_x + \sin\theta\partial_y}\omega_M &= \cos\theta \mathbf{d}p_x + \sin\theta \mathbf{d}p_y,\\ \mathbf{i}_{\partial_\theta}\omega_M &= s\cos\theta \mathbf{d}p_y - s\sin\theta \mathbf{d}p_x,\\ \mathbf{i}_{\partial_{p_y}}\omega_M &= -\mathbf{d}y - s\cos\theta \mathbf{d}\theta,\\ \mathbf{i}_{\partial_{p_x}}\omega_M &= -\mathbf{d}x + s\sin\theta \mathbf{d}\theta.\end{aligned}$$

Hence

$$\begin{aligned}\big\{ &\big(\cos\theta\partial_x + \sin\theta\partial_y, \cos\theta \mathbf{d}p_x + \sin\theta \mathbf{d}p_y\big)\,;\, \big(\partial_\theta, s\cos\theta \mathbf{d}p_y - s\sin\theta \mathbf{d}p_x\big)\,;\\ &\big(\partial_{p_y}, -\mathbf{d}y - s\cos\theta \mathbf{d}\theta\big)\,;\, \big(\partial_{p_x}, -\mathbf{d}x + s\sin\theta \mathbf{d}\theta\big)\,;\, (0, \sin\theta \mathbf{d}x - \cos\theta \mathbf{d}y)\big\}\end{aligned}$$

is a smooth global basis for D.

We consider the action of the Lie group $G = \mathrm{SE}(2)$ on Q, given by

$$\begin{aligned}&\phi : G \times Q \to Q,\\ &\phi((\alpha, r, s), (\theta, x, y)) = (\theta + \alpha, \cos\alpha x - \sin\alpha y + r, \sin\alpha x + \cos\alpha y + s).\end{aligned}$$

Thus, the induced action on $\Phi : G \times T^*Q \to T^*Q$ is given by

$$\begin{aligned}\Phi((\alpha, r, s), (\theta, x, y, p_\theta, p_x, p_y)) = (&\theta + \alpha, \cos\alpha x - \sin\alpha y + r, \sin\alpha x + \cos\alpha y + s,\\ &p_\theta, \cos\alpha p_x - \sin\alpha p_y, \sin\alpha p_x + \cos\alpha p_y).\end{aligned}$$

The action on Q obviously leaves the Lagrangian invariant. We show that the induced action on T^*Q leaves the manifold M invariant: we denote with $\theta', x', y', p'_x, p'_y, p'_\theta$ the coordinates of $\Phi((\alpha, r, s), (\theta, x, y, p_\theta, p_x, p_y))$ and compute

$$\begin{aligned}&s\cos\theta' p'_y - s\sin\theta' p'_x\\ &\quad = s\cos(\theta + \alpha)(\sin\alpha p_x + \cos\alpha p_y) - s\sin(\theta + \alpha)(\cos\alpha p_x - \sin\alpha p_y)\\ &\quad = s(\cos\theta\cos\alpha - \sin\theta\sin\alpha)(\sin\alpha p_x + \cos\alpha p_y)\\ &\qquad - s(\sin\theta\cos\alpha + \cos\theta\sin\alpha)(\cos\alpha p_x - \sin\alpha p_y)\\ &\quad = s\cos\theta p_y - s\sin\theta p_x = p_\theta = p'_\theta.\end{aligned}$$

Since the vertical bundle in this example is $\mathcal{V} = \operatorname{span}\{\partial_\theta, \partial_x, \partial_y\}$, we have $\mathcal{V}\cap\mathcal{H} = \operatorname{span}\{\partial_\theta, \cos\theta\,\partial_x + \sin\theta\,\partial_y\}$ and $(\mathcal{V}\cap\mathcal{H})^{\omega_M} = \ker\{\cos\theta\,\mathbf{d}p_x + \sin\theta\,\mathbf{d}p_y, s\cos\theta\,\mathbf{d}p_y - s\sin\theta\,\mathbf{d}p_x\} = \ker\{\mathbf{d}p_x, \mathbf{d}p_y\}$. Hence the distribution $\mathcal{U} = (\mathcal{V}\cap\mathcal{H})^{\omega_M}\cap\mathcal{H}$ is given by $\operatorname{span}\{\partial_\theta, \cos\theta\,\partial_x + \sin\theta\,\partial_y\}$ and

$$D\cap\mathcal{K}^\perp = \operatorname{span}\left\{\left(\cos\theta\,\partial_x + \sin\theta\,\partial_y, \cos\theta\,\mathbf{d}p_x + \sin\theta\,\mathbf{d}p_y\right), \left(\partial_\theta, s\cos\theta\,\mathbf{d}p_y - s\sin\theta\,\mathbf{d}p_x\right)\right\}.$$

We get the reduced Dirac structure

$$\begin{aligned}
D_{\mathrm{red}} &= \frac{(D\cap\mathcal{K}^\perp)+\mathcal{K}}{\mathcal{K}}\Big/ G \\
&= \operatorname{span}\left\{\left(0, \cos\theta\,\mathbf{d}p_x + \sin\theta\,\mathbf{d}p_y\right), \left(0, s\cos\theta\,\mathbf{d}p_y - s\sin\theta\,\mathbf{d}p_x\right)\right\} \\
&= \operatorname{span}\left\{(0, \mathbf{d}p_x), \left(0, \mathbf{d}p_y\right)\right\}
\end{aligned}$$

on the two-dimensional manifold $\bar{M} := M/G$ with global coordinates (p_x, p_y). Note that this is the graph of the trivial Poisson tensor on $\bar{M}$.

The Chaplygin skate with a rotor on it. We propose here a variation of the previous example by considering the Chaplygin skate with a disk attached to the center of mass of the skate that is free to rotate about the vertical axis. Again, let the contact point of the knife edge have the coordinates $x, y \in \mathbb{R}^2$, let its direction relative to the positive x-axis be θ, and let its center of mass be at distance s from the contact point. Denote by m the mass of the knife edge. Thus its moment of inertia about an axis through the contact point normal to the xy plane is $I = ms^2$. Let ϕ be the angle between a fixed point on the disk and the positive x-axis and J be the moment of inertia of the disk about the vertical axis. The configuration space of this problem is $Q := \mathbb{S}^1\times\mathbb{S}^1\times\mathbb{R}^2$ whose points are denoted by $\mathbf{q} := (\phi, \theta, x, y)$. We have again the following concrete constraints on the velocities:

$$\mathcal{D} := \ker(\sin\theta\,\mathbf{d}x - \cos\theta\,\mathbf{d}y) = \operatorname{span}\left\{\cos\theta\,\partial_x + \sin\theta\,\partial_y, \partial_\theta\right\} \subset TQ.$$

The Lagrangian is the kinetic energy of the knife edge,

$$\begin{aligned}
L(\phi, \theta, x, y, \dot\phi, \dot\theta, \dot x, \dot y) &= \frac{1}{2}m(\dot x - s\dot\theta\sin\theta)^2 + \frac{1}{2}m(\dot y + s\dot\theta\cos\theta)^2 + \frac{1}{2}J(\dot\theta + \dot\phi)^2 \\
&= \frac{1}{2}m(\dot x^2 + \dot y^2) + \frac{1}{2}(I+J)\dot\theta^2 + ms\dot\theta(\dot y\cos\theta - \dot x\sin\theta) \\
&\quad + \frac{1}{2}J\dot\phi^2 + J\dot\phi\dot\theta.
\end{aligned}$$

Compute

$$\begin{aligned}
p_x &= m\dot x - ms\dot\theta\sin\theta, \\
p_y &= m\dot y + ms\dot\theta\cos\theta, \\
p_\theta &= (I+J)\dot\theta + ms(\dot y\cos\theta - \dot x\sin\theta) + J\dot\phi, \\
p_\phi &= J(\dot\phi + \dot\theta).
\end{aligned}$$

Again, if we have $\dot y\cos\theta - \dot x\sin\theta = 0$, we compute

$$p_\theta = (I+J)\dot\theta + J\dot\phi = I\dot\theta + J(\dot\theta+\dot\phi) = I\dot\theta + p_\phi = ms^2\dot\theta + p_\phi$$

and

$$\begin{aligned} p_x\sin\theta &= m\dot x\sin\theta - ms\dot\theta\sin^2\theta = m\dot y\cos\theta - ms\dot\theta(1-\cos^2\theta)\\ &= m\dot y\cos\theta + ms\dot\theta\cos^2\theta - ms\dot\theta = p_y\cos\theta + \frac{1}{s}(p_\phi - p_\theta). \end{aligned}$$

Hence the constraint manifold M is seven-dimensional and given by

$$M := \{(\phi,\theta,x,y,p_\phi,p_\theta,p_x,p_y) \mid p_\theta = sp_y\cos\theta - sp_x\sin\theta + p_\phi\} \subseteq T^*Q.$$

The global coordinates on M are thus $(\phi,\theta,x,y,p_\phi,p_x,p_y)$. The pull back ω_M of the canonical 2-form ω on T^*Q to M has hence the expression

$$\begin{aligned} \omega_M &= \mathbf{d}x\wedge\mathbf{d}p_x + \mathbf{d}y\wedge\mathbf{d}p_y + \mathbf{d}\theta\wedge\mathbf{d}(sp_y\cos\theta - sp_x\sin\theta + p_\phi) + \mathbf{d}\phi\wedge\mathbf{d}p_\phi\\ &= \mathbf{d}x\wedge\mathbf{d}p_x + \mathbf{d}y\wedge\mathbf{d}p_y + s\cos\theta\mathbf{d}\theta\wedge\mathbf{d}p_y - s\sin\theta\mathbf{d}\theta\wedge\mathbf{d}p_x + (\mathbf{d}\theta+\mathbf{d}\phi)\wedge\mathbf{d}p_\phi. \end{aligned}$$

The Dirac structure D modeling this problem is given by (22). Formula (21) gives the vector subbundle

$$\mathcal{H} := (T(\pi_{T^*Q}|_M))^{-1}(\mathcal{D}) = \operatorname{span}\{\partial_\phi, \partial_\theta, \cos\theta\partial_x + \sin\theta\partial_y, \partial_{p_\phi}, \partial_{p_x}, \partial_{p_y}\} \subset TM,$$

or equivalently

$$\mathcal{H}^\circ = \operatorname{span}\{\sin\theta\mathbf{d}x - \cos\theta\mathbf{d}y\}.$$

A computation yields

$$\begin{aligned} \mathbf{i}_{\partial_\phi}\omega_M &= \mathbf{d}p_\phi,\\ \mathbf{i}_{\partial_\theta}\omega_M &= s\cos\theta\mathbf{d}p_y - s\sin\theta\mathbf{d}p_x + \mathbf{d}p_\phi,\\ \mathbf{i}_{\cos\theta\partial_x+\sin\theta\partial_y}\omega_M &= \cos\theta\mathbf{d}p_x + \sin\theta\mathbf{d}p_y,\\ \mathbf{i}_{\partial_{p_\phi}}\omega_M &= -\mathbf{d}\theta - \mathbf{d}\phi,\\ \mathbf{i}_{\partial_{p_y}}\omega_M &= -\mathbf{d}y - s\cos\theta\mathbf{d}\theta,\\ \mathbf{i}_{\partial_{p_x}}\omega_M &= -\mathbf{d}x + s\sin\theta\mathbf{d}\theta. \end{aligned}$$

We get

$$\begin{aligned} D = \operatorname{span}\Big\{ &\big(\partial_\phi, \mathbf{d}p_\phi\big);\ \big(\partial_\theta, s\cos\theta\mathbf{d}p_y - s\sin\theta\mathbf{d}p_x + \mathbf{d}p_\phi\big);\\ &\big(\cos\theta\partial_x + \sin\theta\partial_y, \cos\theta\mathbf{d}p_x + \sin\theta\mathbf{d}p_y\big);\ (\partial_{p_\phi}, -\mathbf{d}\theta - \mathbf{d}\phi);\\ &\big(\partial_{p_y}, -\mathbf{d}y - s\cos\theta\mathbf{d}\theta\big);\ \big(\partial_{p_x}, -\mathbf{d}x + s\sin\theta\mathbf{d}\theta\big);\ (0, \sin\theta\mathbf{d}x - \cos\theta\mathbf{d}y)\Big\}. \end{aligned}$$

We consider the action of the Lie group $G = \mathbb{S}^1\times\mathrm{SE}(2)$ on Q, given by

$$\begin{aligned} &\phi: G\times Q \to Q,\\ &\phi((\beta,\alpha,r,s),(\phi,\theta,x,y)) = (\phi+\beta, \theta+\alpha, \cos\alpha x - \sin\alpha y + r, \sin\alpha x + \cos\alpha y + s). \end{aligned}$$

Thus, the induced action $\Phi : G \times T^*Q \to T^*Q$ on T^*Q is given by

$$\Phi((\beta,\alpha,r,s),(\phi,\theta,x,y,p_\phi,p_\theta,p_x,p_y)) = (\phi+\beta,\theta+\alpha,\cos\alpha x-\sin\alpha y+r,\\ \sin\alpha x+\cos\alpha y+s,p_\theta,\cos\alpha p_x-\sin\alpha p_y,\sin\alpha p_x+\cos\alpha p_y).$$

The Lagrangian is invariant under the lift to TQ of ϕ and it is easy to see, with the considerations in the previous example, that the induced action Φ on T^*Q leaves the manifold M invariant.

Since the vertical bundle in this example is $\mathcal{V} = \operatorname{span}\{\partial_\phi,\partial_\theta,\partial_x,\partial_y\}$, we have $\mathcal{V}\cap\mathcal{H} = \operatorname{span}\{\partial_\phi,\partial_\theta,\cos\theta\partial_x+\sin\theta\partial_y\}$ and $(\mathcal{V}\cap\mathcal{H})^{\omega_M} = \ker\{\mathbf{d}p_\phi,\cos\theta\mathbf{d}p_x+\sin\theta\mathbf{d}p_y, s\cos\theta\mathbf{d}p_y - s\sin\theta\mathbf{d}p_x+\mathbf{d}p_\phi\} = \ker\{\mathbf{d}p_\phi,\mathbf{d}p_x,\mathbf{d}p_y\}$. Hence the distribution $\mathcal{U} = (\mathcal{V}\cap\mathcal{H})^{\omega_M}\cap\mathcal{H}$ is given by $\operatorname{span}\{\partial_\phi,\partial_\theta,\cos\theta\partial_x+\sin\theta\partial_y\}$ and

$$D\cap\mathcal{K}^\perp = \operatorname{span}\big\{\left(\partial_\phi,\mathbf{d}p_\phi\right),\left(\cos\theta\partial_x+\sin\theta\partial_y,\cos\theta\mathbf{d}p_x+\sin\theta\mathbf{d}p_y\right),\\ \left(\partial_\theta, s\cos\theta\mathbf{d}p_y - s\sin\theta\mathbf{d}p_x+\mathbf{d}p_\phi\right)\big\}.$$

We get the reduced Dirac structure

$$\begin{aligned} D_{\text{red}} &= \frac{(D\cap\mathcal{K}^\perp)+\mathcal{K}}{\mathcal{K}}\Big/ G \\ &= \operatorname{span}\left\{\left(0,\mathbf{d}p_\phi\right),\left(0,\cos\theta\mathbf{d}p_x+\sin\theta\mathbf{d}p_y\right),\left(0,s\cos\theta\mathbf{d}p_y-s\sin\theta\mathbf{d}p_x+\mathbf{d}p_\phi\right)\right\} \\ &= \operatorname{span}\left\{\left(0,\mathbf{d}p_\phi\right),\left(0,\mathbf{d}p_x\right),\left(0,\mathbf{d}p_y\right)\right\} \end{aligned}$$

on the three-dimensional manifold $\bar{M} := M/G$ with global coordinates (p_ϕ,p_x,p_y). This is again the graph of the trivial Poisson tensor on $\bar{M}$.

In these six examples we get integrable Dirac structures after reduction. We shall come back to this remark in the last section of the paper.

4. Optimal reduction for nonholonomic systems

Recall the setting of Section 3: Q is a configuration space which is a smooth Riemannian manifold, $\mathcal{D}\subseteq TQ$ is the constraints distribution given as the intersection of the kernels of k linearly independent 1-forms on Q (and $\mathcal{D}$ is hence a vector subbundle of TQ), L is a hyperregular classical Lagrangian equal to the kinetic energy of the given Riemannian metric on Q minus a potential, $M := \mathbb{F}L(\mathcal{D}) \subset T^*Q$ is a submanifold and represents the constraints in phase space T^*Q, and $\omega_M := i^*\omega_{\text{can}} \in \Omega^2(M)$ is the induced 2-form on M, where $i : M \hookrightarrow T^*Q$ is the inclusion and ω_{can} the canonical symplectic form on T^*Q. The vector bundle $\mathcal{H} := TM\cap(T\pi_{T^*Q})^{-1}(\mathcal{D})$ is not integrable but has the property that the restriction $\omega_{\mathcal{H}}$ of ω_M to $\mathcal{H}\times\mathcal{H}$ is nondegenerate. The Dirac structure D associated to this nonholonomic system has fibers

$$D(m) = \{(X(m),\alpha_m)\in T_mM\oplus T_m^*M \mid X\in\Gamma(\mathcal{H}),\ \alpha-\mathbf{i}_X\omega_M\in\Gamma(\mathcal{H}^\circ)\}$$

for all $m \in M$ and is, in general, not integrable. Recall from Proposition 2(i) that $\mathsf{G}_0 = \{0\}$, $\mathsf{P}_1 = T^*M$ and hence *all* functions are admissible.

Consider a G-action $\phi : G \times Q \to Q$ on Q that leaves the constraints and the Lagrangian invariant. The lift $\Phi : G \times T^*Q \to T^*Q$ of the action is defined by $\Phi_g = (T\phi_{g^{-1}})^*$; this is a symplectic action on T^*Q that leaves M invariant. Thus we get a canonical G-action on the Dirac manifold (M, D) and we have for all $g \in G$,

$$\Phi_g^*\omega_M = \Phi_g^*\left(i^*\omega_{\text{can}}\right) = i^*\left(\Phi_g^*\omega_{\text{can}}\right) = i^*\omega_{\text{can}} = \omega_M$$

since the G-action commutes with the inclusion. Note that in this section the G-action on T^*Q is a lift, whereas in Section 3 we needed only that it is a symplectic action.

In this section we shall define a distribution on M that yields the equations of motion and the conserved quantities given by the nonholonomic Noether theorem (see [14], Theorem 2 and also [6], Chapter 5 and the corresponding internet supplement). If this distribution is integrable, we will prove a Marsden–Weinstein reduction theorem that gives a reduced Dirac structure which is the graph of a nondegenerate 2-form (not necessarily closed). This reduction procedure is done from an "optimal" point of view as in [17].

4.1. The nonholonomic Noether theorem

We recall in this subsection the Hamiltonian formulation of the nonholonomic Noether theorem. Let $\mathbf{J} : T^*Q \to \mathfrak{g}^*$ be the canonical momentum map associated to the action of G on T^*Q (see e.g. [23])

$$\mathbf{J}(p)(\xi) = \langle p\,, \xi_Q(\pi(p))\rangle \tag{31}$$

for all $p \in T^*Q$, where $\pi : T^*Q \to Q$ is the projection. For all $\xi \in \mathfrak{g}$, the ξ-component of $\mathbf{J}$ is the map $\mathbf{J}^\xi : T^*Q \to \mathbb{R}$ defined by

$$\mathbf{J}^\xi(p) := \mathbf{J}(p)(\xi) \tag{32}$$

for all $p \in T^*Q$. We shall denote by the same symbol $\mathbf{J}^\xi$ its restriction to the manifold M. For an arbitrary $\xi \in \mathfrak{g}$ we have therefore

$$\mathbf{i}_{\xi_{T^*Q}}\omega_{\text{can}} = \mathbf{d}\mathbf{J}^\xi. \tag{33}$$

Since the action of G on T^*Q leaves the submanifold M invariant, we have $\xi_{T^*Q}(m) \in T_mM$ for all $m \in M$ and hence the fundamental vector field ξ_{T^*Q} is i-related to ξ_M, i.e. $Ti \circ \xi_M = \xi_{T^*Q} \circ i$. Choosing for each vector field $X \in \mathfrak{X}(M)$ an arbitrary extension $X' \in \mathfrak{X}(T^*Q)$ (and hence $X \sim_i X'$) we get for all $m \in M$,

$$\begin{aligned}\mathbf{i}_{\xi_M}\omega_M(X)(m) &= \mathbf{i}_{\xi_M}(i^*\omega_{\text{can}})(X)(m) = \mathbf{i}_{\xi_{T^*Q}}\omega_{\text{can}}(X')(i(m)) = (\mathbf{d}\mathbf{J}^\xi(X'))(i(m))\\ &= (i^*\mathbf{d}\mathbf{J}^\xi)(X)(m) = (\mathbf{d}\mathbf{J}^\xi(X))(m)\end{aligned}$$

which shows that (33) naturally restricts to M

$$\mathbf{i}_{\xi_M}\omega_M = \mathbf{d}\mathbf{J}^\xi. \tag{34}$$

Define for all $p \in M$ the vector subspace $\mathfrak{g}^p := \{\xi \in \mathfrak{g} \mid \xi_M(p) \in (\mathcal{V} \cap \mathcal{H})(p)\} \subseteq \mathfrak{g}$. Then

$$\mathfrak{g}^{\mathcal{H}} := \bigcup_{p \in M} \mathfrak{g}^p$$

is a smooth (not necessarily trivial) vector subbundle of the trivial bundle $M \times \mathfrak{g}$ if and only if $\mathcal{H} + \mathcal{V}$ has constant rank on M, for instance if $\mathcal{H} + \mathcal{V} = TM$. Indeed, note first that $\mathfrak{g}^{\mathcal{H}} = \Lambda^{-1}(\mathcal{V} \cap \mathcal{H})$, where $\Lambda : M \times \mathfrak{g} \to \mathcal{V}$ is the vector bundle isomorphism over M given by $\Lambda(m, \xi) := \xi_M(m)$. However, since $\mathcal{H} + \mathcal{V}$, $\mathcal{H}$, and $\mathcal{V}$ are subbundles of TM, it follows that $\mathcal{V} \cap \mathcal{H}$ is also a subbundle of both TM and $\mathcal{V}$. Consequently, $\mathfrak{g}^{\mathcal{H}} = \Lambda^{-1}(\mathcal{V} \cap \mathcal{H})$ is a subbundle of the trivial vector bundle $M \times \mathfrak{g}$. If $\mathfrak{g}^{\mathcal{H}}$ is a vector bundle over M, then $\mathcal{V} \cap \mathcal{H}$ is also a vector bundle and hence its fibers have constant dimension on M. It follows immediately that the rank of $\mathcal{H} + \mathcal{V}$ is also constant on M.

For the rest of this subsection we assume that $\mathcal{H} + \mathcal{V}$ has constant rank on M and hence that $\mathfrak{g}^{\mathcal{H}}$ is a vector subbundle of the trivial vector bundle $M \times \mathfrak{g}$. If $\xi^{\mathcal{H}}$ is a smooth section of $\mathfrak{g}^{\mathcal{H}}$, then $\boldsymbol{\xi}(p) := (\xi^{\mathcal{H}}(p))_M(p)$ defines a smooth section of $\mathcal{V} \cap \mathcal{H}$. Conversely, if $\{\xi^1, \ldots, \xi^k\}$ is a chosen basis for the Lie algebra $\mathfrak{g}$, then the vector fields $\xi^1_M, \ldots, \xi^k_M$ are global vector fields on M that don't vanish and are everywhere linearly independent. Hence, $\xi^1_M, \ldots, \xi^k_M$ are smooth basis vector fields for the bundle $\mathcal{V}$. Every section $\boldsymbol{\xi}$ of $\mathcal{V} \cap \mathcal{H}$ can hence be written $\boldsymbol{\xi} = \sum_{i=1}^k f_i \xi^i_M$ with smooth (local) functions $f_1, \ldots, f_k$, and corresponds exactly to the section $\xi^{\mathcal{H}} = \sum_{i=1}^k f_i \xi^i$ of $\mathfrak{g}^{\mathcal{H}}$.

Since $\mathcal{V} \cap \mathcal{H}$ is a subbundle of TM with

$$[\Gamma(\mathcal{V} \cap \mathcal{H}), \Gamma(\mathcal{V})] \subseteq \Gamma(\mathcal{V}) = \Gamma((\mathcal{V} \cap \mathcal{H}) + \mathcal{V}),$$

we have that for each $p \in M$ there exists a neighbourhood U of p and spanning sections of $\mathcal{V} \cap \mathcal{H}$ on U that descend to the quotient M/G (see [16]).

Let $\xi^{\mathcal{H}}$ be a smooth section of $\mathfrak{g}^{\mathcal{H}}$. For all $p \in M$ and all $X \in \mathfrak{X}(M)$ the definition of the corresponding $\boldsymbol{\xi}$ and (34) yield

$$\omega_M(p)\,(\boldsymbol{\xi}(p), X(p)) = \mathbf{dJ}^{\xi^{\mathcal{H}}(p)}(p)\,(X)\,. \tag{35}$$

As above, write $\xi^{\mathcal{H}} = \sum_{i=1}^k f_i \xi^i$ with smooth functions $f_1, \ldots, f_k$ and the chosen basis $\{\xi^1, \ldots, \xi^k\}$ of $\mathfrak{g}$. Define the smooth map

$$\begin{aligned} \mathbf{J}^{\xi^{\mathcal{H}}} &: M \to \mathbb{R} \\ p &\mapsto \mathbf{J}^{\xi^{\mathcal{H}}(p)}(p) = \langle \mathbf{J}(p), \xi^{\mathcal{H}}(p) \rangle. \end{aligned}$$

Using (31) and (32) we get

$$\mathbf{J}^{\xi^{\mathcal{H}}}(p) = \mathbf{J}^{\sum_{i=1}^k f_i(p)\xi^i}(p) = \sum_{i=1}^k f_i(p) \mathbf{J}^{\xi^i}(p).$$

If $c : (-\varepsilon; \varepsilon) \to M$ is a solution curve of a vector field $X \in \mathfrak{X}(M)$ with $c(0) = p$, we have

$$\begin{aligned}\mathbf{dJ}^{\xi^{\mathcal{H}}}(p)(X) &= \left.\frac{d}{dt}\right|_{t=0} \mathbf{J}^{\xi^{\mathcal{H}}}(c(t)) = \left.\frac{d}{dt}\right|_{t=0} \sum_{i=1}^{k} f_i(c(t))\mathbf{J}^{\xi^i}(c(t)) \\ &= \sum_{i=1}^{k} \mathbf{d}f_i(p)(\dot{c}(0))\mathbf{J}^{\xi^i}(p) + \sum_{i=1}^{k} f_i(p)\mathbf{dJ}^{\xi^i}(p)(\dot{c}(0)) \\ &= \mathbf{J}^{\sum_{i=1}^{k}\mathbf{d}f_i(p)(X)\xi^i}(p) + \mathbf{dJ}^{\xi^{\mathcal{H}}(p)}(p)(X) = \mathbf{J}^{X\left[\xi^{\mathcal{H}}\right]}(p) + \mathbf{dJ}^{\xi^{\mathcal{H}}(p)}(p)(X),\end{aligned}$$

where we write $X\left[\xi^{\mathcal{H}}\right] := \sum_{i=1}^{k} X[f_i]\xi^i$. Thus (35) becomes for all $p \in M$ and all $X \in \mathfrak{X}(M)$,

$$\omega_M(p)\left(\boldsymbol{\xi}(p), X(p)\right) = \mathbf{dJ}^{\xi^{\mathcal{H}}}(p)(X) - \mathbf{J}^{X\left[\xi^{\mathcal{H}}\right]}(p).$$

Hence, if the one-form $\alpha^{\boldsymbol{\xi}} \in \Omega^1(M)$ is defined by

$$\alpha^{\boldsymbol{\xi}}(X) := \mathbf{dJ}^{\xi^{\mathcal{H}}}(X) - \mathbf{J}^{X\left[\xi^{\mathcal{H}}\right]}$$

for all $X \in \mathfrak{X}(M)$, we have $\mathbf{i}_{\boldsymbol{\xi}}\omega_M = \alpha^{\boldsymbol{\xi}}$ and so the pair $(\boldsymbol{\xi}, \alpha^{\boldsymbol{\xi}})$ is a section of D.

Let h be a G-invariant Hamiltonian and $X_h \in \Gamma(\mathcal{H})$ the solution of the implicit Hamiltonian system $(X, \mathbf{d}h) \in \Gamma(D)$. Then

$$\mathbf{dJ}^{\xi^{\mathcal{H}}}(X_h) - \mathbf{J}^{X_h\left[\xi^{\mathcal{H}}\right]} = \alpha^{\boldsymbol{\xi}}(X_h) = \omega_M(\boldsymbol{\xi}, X_h) = -\mathbf{d}h(\boldsymbol{\xi}) = 0$$

since $\mathbf{d}h(\boldsymbol{\xi})(p) = \langle \mathbf{d}h(p), \boldsymbol{\xi}(p)\rangle = \left\langle \mathbf{d}h(p), \left(\xi^{\mathcal{H}}(p)\right)_M(p)\right\rangle = 0$ by G-invariance of h. Thus, we have proved the following result.

THEOREM 2 (nonholonomic Noether theorem). *Let $\xi^{\mathcal{H}}$ be a section of $\mathfrak{g}^{\mathcal{H}}$ and $X_h \in \Gamma(\mathcal{H})$ the solution of the implicit Hamiltonian system $(X, \mathbf{d}h) \in \Gamma(D)$, where h is a G-invariant Hamiltonian. Then X_h satisfies the nonholonomic Noether momentum equation:*

$$\mathbf{dJ}^{\xi^{\mathcal{H}}}(X_h) - \mathbf{J}^{X_h\left[\xi^{\mathcal{H}}\right]} = 0. \tag{36}$$

Recall from (16) and (18) that $\mathcal{H}$ is defined in terms of the given Lagrangian $L : TQ \to \mathbb{R}$ and hence, only the dynamics defined by the corresponding Hamiltonian H is of interest. Thus, for each other Lagrangian L' we obtain another distribution $\mathcal{H}'$.

REMARK 3. In [6], Theorem 5.5.4, the nonholonomic Noether theorem is formulated in terms of a Lagrangian of a classical mechanical systems (hence equal to the kinetic energy of a metric minus a potential). Let $\mathcal{V}_Q \subseteq TQ$ be the vertical subbundle of the action $\phi : G \times Q \to Q$. Under the *dimension assumption* $\mathcal{D} + \mathcal{V}_Q = TQ$, the distribution $\mathcal{D} \cap \mathcal{V}_Q$ is a smooth subbundle of TQ. Note that this assumption leads automatically to $\mathcal{F} + \mathcal{V}_{T^*Q} = TT^*Q$ and hence to $\mathcal{H} + \mathcal{V} = TM$ (see (17) for the definition of $\mathcal{F}$).

The smooth vector bundle $\mathfrak{g}^{\mathcal{D}} := \bigcup_{p \in M} \mathfrak{g}^{\mathcal{D}}(q)$ is defined pointwise by

$$\mathfrak{g}^{\mathcal{D}}(q) := \{\xi \in \mathfrak{g} \mid \xi_Q(q) \in (\mathcal{V}_Q \cap \mathcal{D})(q)\} \subseteq \mathfrak{g}.$$

Let $(\mathfrak{g}^{\mathcal{D}})^*$ be the dual bundle, that is, its fibers are $(\mathfrak{g}^{\mathcal{D}})^*(q) := (\mathfrak{g}^{\mathcal{D}}(q))^*$ for all $q \in Q$. The *nonholonomic momentum map* $J^{\rm nhc} : TQ \to (\mathfrak{g}^{\mathcal{D}})^*$ is the vector bundle map over Q defined by

$$\langle J^{\rm nhc}(v_q), \xi\rangle = \langle \mathbb{F}L(v_q), \xi_Q(q)\rangle = \frac{\partial L}{\partial \dot{q}^i}(\xi_Q)^i(q) =: J^{\rm nhc}(\xi)(v_q)$$

where $\xi \in \mathfrak{g}^{\mathcal{D}}(q)$. Let $\xi^{\mathcal{D}}$ be a section of the bundle $\mathfrak{g}^{\mathcal{D}}$. Theorem 5.5.4 in [6] states that any solution $c(t) = (q(t), \dot{q}(t))$ of the Lagrange–d'Alembert equations for a nonholonomic system must satisfy, in addition to the given kinematic constraints, the *momentum equation*

$$\frac{d}{dt}J^{\rm nhc}\left(\xi^{\mathcal{D}}(q(t))\right)(c(t)) = \frac{\partial L}{\partial \dot{q}^i}\left[\frac{d}{dt}(\xi^{\mathcal{D}}(q(t))\right]_Q^i. \qquad (37)$$

Since for all $\xi \in \mathfrak{g}$ the vector field ξ_{T^*Q} is the cotangent lift of ξ_Q, we have in local charts

$$\xi_{T^*Q} = \xi_Q^i\frac{\partial}{\partial q^i} - \frac{\partial \xi_Q^i}{\partial q^j}p_i\frac{\partial}{\partial p_j}.$$

Hence, if $\xi_Q(q) \in \mathcal{D}(q)$, we get $\xi_{T^*Q}(\alpha_q) \in \mathcal{F}(\alpha_q)$ for all $\alpha_q \in T_q^*Q$.

Consequently $\boldsymbol{\xi}(p) := \left(\xi^{\mathcal{D}}(\pi_{T^*Q}(p))\right)_M(p)$ for all $p \in M$ defines a smooth section $\boldsymbol{\xi}$ of $\mathcal{V}\cap\mathcal{H}$ and hence a smooth section $\xi^{\mathcal{H}}$ of $\mathfrak{g}^{\mathcal{H}}$. Note that $\xi^{\mathcal{H}}(p) = \xi^{\mathcal{D}}(\pi_{T^*Q}(p))$ for all $p \in M$ and if $\xi^{\mathcal{D}} = \sum_{i=1}^k f_i\xi_Q^i$ with smooth functions $f_1, \ldots, f_k$, then $\xi^{\mathcal{H}} = \sum_{i=1}^k F_i\xi_M^i$ with the smooth functions F_i defined by $F_i = i_M^*\pi_{T^*Q}^*f_i$, where $i_M : M \hookrightarrow T^*Q$ is the inclusion. Let X_H be a solution of the implicit Hamiltonian system $(X, \mathbf{d}H) \in \Gamma(D)$, where H is the G-invariant Hamiltonian on M associated to the Lagrangian L by the Legendre transformation, and $p(t)$ an integral curve of X_H. Then $c(t) := (q(t), \dot{q}(t)) = (\mathbb{F}L)^{-1}(p(t))$ is a solution of the Lagrange–d'Alembert equations. We have for all t

$$\begin{aligned}
0 &= \left(\mathbf{dJ}^{\xi^{\mathcal{H}}}(X_H) - \mathbf{J}^{X_H\left[\xi^{\mathcal{H}}\right]}\right)(p(t)) = \frac{d}{dt}\mathbf{J}^{\xi^{\mathcal{H}}}(p(t)) - \sum_{i=1}^k\frac{d}{dt}(F_i(p(t)))\mathbf{J}^{\xi_i}(p(t))\\
&= \frac{d}{dt}\left\langle p(t), \left(\xi^{\mathcal{H}}(p(t))\right)_Q(q(t))\right\rangle - \left\langle p(t), \left(\frac{d}{dt}\big(F_i(p(t))\big)\xi^i\right)_Q(q(t))\right\rangle\\
&= \frac{d}{dt}\left\langle \mathbb{F}L(c(t)), \left(\xi^{\mathcal{D}}(q(t))\right)_Q(q(t))\right\rangle - \left\langle \mathbb{F}L(c(t)), \left(\frac{d}{dt}\big(f_i(q(t))\big)\xi^i\right)_Q(q(t))\right\rangle\\
&= \frac{d}{dt}\left\langle \mathbb{F}L(c(t)), \left(\xi^{\mathcal{D}}(q(t))\right)_Q(q(t))\right\rangle - \left\langle \mathbb{F}L(c(t)), \left(\frac{d}{dt}\xi^{\mathcal{D}}(q(t))\right)_Q(q(t))\right\rangle\\
&= \frac{d}{dt}J^{\rm nhc}\left(\xi^{\mathcal{D}}(q(t))\right)(c(t)) - \frac{\partial L}{\partial \dot{q}^i}\left[\frac{d}{dt}(\xi^{\mathcal{D}}(q(t))\right]_Q^i.
\end{aligned}$$

Hence our nonholonomic Noether Theorem 2 is the Hamiltonian version of Theorem 5.5.4 in [6], that is, (36) and (37) are equivalent. △

PROPOSITION 3. *Assume that* $\mathcal{V}+\mathcal{H}$ *has constant rank on* M. *Let* $\xi^{\mathcal{H}}$ *be a* G*-equivariant section of* $\mathfrak{g}^{\mathcal{H}}$. *Then the corresponding section* $(\boldsymbol{\xi}, \alpha^{\boldsymbol{\xi}})$ *of* D *is also* G*-equivariant. There are two possibilities:*

(i) $\alpha^{\boldsymbol{\xi}} = \mathbf{i}_{\boldsymbol{\xi}}\omega_M = 0$ *on* $\mathcal{V}\cap\mathcal{H}$. *Then there exist* $\alpha' \in \Gamma(\mathcal{V}^\circ)$ *such that* $(\boldsymbol{\xi}, \alpha')$ *is a* G*-equivariant section of* $D\cap\mathcal{K}^\perp$ *and exactly one section* $\bar{\alpha} \in \Gamma(\mathsf{P}_0^{\mathrm{red}})$ *such that* $\pi^*\bar{\alpha} = \alpha'$. *Conversely, each section of* $\Gamma(\mathsf{P}_0^{\mathrm{red}})$ *pulls back to a section* α' *defined as above and satisfying this condition.*

(ii) $\mathcal{V}\cap\mathcal{H} \not\subseteq \boldsymbol{\xi}^{\omega_M}$ *and hence* $\alpha^{\boldsymbol{\xi}} \neq 0$ *on* $\mathcal{V}\cap\mathcal{H}$. *Then* $\alpha^{\boldsymbol{\xi}}$ *leads to a momentum equation that does not appear in the reduced implicit Hamiltonian system.*

Proof: If $\xi^{\mathcal{H}}$ is G-equivariant, we have $\xi^{\mathcal{H}}(g\cdot p) = \mathrm{Ad}_g\,\xi^{\mathcal{H}}(p)$ for all $p\in M$ and hence we get for the corresponding $\boldsymbol{\xi}$

$$
\begin{aligned}
(\Phi_g^*(\boldsymbol{\xi}))(p) &= T_{g\cdot p}\Phi_{g^{-1}}\boldsymbol{\xi}(g\cdot p) = T_{g\cdot p}\Phi_{g^{-1}}\left(\xi^{\mathcal{H}}(g\cdot p)\right)_M (g\cdot p)\\
&= T_{g\cdot p}\Phi_{g^{-1}}\left(\mathrm{Ad}_g(\xi^{\mathcal{H}}(p))\right)_M (g\cdot p)\\
&= \left(\mathrm{Ad}_{g^{-1}}\circ \mathrm{Ad}_g(\xi^{\mathcal{H}}(p))\right)_M (p) = \left(\xi^{\mathcal{H}}(p)\right)_M (p) = \boldsymbol{\xi}(p).
\end{aligned}
$$

Note that conversely, if $\boldsymbol{\xi}$ is equivariant, then the corresponding section $\xi^{\mathcal{H}}$ of $\mathfrak{g}^{\mathcal{H}}$ is G-equivariant. Since $\Phi_g^*\omega_M = \omega_M$ for all $g\in G$, the section $(\boldsymbol{\xi}, \alpha^{\boldsymbol{\xi}})$ is G-equivariant. Since $\mathcal{V}+\mathcal{H} = TM$, if $\alpha^{\boldsymbol{\xi}} = 0$ on $\mathcal{V}\cap\mathcal{H}$, there exists as in Proposition 2(ii) a section $\beta\in\Gamma(\mathcal{H}^\circ)$ such that $\alpha^{\boldsymbol{\xi}}+\beta \in \Gamma(\mathcal{V}^\circ)$ and hence $(\boldsymbol{\xi}, \alpha^{\boldsymbol{\xi}}+\beta)\in\Gamma(D\cap\mathcal{K}^\perp)$. Since $\boldsymbol{\xi}$, $\alpha^{\boldsymbol{\xi}}$, $\mathcal{H}^\circ$, $\mathcal{V}^\circ$ are all G-invariant and the action is free, we can take β G-invariant. Hence the first statement of (i) holds with $\alpha' := \alpha^{\boldsymbol{\xi}}+\beta$. But because $\boldsymbol{\xi}\in\Gamma(\mathcal{V})$, the section of D_{red} corresponding to $(\boldsymbol{\xi}, \alpha')$ will be $(0, \bar{\alpha})$ with $\bar{\alpha}\in\Omega^1(\bar{M})$ such that $\pi^*\bar{\alpha} = \alpha'$.

On the other hand, if we choose a nonzero section $\bar{\alpha}$ of $\mathsf{P}_0^{\mathrm{red}}$, the codistribution associated to the reduced Dirac structure on $\bar{M}$, we have $(0,\bar{\alpha})\in\Gamma(D_{\mathrm{red}})$ and we find $X\in\Gamma(\mathcal{H})$ such that $X\sim_\pi 0$ and $(X, \pi^*\bar{\alpha})\in\Gamma(D\cap\mathcal{K}^\perp)$. If $X=0$, then we have $\pi^*\bar{\alpha} = 0$ on $\mathcal{H}+\mathcal{V} = TM$, contradicting the fact that $\bar{\alpha}$ is a nonzero section of $\mathsf{P}_0^{\mathrm{red}}$. Therefore X is a nonzero vector field lying in $\Gamma(\mathcal{H}\cap\mathcal{V})$ with $\mathbf{i}_X\omega_M = 0$ on $\mathcal{H}\cap\mathcal{V}$. We conclude from this that the sections of $\mathsf{P}_0^{\mathrm{red}}$ pull back exactly to the G-equivariant sections $\alpha^{\boldsymbol{\xi}}+\beta\in\Gamma(\mathcal{V})^\circ$ induced by sections $\boldsymbol{\xi}$ of $(\mathcal{V}\cap\mathcal{H})\cap(\mathcal{V}\cap\mathcal{H})^{\omega_M}$.

If $\mathcal{V}\cap\mathcal{H}\not\subseteq\boldsymbol{\xi}^{\omega_M}$, i.e. $\boldsymbol{\xi}\notin\Gamma(\mathcal{U})$, and hence $\alpha^{\boldsymbol{\xi}}\neq 0$ on $\mathcal{V}\cap\mathcal{H}$, then there is no $\beta\in\Gamma(\mathcal{H}^\circ)$ such that $(\boldsymbol{\xi}, \alpha^{\boldsymbol{\xi}}+\beta)\in\Gamma(D\cap\mathcal{K}^\perp)$ and (ii) follows immediately (see also Proposition 2(ii)). □

DEFINITION 1. In the conditions of the preceding proposition, we will call *nonholonomic Noether equation* a section $\alpha^{\boldsymbol{\xi}}$ corresponding to a smooth section $\boldsymbol{\xi}$ of $\mathcal{V}\cap\mathcal{H}$. A $\mathcal{H}$*-modified nonholonomic Noether equation* is a 1-form $\alpha'\in\Omega^1(M)$

that can be written $\alpha' = \alpha^{\boldsymbol{\xi}} + \beta$ with a nonholonomic Noether equation $\alpha^{\boldsymbol{\xi}}$ and $\beta \in \Gamma(\mathcal{H}^\circ)$. A *descending ($\mathcal{H}$-modified) nonholonomic Noether equation* is a ($\mathcal{H}$-modified) nonholonomic Noether equation as in Proposition 3 (i).

Note that because of the β-part of a descending $\mathcal{H}$-modified nonholonomic Noether equation, sections of $\mathsf{P}_0^{\text{red}}$ don't pull back exactly to sections $\alpha^{\boldsymbol{\xi}}$ associated to sections $\xi^{\mathcal{H}}$ as in Theorem 2 (the nonholonomic Noether equations). It is possible that they pull back to one-forms that coincide only on $\mathcal{H}$ with some $\alpha^{\boldsymbol{\xi}}$.

PROPOSITION 4. *The codistribution spanned by the Noether equations which descend to the quotient M/G is given by*

$$\pi_2(D \cap (\mathcal{V} \oplus \mathcal{V}^\circ)) = (\flat(\mathcal{V} \cap \mathcal{H}) + \mathcal{H}^\circ) \cap \mathcal{V}^\circ \tag{38}$$

where $\flat : TM \to T^*M$ is associated to ω_M.

Proof: We have seen that a descending ($\mathcal{H}$-modified) nonholonomic Noether equation α' is a G-invariant section of $\mathcal{V}^\circ$ such that there exists a G-equivariant section X of $\mathcal{V} \cap \mathcal{H}$ with $(X, \alpha) \in \Gamma(D \cap \mathcal{K}^\perp)$. So we only have to show equality (38). Let α be a section of the left-hand side. Then there exists $X \in \Gamma(\mathcal{V} \cap \mathcal{H})$ such that $(X, \alpha) \in \Gamma(D \cap (\mathcal{V} \oplus \mathcal{V}^\circ))$ and hence there exists $\beta \in \Gamma(\mathcal{H}^\circ)$ such that $\alpha = \mathbf{i}_X\omega_M + \beta$. Therefore $\alpha \in \Gamma((\flat(\mathcal{V} \cap \mathcal{H}) + \mathcal{H}^\circ) \cap \mathcal{V}^\circ)$. Conversely, let α be a section of $(\flat(\mathcal{V} \cap \mathcal{H}) + \mathcal{H}^\circ) \cap \mathcal{V}^\circ$; then $\alpha = \mathbf{i}_X\omega_M + \beta \in \Gamma(\mathcal{V}^\circ)$ with $X \in \Gamma(\mathcal{V} \cap \mathcal{H})$ and $\beta \in \Gamma(\mathcal{H}^\circ)$. But this means that (X, α) is a section of $D \cap (\mathcal{V} \oplus \mathcal{V}^\circ)$. □

EXAMPLE 1. We compute $\boldsymbol{\xi}$ and $\alpha^{\boldsymbol{\xi}}$ for the constrained particle (see Subsection 3.3.1). In this example, $Q = \mathbb{R}^3$, $M := \{(x, y, z, p_x, p_y, p_z) \mid p_z = yp_x\} \subseteq T^*Q = \mathbb{R}^3 \times \mathbb{R}^3$, and $\mathcal{V} \cap \mathcal{H} = \text{span}\{\partial_x + y\partial_z\}$. If $\xi^1 := (1, 0)$ and $\xi^2 := (0, 1)$ is the chosen basis of $\mathfrak{g} = \mathbb{R}^2$, then $(1, 0)_M = \partial_x$ and $(0, 1)_M = \partial_z$ so that $\mathfrak{g}^{(x,y,z,p_x,p_y)} := \text{span}\{(1, 0) + y(0, 1)\}$ is the fiber of the vector bundle $\mathfrak{g}^{\mathcal{H}}$ at the point $(x, y, z, p_x, p_y) \in M$. Therefore, any section $\xi^{\mathcal{H}}$ of $\mathfrak{g}^{\mathcal{H}}$ has the form $\xi^{\mathcal{H}}(x, y, z, p_x, p_y) = f(x, y, z, p_x, p_y)\,((1, 0) + y(0, 1))$, where $f \in C^\infty(M)$. Consequently

$$\begin{aligned}\boldsymbol{\xi}(x, y, z, p_x, p_y) &= \left(\xi^{\mathcal{H}}(x, y, z, p_x, p_y)\right)_M (x, y, z, p_x, p_y) \\ &= f(x, y, z, p_x, p_y)\,(\partial_x + y\partial_z)\,.\end{aligned}$$

The components of the momentum map $\mathbf{J} : T^*Q \to \mathfrak{g}^*$ are $\mathbf{J}^{(1,0)} = p_x$ and $\mathbf{J}^{(0,1)} = p_z$ so that the restrictions to M of these functions are $\mathbf{J}^{(1,0)} = p_x$ and $\mathbf{J}^{(0,1)} = yp_x$. Therefore $\mathbf{J}^{\xi^{\mathcal{H}}}(x, y, z, p_x, p_y) = f(x, y, z, p_x, p_y)p_x(1 + y^2)$ and if $X \in \mathfrak{X}(M)$, then $X[f(1, 0) + yf(0, 1)] = X[f](1, 0) + X[yf](0, 1)$ and so

$$\mathbf{J}^{X[f(1,0)+yf(0,1)]} = X[f]p_x + X[yf]yp_x = p_x(1 + y^2)X[f] + yp_x f X[y].$$

The 1-form on M which applied to X yields the right-hand side is $p_x(1 + y^2)\mathbf{d}f + yp_x f\mathbf{d}y$ and hence

$$\begin{aligned}\alpha^{\boldsymbol{\xi}}(x, y, z, p_x, p_y) &= \mathbf{d}\mathbf{J}^{\xi^{\mathcal{H}}}(x, y, z, p_x, p_y) - p_x(1+y^2)\mathbf{d}f(x, y, z, p_x, p_y)\\ &\quad - yp_x f(x, y, z, p_x, p_y)\mathbf{d}y\\ &= f(x, y, z, p_x, p_y)\left((1+y^2)\mathbf{d}p_x + yp_x\mathbf{d}y\right).\end{aligned}$$

So the section spanning the codistribution $\mathsf{P}_0^{\mathrm{red}}$ in this example is $(1+y^2)\mathbf{d}p_x+yp_x\mathbf{d}y$, as (26). It is easy to see that in this case $\mathcal{V} \subseteq (\mathcal{V}\cap\mathcal{H})^{\omega_M}$ and hence the nonholonomic Noether equation descends to the quotient. $\diamondsuit$

4.2. The reaction-annihilator distribution

In this section also, we assume that $\mathcal{V}\cap\mathcal{H}$ has constant rank on M. An important problem is to decide when the nonholonomic Noether momentum equation gives a *constant of motion* rather than an *equation of motion*. We have to distinguish between two cases:

(i) The section $\xi^{\mathcal{H}}$ is constant, i.e, $\xi^{\mathcal{H}}(p) = \xi$ for all $p \in M$, where $\xi \in \mathfrak{g}$. Then $\boldsymbol{\xi}(p) = \xi_M(p)$ and so we have $\alpha^{\boldsymbol{\xi}} = \mathbf{d}\mathbf{J}^{\xi}$, so $\mathbf{J}^{\xi}$ is a constant of motion. We will see below that sometimes one can find $\eta \in \mathfrak{g}$ such that $\mathbf{J}^{\eta}$ is a constant of motion for all solutions of G-invariant Hamiltonians, but η_M is not a section of $\mathcal{V}\cap\mathcal{H}$ (see also [15]).

(ii) The other case is that of *gauge symmetries*, that is, nonconstant sections of $\mathfrak{g}^{\mathcal{H}}$ that yield constants of motion (see [15]). Note that if $\xi^{\mathcal{H}} = \sum_{i=1}^{k} f_i\xi_i$ then it leads to a constant of motion if one of the corresponding forms $\alpha^{\boldsymbol{\xi}}+\beta$ is exact, that is, we can find $f \in C^\infty(M)$ such that $\mathbf{d}f = \beta+\sum_{i=1}^{k} f_i\mathbf{d}J(\xi_i)$. However, we do not know of any other characterization of the section so that the momentum equation gives constants of motion rather than an equation of motion.

In the reduction method for nonholonomic systems, the first step is to compute the horizontal annihilator $\mathcal{U}$ of $\mathcal{V}$, that is, the subbundle $\mathcal{U} = (\mathcal{V}\cap\mathcal{H})^{\omega_M}\cap\mathcal{H} \subseteq TM \subseteq TT^*Q$ (see (19)). We have seen in Proposition 2(ii) that any section of $\mathcal{U}$ corresponds to a section of $D\cap\mathcal{K}^{\perp}$: for each $X \in \Gamma(\mathcal{U})$ there exists $\alpha \in \Gamma(\mathcal{V}^\circ)$ such that $(X, \alpha) \in \Gamma(D)$ and hence $\alpha - \mathbf{i}_X\omega_M \in \Gamma(\mathcal{H}^\circ)$. So the method of finding a section $\alpha \in \Gamma(\mathcal{V}^\circ)$ associated to $X \in \Gamma(\mathcal{U})$ is the same as determining $\beta \in \Gamma(\mathcal{H}^\circ)$ such that $\mathbf{i}_X\omega_M + \beta =: \alpha \in \Gamma(\mathcal{V}^\circ)$. As we have seen in Section 3.3.2, case 3, sometimes not the whole of $\mathcal{H}^\circ$ is needed in this construction. This is why we introduce the new codistribution $\mathcal{R}$ on M whose fiber at $p \in M$ equals

$$\mathcal{R}(p) = \{\beta(p) \mid \beta \in \Gamma(\mathcal{H}^\circ) \text{ and there is some } X \in \Gamma(\mathcal{U}) \text{ such that } \beta + \mathbf{i}_X\omega_M \in \Gamma(\mathcal{V}^\circ)\} \subseteq \mathcal{H}^\circ. \quad (39)$$

In general, $\mathcal{R}$ is strictly included in $\mathcal{H}^\circ$.

If $h \in C^\infty(M)^G$ is an admissible function, then there exists $X_h \in \Gamma(\mathcal{H})$ such that $(X_h, \mathbf{d}h) \in \Gamma(D)$. Recall that X_h is unique since $\mathsf{G}_0 = \{0\}$. In addition, since $\mathbf{d}h \in \Gamma(\mathcal{V}^\circ)$, we have $(X_h, \mathbf{d}h) \in \Gamma(D\cap K^{\perp})$ and hence $X_h \in \Gamma(\mathcal{U})$. Thus, there exists $\beta \in \Gamma(\mathcal{H}^\circ)$ such that $\mathbf{d}h = \mathbf{i}_{X_h}\omega_M + \beta$. This is exactly the Hamilton equation for the given nonholonomic system (with the Hamiltonian h) and we have $\beta \in \Gamma(\mathcal{R})$, often interpreted as the *reaction force*. In fact $\mathcal{R}^\circ \subset TM$ is the analogue of the *reaction-annihilator distribution* of [15].

PROPOSITION 5. *We have*

$$\flat(\mathcal{U}) \oplus \mathcal{R} = \mathcal{V}^\circ + \mathcal{R},$$

where $\flat : TM \to T^*M$ *corresponds to* ω_M.

Proof: The sum on the left-hand side is direct since if $X \neq 0 \in \Gamma(\mathcal{U})$, then $X \in \Gamma(\mathcal{H})$ and hence $\mathbf{i}_X\omega_M \notin \Gamma(\mathcal{H}^\circ)$ because $\omega_M|_{\mathcal{H}\times\mathcal{H}}$ is nondegenerate. Thus $\mathbf{i}_X\omega_M \notin \Gamma(\mathcal{R}) \subseteq \Gamma(\mathcal{H}^\circ)$. Second, recall that $\mathsf{P}_1 = T^*M$ (see Proposition 2(i)) so for all $\alpha \in \Gamma(\mathcal{V}^\circ)$ we find $X \in \Gamma(\mathcal{H})$ (actually $X \in \Gamma(\mathcal{U})$) such that $(X, \alpha) \in \Gamma(D)$. Thus, $\pi_2(D \cap \mathcal{K}^\perp) = \mathcal{V}^\circ$.

Now we are ready to prove the formula in the statement. If $X \in \Gamma(\mathcal{U})$, the considerations in Section 3.2 show that there exists $\beta \in \Gamma(\mathcal{H}^\circ)$ such that $\mathbf{i}_X\omega_M + \beta \in \Gamma(\mathcal{V}^\circ)$. The definition (39) of $\mathcal{R}$ yields directly that $\beta \in \Gamma(\mathcal{R})$. This shows $\flat(\mathcal{U}) \oplus \mathcal{R} \subseteq \mathcal{V}^\circ + \mathcal{R}$. For the other inclusion, choose $\alpha \in \Gamma(\mathcal{V}^\circ)$ and $X \in \Gamma(\mathcal{U})$ such that $(X, \alpha) \in \Gamma(D \cap \mathcal{K}^\perp)$. Then the definition of D yields $\beta := \alpha - \mathbf{i}_X\omega_M \in \Gamma(\mathcal{H}^\circ)$ and again, using (39), we conclude that $\beta \in \Gamma(\mathcal{R})$. □

The last lemma leads directly to the equality

$$\mathcal{U}^{\omega_M} \cap \mathcal{R}^\circ = \mathcal{V} \cap \mathcal{R}^\circ.$$

Note that $\mathcal{U}^{\omega_M} = ((\mathcal{H} \cap \mathcal{V})^{\omega_M} \cap \mathcal{H})^{\omega_M} = (\mathcal{H} \cap \mathcal{V}) + \mathcal{H}^{\omega_M}$ since the kernel of ω_M lies in $\mathcal{H}^{\omega_M}$.

Now we are able to state the main theorem of this subsection which is the Hamiltonian analogue of the main statement of [15].

THEOREM 3. *Let* $\xi \in \mathfrak{g}$. *Then the function* $\mathbf{J}^\xi$ *is a constant of motion for every* G*-invariant Hamiltonian* h *if and only if* $\xi_M \in \Gamma(\mathcal{V} \cap \mathcal{R}^\circ)$.

Proof: Choose $\xi \in \mathfrak{g}$ such that $\xi_M \in \Gamma(\mathcal{V} \cap \mathcal{R}^\circ)$. We have seen in the preceding section that $\mathbf{i}_{\xi_M}\omega_M = \mathbf{dJ}^\xi$. For an arbitrary $X \in \Gamma(\mathcal{U})$ choose $\beta \in \Gamma(\mathcal{R})$ with $\mathbf{i}_X\omega_M + \beta =: \alpha \in \Gamma(\mathcal{V}^\circ)$ and get

$$\mathbf{dJ}^\xi(X) = \omega_M(\xi_M, X) = \beta(\xi_M) - \alpha(\xi_M) = 0.$$

This yields the statement since for all G-invariant Hamiltonian h the (unique) solution X_h of the implicit Hamiltonian system $(X, \mathbf{d}h) \in \Gamma(D)$ is a section of $\mathcal{U}$ (with $\alpha = \mathbf{d}h$ the corresponding section of $\mathcal{V}^\circ$ and $\beta = \mathbf{d}h - \mathbf{i}_{X_h}\omega_M$). For the converse implication, choose $\xi \in \mathfrak{g}$ such that $\mathbf{J}^\xi$ is a constant of the motion for the solution curves of every G-invariant Hamiltonian. Note that since $\mathcal{V}$ is an involutive subbundle of TM, the exterior derivatives of all G-invariant functions span pointwise $\mathcal{V}^\circ$ and hence the corresponding solutions span $\mathcal{U}$. This yields $\mathbf{dJ}^\xi = 0$ on $\mathcal{U}$. If we choose $\beta \in \Gamma(\mathcal{R})$, there exists $X \in \Gamma(\mathcal{U})$ such that $\mathbf{i}_X\omega_M + \beta \in \Gamma(\mathcal{V}^\circ)$. Hence we get

$$0 = (\mathbf{i}_X\omega_M + \beta)(\xi_M) = \omega_M(X, \xi_M) + \beta(\xi_M) = -\mathbf{dJ}^\xi(X) + \beta(\xi_M) = 0 + \beta(\xi_M)$$

and therefore $\xi_M \in \Gamma(\mathcal{R}^\circ \cap \mathcal{V})$. □

COROLLARY 1. *Assume that $\mathcal{H}+\mathcal{V}$ has constant rank on M and choose $\xi \in \mathfrak{g}$. If $\mathbf{dJ}^\xi = 0$ on $\mathcal{V}\cap\mathcal{H}$ there exist $\beta \in \Gamma(\mathcal{H}^\circ)$ and $\boldsymbol{\eta} \in \Gamma(\mathcal{V}\cap\mathcal{H})$ such that $\alpha^{\boldsymbol{\eta}} = \mathbf{dJ}^\xi + \beta$.*

Proof: Since $\mathsf{P}_1 = T^*M$ (see Proposition 2(i)) there exists $X \in \Gamma(\mathcal{H})$ such that $(X, \mathbf{dJ}^\xi) \in \Gamma(D)$. Hence we have $\mathbf{dJ}^\xi = \mathbf{i}_X\omega_M + \beta'$ with $\beta' \in \Gamma(\mathcal{H}^\circ)$ and since $\mathbf{dJ}^\xi = 0$ on $\mathcal{U}$ we get $X \in \mathcal{U}^{\omega_M} = (\mathcal{H} \cap \mathcal{V}) + \mathcal{H}^{\omega_M}$. Write $X = V + Y$ with $V \in \Gamma(\mathcal{H}\cap\mathcal{V})$ and $Y \in \Gamma(\mathcal{H}^{\omega_M})$. Since X and V are sections of $\mathcal{H}$, then so is Y. But since $\mathcal{H} \cap \mathcal{H}^{\omega_M} = \{0\}$, this yields $Y = 0$ and hence $X \in \Gamma(\mathcal{H} \cap \mathcal{V})$. We find $\eta^{\mathcal{H}} \in \Gamma(\mathfrak{g}^{\mathcal{H}})$ such that the corresponding section $\boldsymbol{\eta} \in \Gamma(\mathcal{V} \cap \mathcal{H})$ is equal to X and therefore $(\eta, \mathbf{dJ}^\xi) \in \Gamma(D)$. We get $\alpha^{\boldsymbol{\eta}} = \mathbf{dJ}^\xi + \beta$ with $\beta \in \Gamma(\mathcal{H}^\circ)$, a nonholonomic Noether equation corresponding to the section $\eta^{\mathcal{H}} \in \Gamma(\mathfrak{g}^{\mathcal{H}})$. □

4.3. Optimal momentum map for nonholonomic mechanical systems

In this and the next subsection we assume that $\mathcal{H}+\mathcal{V}$ has constant rank on M. Recall from Remark 2 that this implies that $\mathcal{V}\cap\mathcal{H}$ and $\mathcal{U}$ also have constant rank on M.

We show in this subsection that, under certain integrability assumptions, it is possible to restrict the system to “level sets” given by the nonholonomic momentum equations and then perform reduction.

Consider the distribution where all $\alpha^{\boldsymbol{\xi}}, \alpha'$ as in Definition 1 vanish, namely

$$\begin{aligned}\mathcal{D}_G &:= \left[\pi_2(D \cap (\mathcal{V} \oplus \mathcal{V}^\circ))\right]^\circ \overset{(38)}{=} \left[(\flat(\mathcal{V} \cap \mathcal{H}) + \mathcal{H}^\circ) \cap \mathcal{V}^\circ\right]^\circ = ((\mathcal{V} \cap \mathcal{H})^{\omega_M} \cap \mathcal{H}) + \mathcal{V}\\ &= \mathcal{U} + \mathcal{V},\end{aligned}$$

where $\mathcal{U} := (\mathcal{V} \cap \mathcal{H})^{\omega_M} \cap \mathcal{H} \subseteq TM \subseteq TT^*Q$ is the horizontal annihilator of $\mathcal{V}$ (see (19)). Note that $\mathcal{D}_G = (\mathcal{V} \cap \mathcal{H})^{\omega_{\mathcal{H}}} + \mathcal{V}$. Since $\mathcal{U} \subseteq \mathcal{H}$ it follows easily that $\mathcal{D}_G \cap \mathcal{H} = \mathcal{U} + (\mathcal{V} \cap \mathcal{H})$. If $\mathcal{D}_G$ is integrable, its leaves are the level sets of the constants of motion and equations of motion given by the nonholonomic Noether Theorem 2 for sections $\boldsymbol{\xi}$ of $(\mathcal{V} \cap \mathcal{H}) \cap (\mathcal{V} \cap \mathcal{H})^{\omega_M}$: the fiber at $m \in M$ of the distribution $\pi_2(D \cap (\mathcal{V} \oplus \mathcal{V}^\circ))$ equals

$$\{\alpha(m) \mid \alpha \in \Gamma(\mathcal{V}^\circ) \text{ and there exists } X \in \Gamma(\mathcal{V}) \text{ such that } (X, \alpha) \in \Gamma(D)\}.$$

Note that if this distribution is spanned by closed 1-forms, hence locally exact 1-forms, then it can be written as

$$\{(\mathbf{d}f)(m) \mid f \in C^\infty(M)^G \text{ and there exists } X_f \in \Gamma(\mathcal{V}) \text{ such that } (X_f, \mathbf{d}f) \in \Gamma(D)\}.$$

For every $m \in M$, we have

$$\begin{aligned}\dim\left[(\mathcal{V}(m)\cap\mathcal{H}(m))^{\omega_{\mathcal{H}}}\cap(\mathcal{V}(m)\cap\mathcal{H}(m))\right] &= \dim\left[\mathcal{U}(m)\cap(\mathcal{V}(m)\cap\mathcal{H}(m))\right]\\ &= \dim\left[\mathcal{U}(m)\cap\mathcal{V}(m)\right]\\ &= \dim\mathcal{U}(m)+\dim\mathcal{V}(m)-\dim\mathcal{D}_G(m).\end{aligned}$$

Recall that $\mathcal{U}$ and $\mathcal{H}\cap\mathcal{V}$ are vector subbundles of TM by hypothesis. If, in addition, $\mathcal{D}_G$ is integrable, then its fibers $\mathcal{D}_G(m)$ have constant dimension along the leaves of the generalized foliation determined by $\mathcal{D}_G$ and so the computation above shows that the fibers of $(\mathcal{V}\cap\mathcal{H})^{\omega_{\mathcal{H}}}\cap(\mathcal{V}\cap\mathcal{H})$ along a leaf of $\mathcal{D}_G$ are constant. Thus, the same is true for the fibers of $\mathcal{D}_G\cap\mathcal{H}=\mathcal{U}+(\mathcal{V}\cap\mathcal{H})$ since $\mathcal{U}\cap\mathcal{V}\cap\mathcal{H}=(\mathcal{V}\cap\mathcal{H})^{\omega_{\mathcal{H}}}\cap(\mathcal{V}\cap\mathcal{H})$. We shall use this fact in the next subsection where we describe the induced Dirac structure on a leaf.

In order to restrict the system to the leaves of the distribution $\mathcal{D}_G$ and then perform reduction, we have to show several statements, the analogues of those needed for the Dirac optimal reduction. Since $\Phi_g^*\omega_M=\omega_M$ for all $g\in G$, the proof of the following proposition follows easily.

PROPOSITION 6. *The distribution* $(\mathcal{V}\cap\mathcal{H})^{\omega_M}$ *is* G*-invariant in the sense that*

$$\Phi_g^*\left((\mathcal{V}\cap\mathcal{H})^{\omega_M}\right)=(\mathcal{V}\cap\mathcal{H})^{\omega_M}$$

for all $g\in G$. *Since* $\mathcal{V}$ *and* $\mathcal{H}$ *are also* G*-invariant, it follows that the distribution* $\mathcal{D}_G=((\mathcal{V}\cap\mathcal{H})^{\omega_M}\cap\mathcal{H})+\mathcal{V}$ *is* G*-invariant.*

If $\mathcal{D}_G$ is integrable, define the *nonholonomic optimal momentum map*

$$\mathcal{J}: M\to M/\mathcal{D}_G.$$

LEMMA 2. *If* m *and* m' *are in the same leaf of* $\mathcal{D}_G$, *then* $\Phi_g(m)$ *and* $\Phi_g(m')$ *are in the same leaf of* $\mathcal{D}_G$ *for all* $g\in G$. *Hence there is a well-defined action of* G *on* $M/\mathcal{D}_G$:

$$\begin{aligned}\bar{\Phi}&: G\times M/\mathcal{D}_G\to M/\mathcal{D}_G,\\ \bar{\Phi}_g&(\mathcal{J}(m))=\mathcal{J}(g\cdot m).\end{aligned}$$

For all $\rho\in M/\mathcal{D}_G$, *the isotropy subgroup of* ρ *contains* G° *(the connected component of the identity in* G*).*

Proof: Let $g\in G$, $m,m'\in M$ be in the same leaf of $\mathcal{D}_G$, i.e. there exists without loss of generality one vector field $X\in\Gamma(\mathcal{D}_G)$ with flow F^X such that $F_t^X(m)=m'$ for some $t>0$. We have for all $s\in[0,t]$,

$$\begin{aligned}\frac{d}{ds}\left(\Phi_g\circ F_s^X\right)(m)&=T_{F_s^X(m)}\Phi_g\left(X(F_s^X(m))\right)\\ &=(\Phi_{g^{-1}}^*X)(g\cdot F_s^X(m))\in\mathcal{D}_G(g\cdot F_s^X(m)).\end{aligned}$$

Hence the curve $c(s)=\left(\Phi_g\circ F_s^X\right)(m)$ connecting $c(0)=\Phi_g(m)$ and $c(t)=\Phi_g(m')$ lies entirely in the leaf of $\mathcal{D}_G$ through the point $\Phi_g(m)$ and the assertion follows.

The Lie group G° is generated as a group by the exponential of an open neighbourhood of $0 \in \mathfrak{g}$. Thus, we can assume without loss of generality, that for any $g \in G^\circ$ and $m \in M$, there exists some $\xi \in \mathfrak{g}$ such that the curve $\gamma : [0, t] \to M$, $\gamma(s) = \Phi_{\exp(s\xi)}(m)$, has endpoints m and $g \cdot m$ (in reality, the points m and $g \cdot m$ can be joined with finitely many such curves). For all $s \in [0, t]$, we have $\dot{\gamma}(s) = \xi_M(\gamma(s)) \in \mathcal{D}_G(\gamma(s))$ and, arguing as above, we conclude that the whole curve $\gamma([0, t])$ lies in the leaf of $\mathcal{D}_G$ through m. Hence, if $\rho = \mathcal{J}(m)$, the equality $\Phi_g(\mathcal{J}(m)) = \mathcal{J}(g \cdot m) = \mathcal{J}(m)$ proves the statement. □

REMARK 4. The last statement shows that for all $\rho \in M/\mathcal{D}_G$, the isotropy subgroup G_ρ is the union of connected components of G and is therefore closed in G. This implies that the Lie group G_ρ acts properly on the leaf $\mathcal{J}^{-1}(\rho)$. It is obvious that this action is also free. In the optimal reduction results in [24], [17], the induced action on the leaves of the distribution is not necessarily proper. The reason why the action is here always proper is the inclusion $\mathcal{V} \subset \mathcal{D}_G$. △

REMARK 5. Note that if the nonholonomic system satisfies $\mathcal{H} \oplus \mathcal{V} = TM$, then the bundle $\mathcal{U}$ is given by $\mathcal{U} = \{0\}^{\omega_M} \cap \mathcal{H} = \mathcal{H}$ and hence $\mathcal{D}_G = \mathcal{U} + \mathcal{V} = TM$ is trivially integrable with the connected components of M as integral leaves. Hence, if M is connected, the method of reduction presented in the next subsection leads to the same reduced Dirac manifold as the Dirac reduction method of Section 3.2. △

4.4. Optimal reduction for nonholonomic systems

Assume as above that $\mathcal{D}_G$ is integrable and choose $\rho \in M/\mathcal{D}_G$. Since the isotropy subgroup G_ρ contains G°, the distribution $\mathcal{V}_\rho$ spanned by the fundamental vector fields of the action of G on $\mathcal{J}^{-1}(\rho)$ is $\mathcal{V}_\rho = \mathcal{V}|_{\mathcal{J}^{-1}(\rho)}$. Since $\mathcal{U} \subseteq \mathcal{D}_G$, $\mathcal{U}$ also restricts to $\mathcal{U}_\rho$ on the manifold $\mathcal{J}^{-1}(\rho)$. Let $\mathcal{H}_\rho$ be the intersection of $\mathcal{H} := TM \cap (T\pi_{T^*Q})^{-1}(\mathcal{D})$ with $T\mathcal{J}^{-1}(\rho)$.

Since the distribution $(\mathcal{H} \cap (\mathcal{V} \cap \mathcal{H})^{\omega_M} + \mathcal{V}) \cap \mathcal{H} = \mathcal{U} + (\mathcal{V} \cap \mathcal{H}) \subseteq \mathcal{D}_G$ is constant dimensional on the leaves of $\mathcal{D}_G$, the Dirac structure on a leaf $\mathcal{J}^{-1}(\rho)$ of $\mathcal{D}_G$ is given by

$$D_{\mathcal{J}^{-1}(\rho)}(m) = \left\{ (X(m), \alpha_m) \in T\mathcal{J}^{-1}(\rho) \oplus T^*\mathcal{J}^{-1}(\rho) \;\middle|\; \begin{array}{c} X \in \Gamma\left(\mathcal{U}_\rho + (\mathcal{V}_\rho \cap \mathcal{H}_\rho)\right), \\ \alpha - \mathbf{i}_X \omega_{\mathcal{J}^{-1}(\rho)} \in \Gamma\left(\left(\mathcal{U}_\rho + (\mathcal{V}_\rho \cap \mathcal{H}_\rho)\right)^\circ\right) \end{array} \right\}$$

for all $m \in \mathcal{J}^{-1}(\rho)$ (see [5]); here $i_\rho : \mathcal{J}^{-1}(\rho) \hookrightarrow M$ is the inclusion and $\omega_{\mathcal{J}^{-1}(\rho)} := i_\rho^* \omega_M$.

LEMMA 3. *Let $\mathcal{K}_\rho = \mathcal{V}_\rho \oplus \{0\}$ and $\mathcal{K}_\rho^\perp = T\mathcal{J}^{-1}(\rho) \oplus \mathcal{V}_\rho^\circ$ as in Section 2.3. Then $D_{\mathcal{J}^{-1}(\rho)} \cap \mathcal{K}_\rho^\perp$ is a vector bundle over $\mathcal{J}^{-1}(\rho)$.*

Proof: Since $\mathcal{H}+\mathcal{V}$ has constant rank on the n-dimensional manifold M, recall from Remark 2 that $D\cap\mathcal{K}^{\perp}$ is a vector bundle on M. We denote $r=\operatorname{rank}\mathcal{H}$, $n-r=\operatorname{rank}\mathcal{H}^{\circ}$, $l=\operatorname{rank}\mathcal{V}^{\circ}\cap\mathcal{H}^{\circ}$, $u=\operatorname{rank}\mathcal{U}$, and $s=\operatorname{rank}(\mathcal{U}+(\mathcal{V}\cap\mathcal{H}))|_{\mathcal{J}^{-1}(\rho)}$. Let $m\in\mathcal{J}^{-1}(\rho)$. As in Remark 2, choose local basis fields $H_1,\ldots,H_r$ for $\mathcal{H}$ and local basis 1-forms $\beta_1,\ldots,\beta_{n-r}$ for $\mathcal{H}^{\circ}$ defined on a neighbourhood U of m in M. Assume that $H_1,\ldots,H_u$ are local basis fields for $\mathcal{U}$, $H_1,\ldots,H_s$, with $u\leq s\leq r$, are basis fields for $\mathcal{U}+(\mathcal{V}\cap\mathcal{H})$ on $\mathcal{J}^{-1}(\rho)\cap U$, and $\beta_1,\ldots,\beta_l$ a basis of $\mathcal{V}^{\circ}\cap\mathcal{H}^{\circ}=(\mathcal{V}+\mathcal{H})^{\circ}$. Note that the 1-forms $\beta_1,\ldots,\beta_l$ vanish on $\mathcal{U}+\mathcal{V}\subseteq\mathcal{H}+\mathcal{V}$ and that $\beta_{l+1},\ldots,\beta_{n-r}$ do not vanish on $\mathcal{U}+\mathcal{V}$ (otherwise we would have $\beta_j\in\Gamma(\mathcal{U}^{\circ}\cap\mathcal{V}^{\circ}\cap\mathcal{H}^{\circ})=\Gamma(\mathcal{V}^{\circ}\cap\mathcal{H}^{\circ})$ for $j=l+1,\ldots,n-r$, in contradiction to the choice of $\beta_1,\ldots,\beta_{n-r}$). The Dirac structure $D_{\mathcal{J}^{-1}(\rho)}$ is then given on $U\cap\mathcal{J}^{-1}(\rho)$ by (see [3])

$$\operatorname{span}\left\{(\tilde{H}_1, i_\rho^*\mathbf{i}_{H_1}\omega_M),\ldots,(\tilde{H}_s, i_\rho^*\mathbf{i}_{H_s}\omega_M),(0,i_\rho^*\beta_{l+1}),\ldots,(0,i_\rho^*\beta_{n-r})\right\}, \tag{40}$$

where $\tilde{H}_1,\ldots,\tilde{H}_s$ are vector fields on $U\cap\mathcal{J}^{-1}(\rho)$ such that $\tilde{H}_i\sim_{i_\rho}H_i$ for $i=1,\ldots,s$. Note that $i_\rho^*\mathbf{i}_{H_i}\omega_M=\mathbf{i}_{\tilde{H}_i}\omega_{\mathcal{J}^{-1}(\rho)}$ for $i=1,\ldots,s$.

If $(X,\alpha)\in D_{\mathcal{J}^{-1}(\rho)}\cap(T\mathcal{J}^{-1}(\rho)\oplus\mathcal{V}_\rho^{\circ})$ then $X\in\Gamma(\mathcal{U}_\rho+(\mathcal{V}_\rho\cap\mathcal{H}_\rho))$, $\alpha\in\Gamma(\mathcal{V}_\rho^{\circ})$, and $\mathbf{i}_X\omega_{\mathcal{J}^{-1}(\rho)}-\alpha\in\Gamma((\mathcal{U}_\rho+(\mathcal{V}_\rho\cap\mathcal{H}_\rho))^{\circ})$. This is only possible if $\mathbf{i}_X\omega_{\mathcal{J}^{-1}(\rho)}=0$ on

$$\mathcal{V}_\rho\cap\big(\mathcal{U}_\rho+(\mathcal{V}_\rho\cap\mathcal{H}_\rho)\big)=(\mathcal{V}_\rho\cap\mathcal{U}_\rho)+(\mathcal{V}_\rho\cap\mathcal{H}_\rho)=\mathcal{L}_\rho$$

where $\mathcal{L}:=[(\mathcal{V}\cap\mathcal{H})\cap(\mathcal{V}\cap\mathcal{H})^{\omega_M}]+(\mathcal{V}\cap\mathcal{H})$.

We have two different cases. First, if $X\in\Gamma(\mathcal{U}_\rho)$ then for all $m\in\mathcal{J}^{-1}(\rho)$ and $V(m)\in\mathcal{L}(m)$ we have necessarily

$$\begin{aligned}(\mathbf{i}_X\omega_{\mathcal{J}^{-1}(\rho)})(m)(V(m))&=(i_\rho^*\omega_M)(m)(X(m),V(m))\\&=\omega_M(i_\rho(m))(T_m i_\rho X(m),T_m i_\rho V(m))=0\end{aligned}$$

where we have used

$$T_m i_\rho X(m)\in(\mathcal{H}\cap(\mathcal{V}\cap\mathcal{H})^{\omega_M})(i_\rho(m)),\qquad T_m i_\rho V(m)\in\mathcal{L}(i_\rho(m))$$

and the definition of $\mathcal{L}$. Hence, for all $X\in\Gamma(\mathcal{U}_\rho)$ we have $\mathbf{i}_X\omega_{\mathcal{J}^{-1}(\rho)}|_{\mathcal{L}_\rho}=0$ and hence we find $\alpha\in\Gamma(\mathcal{V}_\rho^{\circ})$ such that $(X,\alpha)\in\Gamma\left(D_{\mathcal{J}^{-1}(\rho)}\right)$. Second, for a section X of $\mathcal{V}\cap\mathcal{H}$ that does not take values in $\mathcal{U}_\rho$, the 1-form $\mathbf{i}_X\omega_{\mathcal{J}^{-1}(\rho)}$ does not vanish on $\mathcal{V}_\rho\cap\mathcal{H}_\rho$ and thus neither on $\mathcal{L}_\rho$. Consequently, the sections of $D_{\mathcal{J}^{-1}(\rho)}\cap\mathcal{K}_\rho^{\perp}$ have as first component a section of $\mathcal{U}_\rho$. Since for $i=l+1,\ldots,n-r$ we have $i_\rho^*\beta_i\notin\Gamma(\mathcal{V}_\rho^{\circ})$, we get

$$\begin{aligned}&D_{\mathcal{J}^{-1}(\rho)}\cap\mathcal{K}_\rho^{\perp}\\&\quad=\operatorname{span}\left\{\left(\tilde{H}_1,\mathbf{i}_{\tilde{H}_1}\omega_{\mathcal{J}^{-1}(\rho)}+\sum_{i=l+1}^{n-r}a_1^i i_\rho^*\beta_i\right),\ldots,\left(\tilde{H}_u,\mathbf{i}_{\tilde{H}_u}\omega_{\mathcal{J}^{-1}(\rho)}+\sum_{i=l+1}^{n-r}a_u^i i_\rho^*\beta_i\right)\right\},\end{aligned}$$

where the functions a_j^i are chosen such that $\mathbf{i}_{\tilde{H}_j}\omega_{\mathcal{J}^{-1}(\rho)}+\sum_{i=l+1}^{n-r}a_j^i i_\rho^*\beta_i$ are sections of $\mathcal{V}_\rho^\circ$ for $j = 1, \ldots, u$ and $i = l+1, \ldots, n-r$. Since the vector fields $\tilde{H}_1, \ldots, \tilde{H}_u$ are linearly independent, we have found basis fields for $D_{\mathcal{J}^{-1}(\rho)}$ on U. □

Hence, the reduced Dirac structure D_ρ on $\mathcal{J}^{-1}(\rho)/G_\rho$ is given, according to the general considerations in Section 2.3 (or [7]) by

$$D_\rho = \frac{\left[D_{\mathcal{J}^{-1}(\rho)} \cap \left(T\mathcal{J}^{-1}(\rho) \oplus \mathcal{V}_\rho^\circ\right)\right] + (\mathcal{V}_\rho \oplus \{0\})}{\mathcal{V}_\rho \oplus \{0\}} \Bigg/ G_\rho. \tag{41}$$

The next theorem gives an easier description of this reduced Dirac structure.

We write in the following $i_\rho : \mathcal{J}^{-1}(\rho) \hookrightarrow M$ for the inclusion and $\pi_\rho : \mathcal{J}^{-1}(\rho) \to M_\rho$ for the projection.

THEOREM 4 (Nonholonomic optimal point reduction by Dirac actions). *Assume that $\mathcal{H} + \mathcal{V}$ has constant rank on M and that the Lie group G acts freely and properly on M by Dirac actions. If $\mathcal{D}_G = \mathcal{U} + \mathcal{V}$ is an integrable subbundle of TM, then for any $\rho \in M/\mathcal{D}_G$ we have the following results.*

(i) *The orbit space $M_\rho = \mathcal{J}^{-1}(\rho)/G_\rho$ is a smooth regular Dirac quotient manifold whose Dirac structure D_ρ is given by the graph of a* nondegenerate *(not necessarily closed) 2-form ω_ρ.*

(ii) *Let $h \in C^\infty(M)^G$ be an admissible and G-invariant Hamiltonian and X_h the (unique) solution of the implicit Hamiltonian system $(X_h, \mathbf{d}h) \in \Gamma(D)$. Then $X_h \in \Gamma(\mathcal{U})$ and we have $(X_h|_{\mathcal{J}^{-1}(\rho)}, i_\rho^*\mathbf{d}h) \in \Gamma\left(D_{\mathcal{J}^{-1}(\rho)}\right)$.*

(iii) *The flow F_t of X_h leaves $\mathcal{J}^{-1}(\rho)$ invariant, commutes with the G-action, and therefore induces a flow F_t^ρ on M_ρ uniquely determined by the relation $\pi_\rho \circ F_t \circ i_\rho = F_t^\rho \circ \pi_\rho$.*

(iv) *The flow F_t^ρ is the flow of a vector field X_{h_ρ} in $\mathfrak{X}(M_\rho)$ that is the solution of the Hamiltonian system $\mathbf{i}_{X_{h_\rho}}\omega_\rho = \mathbf{d}h_\rho$, where the function $h_\rho \in C^\infty(M_\rho)$ is given by the equality $h_\rho \circ \pi_\rho = h \circ i_\rho$.*

Proof: According to Remark 4, the G_ρ-action on $\mathcal{J}^{-1}(\rho)$ is free and proper. Thus, the quotient $\mathcal{J}^{-1}(\rho)/G_\rho$ is a regular quotient manifold and the projection $\pi_\rho : \mathcal{J}^{-1}(\rho) \to M_\rho$ is a smooth surjective submersion. We denote from now on by $\omega_{\mathcal{J}^{-1}(\rho)} := i_\rho^*\omega_M$ the pull back of ω_M to $\mathcal{J}^{-1}(\rho)$.

(i) With Lemma 3, we get

$$\begin{aligned}&\frac{\left(D_{\mathcal{J}^{-1}(\rho)} \cap \mathcal{K}_\rho^\perp\right) + \mathcal{K}_\rho}{\mathcal{K}_\rho}\\ &= \left\{ (\widehat{X}, \alpha) \in \Gamma\left(\left(T\mathcal{J}^{-1}(\rho)/\mathcal{V}_\rho\right) \oplus T^*\mathcal{J}^{-1}(\rho)\right) \;\middle|\; \begin{array}{c} \widehat{X} = X (\operatorname{mod} \mathcal{V}_\rho) \text{ with} \\ X \in \Gamma(\mathcal{U}_\rho), \alpha \in \Gamma(\mathcal{V}_\rho^\circ), \\ \text{and } \alpha - \mathbf{i}_X\omega_{\mathcal{J}^{-1}(\rho)} \\ \in \Gamma\left((\mathcal{U}_\rho + (\mathcal{V}_\rho \cap \mathcal{H}_\rho))^\circ\right) \end{array} \right\}.\end{aligned} \tag{42}$$

The G_ρ-quotient of this bundle defines the reduced Dirac structure D_ρ on M_ρ.

Note that the fibers

$$\left(\mathcal{U}_\rho/\mathcal{V}_\rho\right)(m) := (\mathcal{U}_\rho + \mathcal{V}_\rho)(m)/\mathcal{V}(m) = T_m\mathcal{J}^{-1}(\rho)/\mathcal{V}_\rho(m), \qquad m \in \mathcal{J}^{-1}(\rho),$$

of the vector bundle $\mathcal{U}_\rho/\mathcal{V}_\rho$, project surjectively to $T_{\pi_\rho(m)}M_\rho$. Like in Section 2.3, for each G-invariant $X \in \Gamma(\mathcal{U}_\rho)$ we identify $\widehat{X} = X(\operatorname{mod}\mathcal{V}_\rho)$ with the section $\bar{X}$ of M_ρ such that $T\pi_\rho \circ X = \bar{X} \circ \pi_\rho$. Write each G-invariant $\alpha \in \Gamma(\mathcal{V}_\rho^\circ)$ as $\alpha = \pi_\rho^*\bar{\alpha}$ for some $\bar{\alpha} \in \Omega^1\left(M_\rho\right)$.

Next we show that D_ρ is the graph of a nondegenerate 2-form. We begin by giving a formula for this 2-form ω_ρ. Let $\bar{X}, \bar{Y} \in \mathfrak{X}(M_\rho)$ and choose G-invariant $X, Y \in \mathfrak{X}(\mathcal{J}^{-1}(\rho))$ that are π_ρ-related to $\bar{X}$ and $\bar{Y}$, respectively. Write $X = \tilde{X} + V$ and $Y = \tilde{Y} + W$ with $\tilde{X}, \tilde{Y} \in \Gamma(\mathcal{U}_\rho)^G$ and $V, W \in \Gamma(\mathcal{V}_\rho)^G$. Then $\tilde{X}$ and $\tilde{Y}$ are also π_ρ-related to $\bar{X}$ and $\bar{Y}$ and we can write, using the existence of $\bar{\alpha} \in \Omega^1(M_\rho)$ such that $(\bar{X}, \bar{\alpha}) \in \Gamma(D_\rho)$,

$$\omega_\rho(\bar{X}, \bar{Y}) = \bar{\alpha}(\bar{Y}) = (\pi_\rho^*\bar{\alpha})(\tilde{Y}) = \omega_{\mathcal{U}_\rho}(\tilde{X}, \tilde{Y}),$$

since $\tilde{X}$ has to be the (unique) section of $\mathcal{U}_\rho$ associated to the 1-form $\pi_\rho^*\bar{\alpha}$ (see (3)) and where $\omega_{\mathcal{U}_\rho}$ is the restriction of $\omega_{\mathcal{J}^{-1}(\rho)}$ to $\mathcal{U}_\rho \times \mathcal{U}_\rho$.

We prove that ω_ρ is nondegenerate. Let $\bar{X} \in \mathfrak{X}(M_\rho)$ with $\omega_\rho(\bar{X}, \bar{Y}) = 0$ for all $\bar{Y} \in \mathfrak{X}(M_\rho)$. Choose a G-invariant section $\tilde{X} \in \Gamma(\mathcal{U}_\rho)$ as above. Extend $\tilde{X}$ to a local vector field X on M, that is, $X \in \Gamma(\mathcal{U}) \subseteq \mathfrak{X}(M)$ satisfies $\tilde{X} \sim_{i_\rho} X$. For $m \in \mathcal{J}^{-1}(\rho)$ and $v \in \mathcal{U}(m) \subseteq T_mM$ we have

$$\omega_{\mathcal{H}}(m)(X(m), v) = \omega_{\mathcal{U}}(m)(\tilde{X}(m), v) = \omega_\rho\left(\pi_\rho(m)\right)(\bar{X}(\pi_\rho(m)), T_m\pi_\rho(v)) = 0.$$

Thus, the vector $X(m)$ is an element of $\mathcal{U}(m) = (\mathcal{H} \cap \mathcal{V})^{\omega_{\mathcal{H}}}(m)$ that is $\omega_{\mathcal{H}}(m)$-orthogonal to all $v \in \mathcal{U}(m)$ and hence lies in $((\mathcal{H} \cap \mathcal{V})^{\omega_{\mathcal{H}}})^{\omega_{\mathcal{H}}}(m)$. Since $\omega_{\mathcal{H}}$ is nondegenerate, we have $((\mathcal{H} \cap \mathcal{V})^{\omega_{\mathcal{H}}})^{\omega_{\mathcal{H}}} = \mathcal{H} \cap \mathcal{V}$. This yields $\tilde{X}(m) = X(m) \in (\mathcal{H} \cap \mathcal{V})(m)$ and thus the vector $\bar{X}(m)$ is zero in T_mM_ρ.

(ii) Recall that, since $\mathsf{G}_0 = \{0\}$, the solution X_h of the implicit Hamiltonian system $(X, \mathbf{d}h) \in \Gamma(D)$ is unique: if Y is another solution, then $Y - X_h \in \Gamma(\mathsf{G}_0) = \{0_M\}$.

We know already that $X_h \in \Gamma(\mathcal{U}_\rho)$. Furthermore, we have for all $Y \in \Gamma(\mathcal{U}_\rho)$, $V \in \Gamma(\mathcal{V}_\rho \cap \mathcal{H}_\rho)$ and all $m \in \mathcal{J}^{-1}(\rho)$

$$\begin{aligned}\omega_{\mathcal{J}^{-1}(\rho)}(m)(X_h(m), Y(m) + V(m)) &= \omega_M(m)(X_h(m), Y(m) + V(m))\\ &= \mathbf{d}h_m(Y(m) + V(m)) = (i_\rho^*\mathbf{d}h)(m)(Y(m) + V(m))\end{aligned}$$

and the assertion follows.

(iii) The fact that the flow of X_h leaves $\mathcal{J}^{-1}(\rho)$ invariant follows from the preceding statement since we have $X_h \in \Gamma(\mathcal{D}_G)$. By G-invariance of D we have $(\Phi_g^*X_h, \Phi_g^*\mathbf{d}h) \in \Gamma(D)$ for all $g \in G$. Since h is G-invariant, the equality $\Phi_g^*\mathbf{d}h = \mathbf{d}\Phi_g^*h = \mathbf{d}h$ holds and thus we have $\Phi_g^*X_h - X_h \in \Gamma(\mathsf{G}_0) = \{0_M\}$. The vector field X_h is consequently G-equivariant and its flow commutes with the G-action.

(iv) Since $X_h \in \Gamma(\mathcal{U}_\rho)$ and $i_\rho^* \mathbf{d}h \in \mathcal{V}_\rho^\circ$, we have

$$(X_h, \mathbf{d}h) \in \Gamma\left(D_{\mathcal{J}^{-1}(\rho)} \cap (T\mathcal{J}^{-1}(\rho) \oplus \mathcal{V}_\rho^\circ)\right).$$

The flow F_t^ρ on M_ρ induces a vector field $X_{h_\rho} \in \mathfrak{X}(M_\rho)$. Therefore, taking the t-derivative of the relation in (iii) we get

$$X_{h_\rho}(\pi_\rho(m)) = \left.\frac{d}{dt}\right|_{t=0} F_t^\rho(\pi_\rho(m)) = \left.\frac{d}{dt}\right|_{t=0} \left(\pi_\rho \circ F_t\right)(m) = T_m\pi_\rho X_h(m),$$

that is, $X_h \sim_{\pi_\rho} X_{h_\rho}$. Choose $\bar{Y} \in \mathfrak{X}(M_\rho)$, $Y \in \Gamma(\mathcal{U}_\rho)^G$, and $V \in \Gamma(\mathcal{V}_\rho)^G$ such that $T\pi_\rho \circ (Y + V) = \bar{Y} \circ \pi_\rho$. Then, for all $m \in \mathcal{J}^{-1}(\rho)$ we get

$$\begin{aligned}
&\omega_\rho(\pi_\rho(m))\left(X_{h_\rho}(\pi_\rho(m)), \bar{Y}(\pi_\rho(m))\right)\\
&\quad= (\pi_\rho^*\omega_\rho)(m)(X_h(m), Y(m) + V(m)) = (\pi_\rho^*\omega_\rho)(m)(X_h(m), Y(m))\\
&\quad= \omega_{\mathcal{J}^{-1}(\rho)}(m)(X_h(m), Y(m)) = (i_\rho^*\mathbf{d}h)_m(Y(m)) = (\pi_\rho^*\mathbf{d}h_\rho)_m(Y(m) + V(m))\\
&\quad= (\mathbf{d}h_\rho)_{\pi_\rho(m)}(\bar{Y}(\pi_\rho(m))),
\end{aligned}$$

so we have $\mathbf{i}_{X_{h_\rho}}\omega_\rho = \mathbf{d}h_\rho$, as claimed. □

4.5. Examples of optimal reduction for nonholonomic systems

4.5.1. The constrained particle in space

We return to the example treated in Subsection 3.3.1 and use the same notations and conventions. The distribution $\mathcal{V} \cap \mathcal{H}$ is pointwise the span of the vector field $\partial_x + y\partial_z$. Since $\mathcal{V}^\circ$ is spanned by the covector fields $\mathbf{d}y$, $\mathbf{d}p_x$, and $\mathbf{d}p_y$, the considerations in Subsection 3.3.1 yield $\mathbf{i}_{\partial_x + y\partial_z}\omega_M = -(1 + y^2)\mathbf{d}p_x - yp_x\mathbf{d}y \in \Gamma(\mathcal{V}^\circ)$ and hence $(\partial_x + y\partial_z, -(1+y^2)\mathbf{d}p_x - yp_x\mathbf{d}y) \in \Gamma(D \cap (\mathcal{V} \oplus \mathcal{V}^\circ))$. Hence the distribution $\mathcal{D}_G$ is in this case $\ker\{-(1+y^2)\mathbf{d}p_x - yp_x\mathbf{d}y\} = \ker\{\mathbf{d}f\}$, where $f(x, y, z, p_x, p_y) = \sqrt{1 + y^2}\,p_x$ is the constant of motion (in agreement with [2]). Note that by Example 1, $\mathbf{d}f$ is the 1-form giving the Nonholonomic Noether Theorem. Hence

$$\mathcal{D}_G = \operatorname{span}\{\partial_{p_y}, \partial_x, \partial_z, yp_x\partial_{p_x} - (y^2 + 1)\partial_y\}$$

is obviously involutive and constant dimensional (and consequently integrable). This shows that $M/\mathcal{D}_G = \mathbb{R}$. The Dirac structure on a leaf $f^{-1}(\mu)$, $\mu \in M/\mathcal{D}_G = \mathbb{R}$, of this distribution is given by

$$\begin{aligned}
&D_{f^{-1}(\mu)} =\\
&\left\{(X, \alpha) \in \Gamma(Tf^{-1}(\mu) \oplus T^*f^{-1}(\mu)) \mid X \in \Gamma(\mathcal{H} \cap \mathcal{D}_G), \alpha - \mathbf{i}_X i_\mu^*\omega_M \in \Gamma((\mathcal{H} \cap \mathcal{D}_G)^\circ)\right\}
\end{aligned}$$

and a computation yields

$$\begin{aligned}
&D_{f^{-1}(\mu)} =\\
&\quad\operatorname{span}\left\{\left(\partial_{p_y}, -\mathbf{d}y\right), (\partial_x + y\partial_z, 0)\,(0, \mathbf{d}z - y\mathbf{d}x), \left((1 + y^2)\partial_y - yp_x\partial_{p_x}, (1 + y^2)\mathbf{d}p_y\right)\right\}
\end{aligned}$$

because the 1-form $(1+y^2)\mathbf{d}p_x + yp_x\mathbf{d}y$ vanishes on $T\left(f^{-1}(\mu)\right)$. Since G is in this case connected, we have $G_\mu = G$ (see Remark 4). Consider the codistribution $\mathcal{V}^\circ$ on $f^{-1}(\mu)$ and get

$$\begin{aligned}&D_{f^{-1}(\mu)} \cap (Tf^{-1}(\mu) \oplus \mathcal{V}^\circ)\\ &\qquad= \operatorname{span}\left\{\left(\partial_{p_y}, -\mathbf{d}y\right), (\partial_x + y\partial_z, 0)\left((1+y^2)\partial_y - yp_x\partial_{p_x}, (1+y^2)\mathbf{d}p_y\right)\right\}.\end{aligned}$$

Hence the reduced Dirac structure D_μ on $M_\mu = f^{-1}(\mu)/G$ is given by the formula

$$\begin{aligned}D_\mu &= \frac{\left[D_{f^{-1}(\mu)} \cap \left(T\left(f^{-1}(\mu)\right) \oplus \mathcal{V}^\circ\right)\right] + \mathcal{V} \oplus \{0\}}{\mathcal{V} \oplus \{0\}} \Bigg/ G\\ &= \operatorname{span}\left\{\left(\partial_{p_y}, -\mathbf{d}y\right), \left((1+y^2)\partial_y - yp_x\partial_{p_x}, (1+y^2)\mathbf{d}p_y\right)\right\}.\end{aligned}$$

This corresponds exactly to a symplectic leaf (with its associated Dirac structure) of the Poisson structure (26) obtained in the first part of this example (see Subsection 3.3.1).

Finally we compute $\mathcal{R}$ for this example. Since $\mathcal{H}^\circ$ is one-dimensional, we get $\mathcal{R} = \mathcal{H}^\circ$ or $\mathcal{R} = \{0\}$. Recall that D is the span of

$$\begin{aligned}&\left\{\left(\partial_{p_y}, -\mathbf{d}y\right), \left(\partial_x + y\partial_z, (1+y^2)\mathbf{d}p_x + yp_x\mathbf{d}y\right), (0, \mathbf{d}z - y\mathbf{d}x), \left(\partial_y, \mathbf{d}p_y - p_x\mathbf{d}z\right),\right.\\ &\qquad\qquad\left.\left(\partial_{p_x}, -y\mathbf{d}z - \mathbf{d}x\right)\right\}\end{aligned}$$

where we have computed:

$$\begin{aligned}\mathbf{i}_{\partial_x + y\partial_z}\omega_M &= (1+y^2)\mathbf{d}p_x + yp_x\mathbf{d}y,\\ \mathbf{i}_{\partial_y}\omega_M &= \mathbf{d}p_y - p_x\mathbf{d}z,\\ \mathbf{i}_{\partial_{p_y}}\omega_M &= -\mathbf{d}y,\\ \mathbf{i}_{\partial_{p_x}}\omega_M &= -y\mathbf{d}z - \mathbf{d}x.\end{aligned}$$

Since $\mathcal{U} = \operatorname{span}\{\partial_{p_y}, (1+y^2)\partial_y - yp_x\partial_{p_x}, \partial_x + y\partial_z\}$, we conclude from

$$\begin{aligned}\mathbf{i}_{(1+y^2)\partial_y - yp_x\partial_{p_x}}\omega_M &= (1+y^2)(\mathbf{d}p_y - p_x\mathbf{d}z) - yp_x(-y\mathbf{d}z - \mathbf{d}x)\\ &= (1+y^2)\mathbf{d}p_y - p_x(\mathbf{d}z - y\mathbf{d}x),\end{aligned}$$

that the distribution $\mathcal{R}$ is equal to $\mathcal{H}^\circ$. The constant of motion that we have found above is a *gauge constant of motion*.

4.5.2. The vertical rolling disk

In this subsection we shall determine the nonholonomic Momentum equations for the example of the vertical rolling disk studied in Subsection 3.3.2. The Dirac structure for this nonholonomic system is given by

$$D = \operatorname{span}\left\{\left(\partial_\phi, \mathbf{d}p_\phi + \frac{\mu R\sin\phi}{I}p_\theta\mathbf{d}x - \frac{\mu R\cos\phi}{I}p_\theta\mathbf{d}y\right), \left(\partial_{p_\phi}, -\mathbf{d}\phi\right),\right.$$
$$\left(\partial_{p_\theta}, -\frac{\mu R\cos\phi}{I}\mathbf{d}x - \frac{\mu R\sin\phi}{I}\mathbf{d}y - \mathbf{d}\theta\right), (0, \mathbf{d}x - R\cos\phi\mathbf{d}\theta),$$
$$\left.(0, \mathbf{d}y - R\sin\phi\mathbf{d}\theta), \left(\partial_\theta + R\cos\phi\partial_x + R\sin\phi\partial_y, (1 + \frac{\mu R^2}{I})\mathbf{d}p_\theta\right)\right\}$$

where we have computed

$$\begin{aligned}
\mathbf{i}_{\partial_\phi}\omega_M &= \mathbf{d}p_\phi + \frac{\mu R\sin\phi}{I}p_\theta\mathbf{d}x - \frac{\mu R\cos\phi}{I}p_\theta\mathbf{d}y,\\
\mathbf{i}_{\partial_{p_\theta}}\omega_M &= -\frac{\mu R}{I}\cos\phi\mathbf{d}x - \frac{\mu R}{I}\sin\phi\mathbf{d}y - \mathbf{d}\theta,\\
\mathbf{i}_{\partial_{p_\phi}}\omega_M &= -\mathbf{d}\phi,\\
\mathbf{i}_{\partial_\theta + R\cos\phi\partial_x + R\sin\phi\partial_y}\omega_M &= \mathbf{d}p_\theta + R\cos\phi\left(\frac{\mu R\cos\phi}{I}\mathbf{d}p_\theta - \frac{\mu R\sin\phi}{I}\mathbf{d}\phi\right)\\
&\quad + R\sin\phi\left(+\frac{\mu R\sin\phi}{I}\mathbf{d}p_\theta + \frac{\mu R\cos\phi}{I}\mathbf{d}\phi\right),\\
&= \left(1 + \frac{\mu R^2}{I}\right)\mathbf{d}p_\theta.
\end{aligned}$$

We consider again the three possible Lie group actions:

1. *The case* $G = \mathbb{R}^2$ [10].

Here, $\mathcal{V}^\circ = \operatorname{span}\{\mathbf{d}p_\phi, \mathbf{d}\phi, \mathbf{d}p_\theta, \mathbf{d}\theta\}$ but there are no nontrivial horizontal symmetries and hence the distribution $\mathcal{D}_G$ is simply the whole bundle TM. We next compute $\mathcal{R}$. The vector bundle $D \cap \mathcal{K}^\perp$ is given in this case by

$$\operatorname{span}\left\{\left(\partial_\phi, \mathbf{d}p_\phi\right), \left(\partial_{p_\theta}, -\left(1 + \frac{\mu R^2}{I}\right)\mathbf{d}\theta\right), \left(\partial_{p_\phi}, -\mathbf{d}\phi\right),\right.$$
$$\left.\left(\partial_\theta + R\cos\phi\partial_x + R\sin\phi\partial_y, \left(1 + \frac{\mu R^2}{I}\right)\mathbf{d}p_\theta\right)\right\}.$$

To get this, we have added

$$\frac{\mu R}{I}p_\theta\cos\phi(\mathbf{d}x - R\cos\phi\mathbf{d}\theta) + \frac{\mu R}{I}\sin\phi(\mathbf{d}y - R\sin\phi\mathbf{d}\theta) \tag{43}$$

to $\mathbf{i}_{\partial_{p_\theta}}\omega_M$ and

$$-\frac{\mu R}{I}p_\theta\sin\phi(\mathbf{d}x - R\cos\phi\mathbf{d}\theta) + \frac{\mu R}{I}\cos\phi(\mathbf{d}y - R\sin\phi\mathbf{d}\theta) \tag{44}$$

to $\mathbf{i}_{\partial_\phi}\omega_M$. This yields $\mathcal{R} = \mathcal{H}^\circ$ and thus the distribution $\mathcal{R}^\circ \cap \mathcal{V}$ is equal to $\mathcal{H} \cap \mathcal{V}$ and hence trivial.

2. *The case* $G = \mathrm{SE}(2)$ [6].

In this case, we have $\mathcal{V}^\circ = \operatorname{span}\{\mathbf{d}p_\phi, \mathbf{d}p_\theta, \mathbf{d}\theta\}$ and $\mathcal{H} \cap \mathcal{V} = \operatorname{span}\{\partial_\phi\}$. We get $\mathcal{H} \cap \mathcal{V} \subseteq (\mathcal{H} \cap \mathcal{V})^{\omega_M}$. A direct computation gives

$$\mathbf{i}_{\partial_\phi}\omega_M = \mathbf{d}p_\phi + \frac{\mu R \sin\phi}{I} p_\theta \mathbf{d}x - \frac{\mu R \cos\phi}{I} p_\theta \mathbf{d}y.$$

Adding

$$\frac{\mu R}{I} p_\theta (\cos\phi \mathbf{d}y - \sin\phi \mathbf{d}x) d$$
$$= \frac{\mu R}{I} p_\theta \left(\cos\phi(\mathbf{d}y - R\sin\phi\mathbf{d}\theta) - \sin\phi(\mathbf{d}x - R\cos\phi\mathbf{d}\theta)\right) \in \Gamma(\mathcal{H}^\circ)$$

to this expression we see that $\Gamma(D \cap (\mathcal{V} \oplus \mathcal{V}^\circ))$ is spanned by $(\partial_\phi, \mathbf{d}p_\phi)$ and thus the distribution $\mathcal{D}_G = \ker \mathbf{d}p_\phi$ is obviously integrable. For a value $\rho \in \mathbb{R}$ of the map p_ϕ, the reduced Dirac structure on M_ρ is spanned by

$$\left(\partial_\theta, \left(1 + \frac{\mu R^2}{I}\right)\mathbf{d}p_\theta\right) \quad \text{and} \quad \left(\partial_{p_\theta}, -\left(1 + \frac{\mu R^2}{I}\right)\mathbf{d}\theta\right).$$

The nonholonomic Noether theorem yields a constant of motion but this constant does not arise from an element of $\mathfrak{g}$ whose corresponding fundamental vector field is lying in $\Gamma(\mathcal{V} \cap \mathcal{R}^\circ)$: we have computed in §3 that, in this case, $\mathcal{U}$ is the span of the three vector fields ∂_ϕ, ∂_{p_θ}, and $\partial_\theta + R\cos\phi\partial_x + R\sin\phi\partial_y$. Again, we have to add (43) to $\mathbf{i}_{\partial_{p_\theta}}\omega_M$ and (44) to $\mathbf{i}_{\partial_\phi}\omega_M$ in order to get sections of $\mathcal{V}^\circ$. Thus, we need the whole of $\mathcal{H}^\circ$ in the construction of $D \cap \mathcal{K}^\perp$.

3. *The case* $G = \mathbb{S}^1 \times \mathbb{R}^2$ [6].

Here, we have $\mathcal{V}^\circ = \operatorname{span}\{\mathbf{d}p_\phi, \mathbf{d}p_\theta, \mathbf{d}\phi\}$ and $\mathcal{H} \cap \mathcal{V}$ is again one-dimensional: this time it is the span of the vector field $\partial_\theta + R\cos\phi\partial_x + R\sin\phi\partial_y$. Thus, $\Gamma(D \cap (\mathcal{V} \oplus \mathcal{V}^\circ))$ is spanned by

$$\left(\partial_\theta + R\cos\phi\partial_x + R\sin\phi\partial_y, (1 + \frac{\mu R^2}{I})\mathbf{d}p_\theta\right)$$

and the distribution $\mathcal{D}_G = \ker\{(1 + \frac{\mu R^2}{I})\mathbf{d}p_\theta\}$ is again integrable. For a value $\rho \in \mathbb{R}$ of the map p_θ, the reduced Dirac structure on M_ρ is spanned by $(\partial_\phi, \mathbf{d}p_\phi)$ and $(\partial_{p_\phi}, -\mathbf{d}\phi)$.

In this case, we have $\mathcal{U} = \operatorname{span}\{\partial_\phi, \partial_{p_\phi}, \partial_\theta + R\cos\phi\partial_x + R\sin\phi\partial_y\}$. We get $\mathcal{R} = \operatorname{span}\{\frac{\mu R}{I} p_\theta(\sin\phi\mathbf{d}x - \cos\phi\mathbf{d}y)\}$ from the considerations for the second case. Thus we have $\mathcal{R}^\circ = \operatorname{span}\{\partial_\phi, \partial_{p_\phi}, \partial_{p_\theta}, \partial_\theta, \cos\phi\partial_x + \sin\phi\partial_y\}$ and our constant of motion p_θ really arises from a fundamental vector field lying in $\mathcal{R}^\circ$.

4. *The case* $G = \mathrm{SE}(2) \times \mathbb{S}^1$ [6].

In this last case, we have $\mathcal{V}^\circ = \operatorname{span}\{\mathbf{d}p_\phi, \mathbf{d}p_\theta\}$ and $\mathcal{H} \cap \mathcal{V}$ is this time two-dimensional: it is the span of the vector fields $\partial_\theta + R\cos\phi\partial_x + R\sin\phi\partial_y$ and ∂_ϕ. Thus, $\Gamma(D \cap (\mathcal{V} \oplus \mathcal{V}^\circ))$ is spanned by $(\partial_\theta + R\cos\phi\partial_x + R\sin\phi\partial_y, (1 + \frac{\mu R^2}{I})\mathbf{d}p_\theta)$ and

$(\partial_\phi, \mathbf{d}p_\phi)$ and the distribution $\mathcal{D}_G = \ker\{(1+\frac{\mu R^2}{I})\mathbf{d}p_\theta, \mathbf{d}p_\phi\}$ is integrable. Here, the reduced manifolds are single points.

We have $\mathcal{U} = \operatorname{span}\{\partial_\phi, \partial_\theta + R\cos\phi\partial_x + R\sin\phi\partial_y\}$ and we get as above $\mathcal{R}^\circ = \operatorname{span}\{\partial_\phi, \partial_{p_\phi}, \partial_{p_\theta}, \partial_\theta, \cos\phi\partial_x + \sin\phi\partial_y\}$.

4.5.3. The Chaplygin skate

We continue here the examples of Subsection 3.3.3.

The standard Chaplygin skate. We have seen that $\mathcal{V}\cap\mathcal{H} = \operatorname{span}\{\partial_\theta, \cos\theta\,\partial_x+\sin\theta\,\partial_y\}$. If we choose the basis $\xi^1 := (1,0,0)$, $\xi^2 := (0,1,0)$, $\xi^3 := (0,0,1)$ of the Lie algebra $\mathfrak{se}(2)$ we get $\xi^1_M = \partial_\theta - y\partial_x + x\partial_y$, $\xi^2_M = \partial_x$, and $\xi^3_M = \partial_y$. Hence, the sections $\xi^1 + y\xi^2 - x\xi^3$ and $\cos\theta\xi^2 + \sin\theta\xi^3 \in \Gamma(\mathfrak{g}^{\mathcal{H}})$ are spanning sections of $\mathfrak{g}^{\mathcal{H}}$ and the corresponding nonholonomic Noether equations are $s\cos\theta\mathbf{d}p_y - s\sin\theta\mathbf{d}p_x$ and $\cos\theta\mathbf{d}p_x + \sin\theta\mathbf{d}p_y$, respectively. Consequently, the two spanning sections $-s^{-1}\sin\theta\xi^1 + (\cos^2\theta - s^{-1}y\sin\theta)\xi^2 + \sin\theta(\cos\theta + s^{-1}x)\xi^3$ and $s^{-1}\cos\theta\xi^1 + \cos\theta(\sin\theta + s^{-1}y)\xi^2 + (\sin^2\theta - s^{-1}x\cos\theta)\xi^3$ of $\mathfrak{g}^{\mathcal{H}}$ lead to the nonholonomic Noether equations $\mathbf{d}p_x$ and $\mathbf{d}p_y$ respectively. Thus, $\mathcal{D}_G = \mathcal{U} + \mathcal{V} = \operatorname{span}\{\partial_x, \partial_y, \partial_\theta\}$ is found easily because $\mathcal{D}_G$ is the kernel of $\{\mathbf{d}p_x, \mathbf{d}p_y\}$. This is obviously integrable. The induced Dirac structure on a leaf $f^{-1}(a,b)$ (where f is the projection on (p_x, p_y)) of $\mathcal{D}_G$ is given by

$$D_{f^{-1}(a,b)} = \operatorname{span}\{(\cos\theta\,\partial_x + \sin\theta\,\partial_y, 0), (\partial_\theta, 0), (0, \sin\theta\,\mathbf{d}x - \cos\theta\,\mathbf{d}y)\}.$$

Here the reduced space $M_{(a,b)}$ is a single point. The reduced Dirac structure is hence trivial, as can also be seen from the formula

$$\frac{(D_{f^{-1}(a,b)} \cap \mathcal{K}^\perp_{(a,b)}) + \mathcal{K}_{(a,b)}}{\mathcal{K}_{(a,b)}} / G_{(a,b)}.$$

We could also consider the action of $\mathbb{S}^1$ on M given by $\Phi : \mathbb{S}^1 \times M \to M$, $(\alpha, \theta, x, y, p_x, p_y) \mapsto (\theta + \alpha, x\cos\alpha - y\sin\alpha, x\sin\alpha + y\cos\alpha, p_x\cos\alpha - p_y\sin\alpha, p_x\sin\alpha + p_y\cos\alpha)$. Here, we would have $\mathcal{V}\cap\mathcal{H} = \{0\}$ except for the points satisfying $x = -\sin\theta$ and $y = -\cos\theta$, so the condition that $\mathcal{V}\cap\mathcal{H}$ has constant rank is not satisfied (we have also $\mathcal{V} + \mathcal{H} \neq TM$).

If we consider the action of $\mathbb{R}^2$ on M given by $\Phi : \mathbb{R}^2 \times M \to M$, $(r, s, \theta, x, y, p_x, p_y) \mapsto (\theta, x + r, y + s, p_x, p_y)$, we have $\mathcal{V} = \operatorname{span}\{\partial_x, \partial_y\}$. Hence, $\mathcal{V} + \mathcal{H} = TM$ and $\mathcal{V}\cap\mathcal{H} = \{\cos\theta\,\partial_x + \sin\theta\,\partial_y\}$ has constant rank on M. The distribution $\mathcal{D}_G$ is given by $\mathcal{D}_G = (\ker\{\cos\theta\mathbf{d}p_x + \sin\theta\mathbf{d}p_y\} \cap \mathcal{H}) + \mathcal{V} = \operatorname{span}\{\sin\theta\,\partial_{p_x} - \cos\theta\,\partial_{p_y}, \partial_\theta, \partial_x, \partial_y\}$. This vector bundle is not involutive and hence it is not integrable. Since $\mathcal{U} = \ker\{\cos\theta\mathbf{d}p_x + \sin\theta\mathbf{d}p_y\}\cap\mathcal{H} = \operatorname{span}\{\sin\theta\,\partial_{p_x} - \cos\theta\,\partial_{p_y}, \partial_\theta, \cos\theta\,\partial_x + \sin\theta\,\partial_y\}$, it is easy to see that $\mathcal{R} = \mathcal{H}^\circ$ and hence $\mathcal{R}^\circ\cap\mathcal{V} = \mathcal{V}\cap\mathcal{H} = \{\cos\theta\,\partial_x + \sin\theta\,\partial_y\}$, which confirms the fact that the nonholonomic Noether equation yields in this case no constant of motion.

The Chaplygin skate with a rotor on it. We have $\mathcal{V}\cap\mathcal{H} = \operatorname{span}\{\partial_\phi, \partial_\theta, \cos\theta\,\partial_x + \sin\theta\,\partial_y\}$. If we choose the basis $\xi^1 := (1,0,0,0)$, $\xi^2 := (0,1,0,0)$, $\xi^3 := (0,0,1,0)$,

$\xi^4 := (0,0,0,1)$ of the Lie algebra $\mathbb{R}\times\mathfrak{se}(2)$ we get $\xi^1_M = \partial_\phi$, $\xi^2_M = \partial_\theta - y\partial_x + x\partial_y$, $\xi^3_M = \partial_x$, and $\xi^4_M = \partial_y$. Hence, the sections ξ^1, $\xi^2 + y\xi^3 - x\xi^4$, and $\cos\theta\xi^3 + \sin\theta\xi^4 \in \Gamma(\mathfrak{g}^{\mathcal{H}})$ are spanning sections of $\mathfrak{g}^{\mathcal{H}}$ and the corresponding nonholonomic Noether equations are $\mathbf{d}p_\phi$, $s\cos\theta\mathbf{d}p_y - s\sin\theta\mathbf{d}p_x + \mathbf{d}p_\phi$, and $\cos\theta\mathbf{d}p_x + \sin\theta\mathbf{d}p_y$, respectively. Thus, the three spanning sections ξ^1, $s^{-1}\sin\theta\xi^1 - s^{-1}\sin\theta\xi^2 + (\cos^2\theta - s^{-1}y\sin\theta)\xi^3 + \sin\theta(\cos\theta + s^{-1}x)\xi^4$, and $-s^{-1}\cos\theta\xi^1 + s^{-1}\cos\theta\xi^2 + \cos\theta(\sin\theta + s^{-1}y)\xi^3 + (\sin^2\theta - s^{-1}x\cos\theta)\xi^4$ of $\mathfrak{g}^{\mathcal{H}}$ lead to the nonholonomic Noether equations $\mathbf{d}p_\phi$, $\mathbf{d}p_x$, and $\mathbf{d}p_y$, respectively. Thus, $\mathcal{D}_G = \mathcal{U} + \mathcal{V} = \operatorname{span}\{\partial_x, \partial_y, \partial_\theta, \partial_\phi\}$ is found easily because $\mathcal{D}_G$ is the kernel of $\{\mathbf{d}p_\phi, \mathbf{d}p_x, \mathbf{d}p_y\}$. This is obviously integrable. The induced Dirac structure on a leaf $f^{-1}(a,b,c)$ (where f is the projection on (p_ϕ, p_x, p_y)) of $\mathcal{D}_G$ is given by

$$D_{f^{-1}(a,b,c)} = \operatorname{span}\{(\cos\theta\partial_x + \sin\theta\partial_y, 0), (\partial_\theta, 0), (\partial_\phi, 0), (0, \sin\theta\mathbf{d}x - \cos\theta\mathbf{d}y)\}.$$

Here the reduced space $M_{(a,b,c)}$ is a single point. The reduced Dirac structure is hence trivial, as can also be seen from the formula

$$\frac{(D_{f^{-1}(a,b,c)} \cap \mathcal{K}^{\perp}_{(a,b,c)}) + \mathcal{K}_{(a,b,c)}}{\mathcal{K}_{(a,b,c)}} \Big/ G_{(a,b,c)}.$$

Finally, note that in the last two examples, we have $\mathcal{R} = \{0\}$ and hence $\mathcal{R}^\circ \cap \mathcal{V} = \mathcal{V}$. This is why we get in the first of the two examples the three constants of motion p_x, p_y, and p_θ belonging to the three elements ξ^2, ξ^3, and ξ^1 of $\mathfrak{g}$ and in the second example the constants p_x, p_y, p_θ, and p_ϕ belonging to the four elements ξ^3, ξ^4, ξ^2, and ξ^1 of $\mathfrak{g}$. Note that in this case, the constancy of p_ϕ follows already from the existence of the constant section ξ^1 of $\mathfrak{g}^{\mathcal{H}}$.

Like in the previous example, the other symmetry groups of the system (the "θ-symmetry" $\mathbb{S}^1$, "the ϕ-symmetry" $\mathbb{S}^1$, $\mathbb{S}^1\times\mathbb{S}^1$, $\mathbb{S}^1\times\mathbb{R}^2$, SE(2)) are not interesting for the method of reduction presented in this section.

4.5.4. The Heisenberg particle

At last, we present an example where the reduced form is not closed. It can be found in [6]. The configuration space Q is $\mathbb{R}^3$ with coordinates (x,y,z) subject to the constraint $\dot z = y\dot x - x\dot y$. The Lagrangian on TQ is given by $L(x,y,z,\dot x,\dot y,\dot z) = \frac{1}{2}(\dot x^2 + \dot y^2 + \dot z^2)$ and hence the Legendre transformation yields

$$p_x = \dot x, \qquad p_y = \dot y, \qquad p_z = \dot z.$$

For (x,y,z,p_x,p_y,p_z), we have $p_z = yp_x - xp_y$. Hence, we have the global coordinates (x,y,z,p_x,p_y) for M and the 2-form ω_M is given by

$$\begin{aligned}\omega_M &= \mathbf{d}x\wedge\mathbf{d}p_x + \mathbf{d}y\wedge\mathbf{d}p_y + \mathbf{d}z\wedge\mathbf{d}(yp_x - xp_y)\\ &= \mathbf{d}x\wedge(\mathbf{d}p_x + p_y\mathbf{d}z) + \mathbf{d}y\wedge(\mathbf{d}p_y - p_x\mathbf{d}z) + \mathbf{d}z\wedge(y\mathbf{d}p_x - x\mathbf{d}p_y).\end{aligned}$$

The vector bundle $\mathcal{H}$ is given by $\mathcal{H} = \ker\{\mathbf{d}z - y\mathbf{d}x + x\mathbf{d}y\} = \operatorname{span}\{y\partial_z + \partial_x, x\partial_z - \partial_y, \partial_{p_x}, \partial_{p_y}\}$ and we compute

$$\begin{aligned}
\mathbf{i}_{y\partial_z+\partial_x}\omega_M &= y(y\mathbf{d}p_x - x\mathbf{d}p_y) - yp_y\mathbf{d}x + yp_x\mathbf{d}y + (\mathbf{d}p_x + p_y\mathbf{d}z),\\
\mathbf{i}_{x\partial_z-\partial_y}\omega_M &= x(y\mathbf{d}p_x - x\mathbf{d}p_y) - xp_y\mathbf{d}x + xp_x\mathbf{d}y - (\mathbf{d}p_y - p_x\mathbf{d}z),\\
\mathbf{i}_{\partial_{p_x}}\omega_M &= -\mathbf{d}x - y\mathbf{d}z,\\
\mathbf{i}_{\partial_{p_y}}\omega_M &= -\mathbf{d}y + x\mathbf{d}z.
\end{aligned}$$

Thus, we get the smooth global spanning sections

$$\begin{aligned}
&\big\{\big(y\partial_z + \partial_x, (y^2+1)\mathbf{d}p_x - xy\mathbf{d}p_y - yp_y\mathbf{d}x + yp_x\mathbf{d}y + p_y\mathbf{d}z\big), \big(\partial_{p_x}, -\mathbf{d}x - y\mathbf{d}z\big),\\
&\quad \big(x\partial_z - \partial_y, -(x^2+1)\mathbf{d}p_y + xy\mathbf{d}p_x - xp_y\mathbf{d}x + xp_x\mathbf{d}y + p_x\mathbf{d}z\big), \big(\partial_{p_y}, -\mathbf{d}y + x\mathbf{d}z\big),\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (0, \mathbf{d}z - y\mathbf{d}x + x\mathbf{d}y)\big\}
\end{aligned}$$

for the Dirac structure D.

Consider the action $\phi : \mathbb{R} \times Q \to Q$ of the Lie group $G = \mathbb{R}$ on Q, given by $\phi(r, x, y, z) := (x, y, z + r)$. This action obviously leaves the Lagrangian and the constraints invariant. The induced action $\Phi : G \times M \to M$ is given by $\Phi(r, x, y, z, p_x, p_y) := (x, y, z + r, p_x, p_y)$ and hence the vertical bundle in this example equals $\mathcal{V} = \operatorname{span}\{\partial_z\}$. We get $\mathcal{H} \cap \mathcal{V} = \{0\}$ and hence $\mathcal{U} = \mathcal{H}$. The two methods of reduction (in Section 3.2 and Section 4.4) lead in this case to the same result since the distribution $\mathcal{D}_G = \mathcal{U} + \mathcal{V} = \mathcal{H} + \mathcal{V} = TM$ is trivially integrable with M as single leaf. The reduced Dirac structure D_{red} on $\bar{M}$ with coordinates (x, y, p_x, p_y) is thus given by

$$\begin{aligned}
D_{\text{red}} &= \left.\frac{(D \cap \mathcal{K}^{\perp}) + \mathcal{K}}{\mathcal{K}}\right/ G\\
&= \operatorname{span}\big\{\big(\partial_x, (y^2+1)\mathbf{d}p_x - xy\mathbf{d}p_y - yp_y\mathbf{d}x + yp_x\mathbf{d}y + p_y(y\mathbf{d}x - x\mathbf{d}y)\big),\\
&\qquad \big(\partial_{p_x}, -\mathbf{d}x - y(y\mathbf{d}x - x\mathbf{d}y)\big),\\
&\qquad \big(-\partial_y, -(x^2+1)\mathbf{d}p_y + xy\mathbf{d}p_x - xp_y\mathbf{d}x + xp_x\mathbf{d}y + p_x(y\mathbf{d}x - x\mathbf{d}y)\big),\\
&\qquad \big(\partial_{p_y}, -\mathbf{d}y + x(y\mathbf{d}x - x\mathbf{d}y)\big)\big\}\\
&= \operatorname{span}\big\{\big(\partial_x, (y^2+1)\mathbf{d}p_x - xy\mathbf{d}p_y + (yp_x - xp_y)\mathbf{d}y\big),\\
&\qquad \big(\partial_{p_x}, -(1+y^2)\mathbf{d}x + xy\mathbf{d}y\big),\\
&\qquad \big(-\partial_y, -(x^2+1)\mathbf{d}p_y + xy\mathbf{d}p_x + (yp_x - xp_y)\mathbf{d}x\big),\\
&\qquad \big(\partial_{p_y}, -(1+x^2)\mathbf{d}y + xy\mathbf{d}x\big)\big\}.
\end{aligned}$$

Note that this is the graph of the 2-form

$$\omega_{\text{red}} = (1+y^2)\mathbf{d}x\wedge\mathbf{d}p_x + (1+x^2)\mathbf{d}y\wedge\mathbf{d}p_y + (yp_x - xp_y)\mathbf{d}x\wedge\mathbf{d}y - xy(\mathbf{d}x\wedge\mathbf{d}p_y + \mathbf{d}y\wedge\mathbf{d}p_x).$$

A direct computation shows that the determinant of ω_{red} equals $(1 + x^2 + y^2)^2 \neq 0$

on $\bar{M}$ which shows that the form $\omega_{\rm red}$ is nondegenerate. The equalities

$$\mathbf{d}\omega_{\rm red}(\partial_x, \partial_y, \partial_{p_x}) = -2y \qquad \text{and} \qquad \mathbf{d}\omega_{\rm red}(\partial_x, \partial_y, \partial_{p_y}) = 2x$$

show that $\omega_{\rm red}$ is not closed.

Note also that in this example we have $\mathcal{R} = \mathcal{H}^\circ$ and hence $\mathcal{R}^\circ \cap \mathcal{V} = \mathcal{H} \cap \mathcal{V} = \{0\}$.

REFERENCES

[1] V. I. Arnol′d, V. V. Kozlov and A. I. Neĭshtadt: *Dynamical Systems. III*, volume 3 of *Encyclopaedia of Mathematical Sciences*. Springer, Berlin 1988. Translated from the Russian by A. Iacob.

[2] L. Bates and J. Śniatycki: Nonholonomic reduction, *Rep. Math. Phys.* **32** (1993), 99.

[3] G. Blankenstein: *Implicit Hamiltonian Systems: Symmetry and Interconnection*, Ph.D.Thesis, University of Twente, 2000.

[4] G. Blankenstein and T.S. Ratiu: Singular reduction of implicit Hamiltonian systems, *Rep. Math. Phys.* **53** (2004), 211.

[5] G. Blankenstein and A. J. van der Schaft; Symmetry and reduction in implicit generalized Hamiltonian systems, *Rep. Math. Phys.* **47** (2001), 57.

[6] A. M. Bloch: *Nonholonomic Mechanics and Control*, volume 24 of *Interdisciplinary Applied Mathematics*. Springer, New York 2003. With the collaboration of J. Baillieul, P. Crouch and J. Marsden, With scientific input from P. S. Krishnaprasad, R. M. Murray and D. Zenkov, *Systems and Control.*

[7] H. Bursztyn, G. R. Cavalcanti and M. Gualtieri: Reduction of Courant algebroids and generalized complex structures, *Adv. Math.* **211** (2007), 726.

[8] H. Bursztyn and M. Crainic: Dirac structures, momentum maps, and quasi-Poisson manifolds. In *The breadth of symplectic and Poisson geometry*, volume 232 of *Progr. Math.*, pages 1–40. Birkhäuser Boston, Boston, MA 2005.

[9] H. Bursztyn, M. Crainic, A. Weinstein and C. Zhu: Integration of twisted Dirac brackets, *Duke Math. J.* **123** (2004), 549.

[10] F. Cantrijn, D. M. de Diego, M. de León and J. C. Marrero: Reduction of nonholonomic mechanical systems with symmetries, *Rep. Math. Phys.* **42** (1998), 25. Pacific Institute of Mathematical Sciences Workshop on Nonholonomic Constraints in Dynamics (Calgary 1997).

[11] H. Cendra, J. E. Marsden, T.S. Ratiu and H. Yoshimura: In preparation, 2008.

[12] T. J. Courant: Dirac manifolds, *Trans. Am. Math. Soc.* **319** (1990), 631.

[13] T. J. Courant and A. Weinstein: Beyond Poisson structures. In *Action hamiltoniennes de groupes. Troisième théorème de Lie (Lyon 1986)*, volume 27 of *Travaux en Cours*, pages 39–49. Hermann, Paris 1988.

[14] R. Cushman, D. Kemppainen, J. Śniatycki and L. Bates: Geometry of nonholonomic constraints. In *Proceedings of the XXVII Symposium on Mathematical Physics (Toruń 1994)*, volume 36, pages 275–286 (1995).

[15] F. Fassò, A. Ramos and N. Sansonetto: The reaction-annihilator distribution and the nonholonomic Noether theorem for lifted actions, *Regul. Chaotic Dyn.* **12** (2007), 579.

[16] M. Jotz, T. Ratiu and M. Zambon: Invariant frames for vector bundles and applications, To appear in *Geometriae Dedicata* (2011).

[17] M. Jotz and T.S. Ratiu: Optimal Dirac reduction, To appear in *International Mathematics Research Notices* (2011).

[18] M. Jotz and T.S. Ratiu: Induced Dirac structure on isotropy type manifolds, *Transform. Groups* **16** (2011), 175.

[19] M. Jotz, T.S. Ratiu and J. Śniatycki: Singular Dirac reduction, *Trans. Amer. Math. Soc.* **363** (2011), 2967.

[20] V. V. Kozlov and N. N. Kolesnikov: On theorems of dynamics, *J. Appl. Math. Mech.* **42** (1978), 28.

[21] P. Libermann and C.-M. Marle: *Symplectic Geometry and Analytical Mechanics. Transl. from the French by Bertram Eugene Schwarzbach.* Mathematics and its Applications, 35. Dordrecht etc.: D. Reidel Publishing Company, a member of the Kluwer Academic Publishers Group. XVI, 526 p. , 1987.

[22] Z.-J. Liu, A. Weinstein and P. Xu: Manin triples for Lie bialgebroids, *J. Differential Geom.* **45** (1997), 547.

[23] J. E. Marsden and T.S. Ratiu: *Introduction to Mechanics and Symmetry. A Basic Exposition of Classical Mechanical Systems.* 2nd ed. Texts in Applied Mathematics. 17. New York, NY: Springer. xviii, 582 p. , 1999.

[24] J.-P. Ortega and T.S. Ratiu: *Momentum Maps and Hamiltonian Reduction,* Progress in Mathematics (Boston, Mass.) 222. Boston, MA: Birkhäuser. xxxiv, 497 p., 2004.

[25] M. J. Pflaum: *Analytic and Geometric Study of Stratified Spaces*, volume 1768 of *Lecture Notes in Mathematics*. Springer, Berlin 2001.

[26] R. M. Rosenberg: *Analytical Dynamics of Discrete Systems*, volume 4 of *Mathematical Concepts and Methods in Science and Engineering*. Plenum Press, New York 1977.

[27] G.W. Schwarz: Lifting smooth homotopies of orbit spaces, *Inst. Hautes Études Sci. Publ. Math.* **51** (1980), 37.

[28] P. Ševera and A. Weinstein: Poisson geometry with a 3-form background, *Progr. Theoret. Phys. Suppl.* **144** (2001), 145. Noncommutative geometry and string theory (Yokohama 2001).

[29] P. Stefan: Accessibility and foliations with singularities, *Bull. Amer. Math. Soc.* **80** (1974), 1142.

[30] P. Stefan: Accessible sets, orbits, and foliations with singularities, *Proc. London Math. Soc.* (3) **29** (1974), 699.

[31] P. Stefan: Integrability of systems of vector fields, *J. London Math. Soc.* (2) **21** (1980), 544.

[32] M. Stiénon and P. Xu: Reduction of generalized complex structures, *J. Geom. Phys.* **58** (2008), 105.

[33] H. J. Sussmann: Orbits of families of vector fields and integrability of distributions, *Trans. Amer. Math. Soc.* **180** (1973), 171.

[34] I. Vaisman: *Lectures on the Geometry of Poisson Manifolds*, volume 118 of *Progress in Mathematics*. Birkhäuser, Basel 1994.

[35] H. Yoshimura and J. E. Marsden: Dirac structures in Lagrangian mechanics. I. Implicit Lagrangian systems, *J. Geom. Phys.* **57** (2006), 133.

[36] H. Yoshimura and J. E. Marsden: Reduction of Dirac structures and the Hamilton-Pontryagin principle, *Rep. Math. Phys.* **60** (2007), 381.

[37] H. Yoshimura and J. E. Marsden: Dirac cotangent bundle reduction, *J. Geom. Mech.* **1** (2009), 87.