Faculty of Mathematical Sciences

University of Twente

University for Technical and Social Sciences

P.O. Box 217
7500 AE Enschede
The Netherlands
Phone: +31-53-4893400
Fax: +31-53-4893114
Email: memo@math.utwente.nl

Memorandum No. 1489

Symmetry and reduction in implicit generalized Hamiltonian systems

G. Blankenstein and A.J. van der Schaft

June 1999

ISSN 0169-2690

Symmetry and Reduction in Implicit Generalized Hamiltonian Systems

G. Blankenstein and A.J. van der Schaft

June 8, 1999

Abstract

In this paper the notion of symmetry for implicit generalized Hamiltonian systems will be studied and a reduction theorem, generalizing the 'classical' reduction theorems of symplectic and Poisson Hamiltonian systems, will be derived.

1991 Mathematics Subject Classification: 70H33, 58F05, 70H05, 70F25, 34A09, 34C20, 34A05

Keywords: constraints, Dirac structures, Hamiltonian systems, implicit systems, reduction, symmetry

1 Introduction

The notions of symmetry, conserved quantity and reduction belong to the most important tools in the study of mechanical systems. Noether's theorem for the Euler-Lagrange or Hamiltonian equations states that there is a one-to-one correspondence between continuous *symmetries* of the system and *conserved quantities*. This leads to the observation that the existence of a *one*-dimensional symmetry action implies the possibility of *reducing* the equations by *two* dimensions; namely by restricting to the level sets of the conserved quantity and then factoring out the symmetry, or, as can be shown to lead to the same result, first factoring out the symmetry and then restricting to the level sets.
Classically this theory of reduction by symmetry has been very important in the actual *solving* of Euler-Lagrange or Hamiltonian equations, but also turned out to be an indispensable tool in the *stability analysis*, see e.g. [13, 1]. Also for *simulation* these notions have proved to be valuable since reliable numerical integration routines ideally should respect the conserved quantities and symmetries. On the other hand, modular or 'object-oriented' modeling of (electro-)mechanical systems almost invariably leads, at least in first instance, to *mixed* sets of differential and *algebraic* equations (DAE's), and in many situations one would prefer *not* to eliminate the constraints and reduce the system to an ordinary Hamiltonian system (if this is possible at all!). In previous work, see e.g. [15, 16, 10, 4], it has been shown how the underlying Hamiltonian structure of DAE's can be made explicit using the geometric notion of a Dirac structure. The question then comes up if, and how, the tools of symmetries and reduction as used for ordinary, explicit, Hamiltonian equations can be extended to such Hamiltonian DAE's. This should be equally important for their stability analysis and should also have important consequences for the choice of integration routines of such DAE's and their properties.
In this paper we will investigate the notion of symmetries of implicit generalized Hamiltonian systems. We continue up on the results in [14]. Furthermore we will investigate the reduction possibilities of implicit generalized Hamiltonian system and prove the analog of the 'classical' reduction theorems of symplectic and Poisson Hamiltonian systems in [1, 7, 9, 12]. The paper is organized as follows.

In section 2 we will give an introduction to Dirac structures and implicit generalized Hamiltonian systems. In section 3 we will investigate the notion of symmetries of an implicit generalized Hamiltonian system. We will state some important results obtained in [14] and will derive some new ones. Furthermore we will introduce the notion of first integrals (or conserved quantities) and Casimir functions which are important for the reduction process described in sections 4 and 5. In section 4 we will derive the basic results on reduction of Dirac structures and implicit generalized Hamiltonian systems. We will combine these results in section 5 to derive our main result on reduction of implicit generalized Hamiltonian systems. This result will generalize the 'classical' reduction theorems of explicit Hamiltonian systems described in [1, 7, 9, 12]. In section 6 we will take a closer look at a specific Casimir function introduced in section 5. Finally, in section 7 we will specialize the main reduction result of section 5 to implicit generalized Hamiltonian systems satisfying an additional regularity assumption on the constraints, which makes these systems explicit in some sense. We will compare the reduction result in this case with the 'classical' explicit case. Conclusions are given in section 8.

2 Implicit generalized Hamiltonian systems

In this section we will give an introduction to Dirac structures and implicit generalized Hamiltonian systems. For more information we refer to [16, 10, 14, 3, 5]. Let $\mathcal{X}$ be an n-dimensional manifold with tangent bundle $T\mathcal{X}$ and cotangent bundle $T^*\mathcal{X}$. Define $T\mathcal{X} \oplus T^*\mathcal{X}$ as the smooth vector bundle over $\mathcal{X}$ with fiber at each $x \in \mathcal{X}$ given by $T_x\mathcal{X} \times T_x^*\mathcal{X}$. Let X be a smooth vector field and α a smooth one-form on $\mathcal{X}$ respectively. We say that the pair (X, α) belongs to a subspace $D \subset T\mathcal{X} \oplus T^*\mathcal{X}$, denoted $(X, \alpha) \in D$, if $(X(x), \alpha(x)) \in D(x)$, $\forall x \in \mathcal{X}$. Let D be a linear subspace of $T\mathcal{X} \oplus T^*\mathcal{X}$, that is, $(X, \alpha), (Y, \beta) \in D$ implies $h_1(X, \alpha) + h_2(Y, \beta) \in D$ for all $h_1, h_2 \in C^\infty(\mathcal{X})$. Define the linear subspace $D^\perp$ as follows

$$D^\perp = \{(Y, \beta) \in T\mathcal{X} \oplus T^*\mathcal{X} \mid \langle \alpha, Y \rangle + \langle \beta, X \rangle = 0, \ \forall \ (X, \alpha) \in D\},$$

where $\langle \cdot, \cdot \rangle$ denotes the natural pairing between a one-form and a vector field on $\mathcal{X}$.

Definition 1. *[5] A generalized Dirac structure on $\mathcal{X}$ is a linear subspace $D \subset T\mathcal{X} \oplus T^*\mathcal{X}$ such that $D = D^\perp$.*

Remark 1. In definition 1, $T\mathcal{X}$, respectively $T^*\mathcal{X}$, is identified with the set of smooth vector fields, respectively one-forms, on $\mathcal{X}$. So a Dirac structure is a set of pairs (X, α), with X a smooth vector field and α a smooth one-form on $\mathcal{X}$, such that D is linear and $D = D^\perp$.

From the condition $D = D^\perp$ it follows that D is constant dimensional, with $\dim D(x) = n$, $\forall x \in \mathcal{X}$, see also [5].

Proposition 1. *Let D be a generalized Dirac structure on an n-dimensional manifold $\mathcal{X}$. Then D is constant dimensional with $\dim D(x) = n$, $\forall x \in \mathcal{X}$.*

Proof. (i) Assume that $\dim D(x_0) = p > n$ for some $x_0 \in \mathcal{X}$. Because D is smooth then also $\dim D(x) \geq p > n$, $\forall x \in U$, where $U \subset \mathcal{X}$ is some neighborhood of x_0. But this implies that $\dim D^\perp(x) < n$, $\forall x \in U$, which contradicts the fact that $D(x) = D^\perp(x)$.
(ii) Assume that $\dim D(x) < n$, $\forall x \in U$, for some open subset $U \subset \mathcal{X}$. Then there is an open subset $\tilde{U} \subset U$ such that $\dim D^\perp(x) > n$, $\forall x \in \tilde{U}$. This contradicts the fact that $D(x) = D^\perp(x)$.
(iii) Finally, assume that $\dim D(x) < n$, $\forall x \in A$, where $A \subset \mathcal{X}$ contains no interior points (i.e. A does not contain any open subset of $\mathcal{X}$), and there exists an open subset $U \subset \mathcal{X}$, with $U \cap A$ nonempty, such that $\dim D(x) = n$, $\forall x \in U \setminus A$. Note that this is the only case still remaining from

(i) and (ii). Then $\dim D^{\perp}(x) = \dim D(x) = n$, $\forall x \in U \setminus A$ (because $D^{\perp}(x) = D(x)$). Now, if D looses dimension on A then $D^{\perp}$ can only gain dimension on A, so $\dim D(x) \geq n$, $\forall x \in A$. Actually, from the fact that $D^{\perp}$ is smooth it follows that $\dim D^{\perp}(x) = n$, $\forall x \in A$. This however contradicts the fact that $D(x) = D^{\perp}(x)$ (since $\dim D(x) < n$, $\forall x \in A$). This ends the proof. $\square$

From proposition 1 it follows that D is a subbundle of $T\mathcal{X} \oplus T^*\mathcal{X}$. Since $D = D^{\perp}$ it immediately follows that for every pair $(X, \alpha) \in D$

$$\langle \alpha, X \rangle = 0. \tag{1}$$

Proposition 1 has the following obvious but important consequence.

Proposition 2. *Let D be a generalized Dirac structure. Then $D^{\perp}(x) = [D(x)]^{\perp}$, $\forall x \in \mathcal{X}$. Here $[D(x)]^{\perp}$ means the pointwise perpendicular to $D(x)$, i.e.*

$$[D(x)]^{\perp} = \{(w, w^*) \in T_x\mathcal{X} \times T_x^*\mathcal{X} \mid \langle v^*, w \rangle + \langle w^*, v \rangle = 0,\ \forall (v, v^*) \in D(x)\}.$$

Proof. It immediately follows that $D^{\perp}(x) \subset [D(x)]^{\perp}$. Both $D^{\perp}(x)$ and $[D(x)]^{\perp}$ are linear (over $\mathbb{R}$) subspaces of $T_x\mathcal{X} \times T_x^*\mathcal{X}$. Furthermore, since $\dim D(x) = n$ ($= \dim D^{\perp}(x)$ since $D = D^{\perp}$) it follows that $\dim[D(x)]^{\perp} = n$. This implies that $D^{\perp}(x) = [D(x)]^{\perp}$. $\square$

A generalized Dirac structure is called closed, or just a Dirac structure, if the following condition holds.

Definition 2. *A generalized Dirac structure D on an n-dimensional manifold $\mathcal{X}$ is called closed if*

$$\langle L_{X_1}\alpha_2, X_3 \rangle + \langle L_{X_2}\alpha_3, X_1 \rangle + \langle L_{X_3}\alpha_1, X_2 \rangle = 0,$$

for all pairs $(X_1, \alpha_1), (X_2, \alpha_2)$ and (X_3, α_3) in D.

Here $L_X\alpha$ denotes the Lie derivative of a one-form α with respect to a vector field X. We have the following theorem.

Theorem 3. *[5, 3, 4] A generalized Dirac structure D on $\mathcal{X}$ is closed if and only if*

$$([X_1, X_2], i_{X_1}d\alpha_2 - i_{X_2}d\alpha_1 + d\langle \alpha_2, X_1 \rangle) \in D, \quad \forall\ (X_1, \alpha_1), (X_2, \alpha_2) \in D.$$

Example 1. Let ω be a nondegenerate two-form on $\mathcal{X}$, then

$$D = \{(X, \alpha) \in T\mathcal{X} \oplus T^*\mathcal{X} \mid \alpha = i_X\omega\}$$

is a generalized Dirac structure on $\mathcal{X}$. D is closed if and only if $d\omega = 0$. This corresponds to a symplectic structure $(\mathcal{X}, \omega)$.

Example 2. Let $J(x) : T_x^*\mathcal{X} \to T_x\mathcal{X},\ x \in \mathcal{X}$, be a skew-symmetric vector bundle map, then

$$D = \{(X, \alpha) \in T\mathcal{X} \oplus T^*\mathcal{X} \mid X(x) = J(x)\alpha(x), \quad \forall x \in \mathcal{X}\}$$

is a generalized Dirac structure on $\mathcal{X}$. This corresponds to a Poisson structure $(\mathcal{X}, \{\cdot, \cdot\})$, where $J(x)$ is the structure matrix of the Poisson bracket $\{\cdot, \cdot\}$. D is closed if and only if the bracket satisfies the Jacobi identity.

Examples 1 and 2 show that the notion of a (generalized) Dirac structure is a generalization of the classical symplectic and Poisson structures.

Corresponding to a generalized Dirac structure D on $\mathcal{X}$ we define the following (co-)distributions

$$\begin{aligned}\mathsf{G}_0 &= \{X \in T\mathcal{X} \mid (X,0) \in D\},\\ \mathsf{G}_1 &= \{X \in T\mathcal{X} \mid \exists\, \alpha \in T^*\mathcal{X} \text{ such that } (X,\alpha) \in D\},\\ \mathsf{P}_0 &= \{\alpha \in T^*\mathcal{X} \mid (0,\alpha) \in D\},\\ \mathsf{P}_1 &= \{\alpha \in T^*\mathcal{X} \mid \exists\, X \in T\mathcal{X} \text{ such that } (X,\alpha) \in D\}.\end{aligned}$$

Define the annihilator of a smooth distribution $L \subset T\mathcal{X}$ as the smooth codistribution

$$\text{ann } L = \{\alpha \in T^*\mathcal{X} \mid \langle \alpha, X\rangle = 0,\ \forall\, X \in L\},$$

and the kernel of a smooth codistribution $K \subset T^*\mathcal{X}$ as the smooth distribution

$$\ker K = \{X \in T\mathcal{X} \mid \langle \alpha, X\rangle = 0,\ \forall\, \alpha \in K\}.$$

It follows that by definition $\mathsf{G}_0 = \ker \mathsf{P}_1$ and $\mathsf{P}_0 = \text{ann } \mathsf{G}_1$. Furthermore, we have that $\mathsf{P}_1 \subset \text{ann } \mathsf{G}_0$ and $\mathsf{G}_1 \subset \ker \mathsf{P}_0$, with equality if and only if P_1, respectively G_1, is constant dimensional [4]. From theorem 3 it follows that $\mathsf{G}_0, \mathsf{G}_1$ and P_1 are involutive if D is closed (if G_1 is constant dimensional it follows that also P_0 is involutive).

We have the following two important representations of a generalized Dirac structure.

Theorem 4. *[4] Let D be a generalized Dirac structure on a manifold $\mathcal{X}$.*

(a) If G_1 is constant dimensional, then there exists a skew-symmetric linear map $\omega(x) : \mathsf{G}_1(x) \subset T_x\mathcal{X} \to (\mathsf{G}_1(x))^* \subset T_x^*\mathcal{X},\ x \in \mathcal{X}$, with kernel G_0, such that

$$D = \{(X,\alpha) \in T\mathcal{X} \oplus T^*\mathcal{X} \mid \alpha(x) - \omega(x)X(x) \in \text{ann } \mathsf{G}_1(x),\ \forall x \in \mathcal{X},\ X \in \mathsf{G}_1\}. \tag{2}$$

(b) If P_1 is constant dimensional, then there exists a skew-symmetric linear map $J(x) : \mathsf{P}_1(x) \subset T_x^*\mathcal{X} \to (\mathsf{P}_1(x))^* \subset T_x\mathcal{X},\ x \in \mathcal{X}$, with kernel P_0, such that

$$D = \{(X,\alpha) \in T\mathcal{X} \oplus T^*\mathcal{X} \mid X(x) - J(x)\alpha(x) \in \ker \mathsf{P}_1(x),\ \forall x \in \mathcal{X},\ \alpha \in \mathsf{P}_1\}. \tag{3}$$

Conversely, if D is defined as in (2) for some skew-symmetric linear map $\omega(x) : T_x\mathcal{X} \to T_x^*\mathcal{X},\ x \in \mathcal{X}$, and constant dimensional distribution $\mathsf{G}_1 \subset T\mathcal{X}$, respectively if D is defined as in (3) for some skew-symmetric linear map $J(x) : T_x^*\mathcal{X} \to T_x\mathcal{X},\ x \in \mathcal{X}$, and constant dimensional codistribution $\mathsf{P}_1 \subset T^*\mathcal{X}$, then D is a generalized Dirac structure on $\mathcal{X}$.

Note that if $\mathsf{G}_1 = T\mathcal{X}$ and $\mathsf{G}_0 = 0$, then we are in the situation of example 1, whereas if $\mathsf{P}_1 = T^*\mathcal{X}$, then we are in the situation of example 2.

The set of *admissible* functions corresponding to a generalized Dirac structure D is defined as

$$\mathcal{A}_D = \{H \in C^\infty(\mathcal{X}) \mid dH \in \mathsf{P}_1\}.$$

There is a well defined generalized Poisson bracket on $\mathcal{A}_D$ given by [4]

$$\{H_1, H_2\}_D = \langle dH_1, X_2\rangle = -\langle dH_2, X_1\rangle,$$

where $H_1, H_2 \in \mathcal{A}_D$, i.e. $(X_1, dH_1), (X_2, dH_2) \in D$.

Now we will define the notion of an implicit generalized Hamiltonian system.

Definition 3. *[4] Let D be a (generalized) Dirac structure on a manifold $\mathcal{X}$. Let $H \in C^\infty(\mathcal{X})$ be a smooth function on $\mathcal{X}$, called the Hamiltonian or energy function. Then the implicit (generalized) Hamiltonian system corresponding to $(\mathcal{X}, D, H)$ is defined by the specification*

$$(\dot{x}, dH(x)) \in D(x), \quad x \in \mathcal{X}.$$

Usually we will use the terminology *implicit (generalized) Hamiltonian system* $(\mathcal{X}, D, H)$, by which we mean the implicit (generalized) Hamiltonian system corresponding to $(\mathcal{X}, D, H)$ as defined in definition 3.

Example 3. Consider the generalized Dirac structure in example 1, then the corresponding implicit generalized Hamiltonian system is precisely the classical Hamiltonian system defined by the two-form ω

$$dH = \omega(X_H, \cdot), \tag{4}$$

where X_H is the vector field corresponding to the solution $x(t)$, i.e. $\dot{x} = X_H(x)$. D is closed if and only if there exist local coordinates (q, p) for x for which the system (4) for an arbitrary Hamiltonian H takes the form

$$\dot{q} = \frac{\partial H}{\partial p}(q, p), \quad \dot{p} = -\frac{\partial H}{\partial q}(q, p),$$

which are just the classical canonical Hamiltonian equations.

Example 4. Consider the generalized Dirac structure in example 2, then the corresponding implicit generalized Hamiltonian system is given by

$$\dot{x} = J(x)\frac{\partial H}{\partial x}(x). \tag{5}$$

This is precisely the classical Hamiltonian dynamics given by the Poisson bracket, i.e. $\dot{x} = \{x, H\}$. Again, D is closed if and only if there exist local coordinates (q, p, r) for x for which (5) for an arbitrary Hamiltonian H takes the form

$$\dot{q} = \frac{\partial H}{\partial p}(q, p, r), \ \dot{p} = -\frac{\partial H}{\partial q}(q, p, r), \ \dot{r} = 0.$$

Let us reflect on definition 3 a bit more. First we will define the concept of a *solution* of the implicit (generalized) Hamiltonian system $(\mathcal{X}, D, H)$.

Definition 4. *A solution of the implicit (generalized) Hamiltonian system $(\mathcal{X}, D, H)$ is defined as a smooth time function $x : I \subset \mathbb{R} \to \mathcal{X}$ such that*

$$(X_H, dH)(x(t)) \in D(x(t)), \quad \forall t \in I,$$

where $X_H(x(t)) = \dot{x}(t)$, $\forall t \in I$, and where I is the interval of existence of $x(t)$, i.e. the domain of x.

By (1) it follows that we have the usual invariance of the Hamiltonian, or conservation of energy, along solutions

$$\frac{dH}{dt}(x(t)) = \langle dH(x(t)), X_H(x(t))\rangle = 0, \quad \forall t \in I.$$

In general, the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ defines a mixed set of differential and algebraic equations (DAE's). Take for instance the Dirac structure given in (3). The corresponding implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$, for any $H \in C^\infty(\mathcal{X})$, is given by

$$\begin{aligned} \dot{x} &= J(x)\frac{\partial H}{\partial x}(x) + g(x)\lambda, && (6) \\ 0 &= g^T(x)\frac{\partial H}{\partial x}(x), && (7) \end{aligned}$$

where $g(x)$ is any full rank matrix such that $\mathrm{Im}\ g(x) = \mathsf{G}_0(x) = \ker \mathsf{P}_1(x)$. (6,7) defines a set of DAE's, where the algebraic equations are given by (7). The variables λ can be seen as Lagrange multipliers, required to keep the constraint equations (7) to be satisfied for all time. In [4] it is shown that (6,7) can be used to describe a mechanical system with kinematic constraints (the corresponding Dirac structure is closed if and only if the constraints are holonomic [4]). In that case, λ can be interpreted as the constraint forces.

In general, define the *constraint manifold* (corresponding to an implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$)

$$\mathcal{X}_c = \{x \in \mathcal{X} \mid dH(x) \in \mathsf{P}_1(x)\}.$$

Then it follows that every solution $x(t)$ of $(\mathcal{X}, D, H)$ necessarily is contained in $\mathcal{X}_c$. Notice that not through every point of $\mathcal{X}_c$ there has to go a solution of $(\mathcal{X}, D, H)$. Also notice that in general the solutions of $(\mathcal{X}, D, H)$ are not unique. This happens for instance if the Lagrange multipliers λ in (6,7) are not uniquely determined. If λ *is* uniquely determined, then the solutions of $(\mathcal{X}, D, H)$ *are* unique. This is the case when the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ satisfies assumption 5 (see next). In that case there goes through *every* point $x_c \in \mathcal{X}_c$ a *unique* solution $x(t)$ of $(\mathcal{X}, D, H)$, see proposition 6.

An implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ can be reduced to an *explicit* generalized Hamiltonian system on $\mathcal{X}_c$ provided the following assumption is satisfied.

Assumption 5. Consider the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$, with D a generalized Dirac structure on $\mathcal{X}$. Assume that P_1 is constant dimensional, so that D can be represented as in theorem 4b. Let $\mathsf{G}_0(x) = \mathrm{Im}\ g(x) = \mathrm{span}\ \{g_1(x), \ldots, g_m(x)\}$, with $g_1, \ldots, g_m$ linearly independent vector fields on $\mathcal{X}$ (note that $\mathsf{G}_0 = \ker \mathsf{P}_1$ is constant dimensional because P_1 is constant dimensional). Assume that the $m \times m$ matrix $[L_{g_i} L_{g_j} H(x)]_{i,j=1,\ldots,m}$ is invertible for all $x \in \mathcal{X}_c$.

Proposition 6. *[14] Consider the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ and let assumption 5 be satisfied. Then $(\mathcal{X}, D, H)$ reduces to an explicit generalized Hamiltonian system on $\mathcal{X}_c$, denoted by $(\mathcal{X}_c, D_c, H_c)$, given by*

$$\dot{x}_c = J_c(x_c)\frac{\partial H_c}{\partial x_c}(x_c) =: X_{H_c}(x_c), \tag{8}$$

where $x_c \in \mathcal{X}_c$, $J_c(x_c) : T^*_{x_c}\mathcal{X}_c \to T_{x_c}\mathcal{X}_c$ and $H_c : \mathcal{X}_c \to \mathbb{R}$ denotes the restriction of H to $\mathcal{X}_c$.

Proposition 6 becomes very transparent if we consider an implicit Hamiltonian system $(\mathcal{X}, D, H)$, i.e. with a generalized Dirac structure D which is closed. Then around every point $x \in \mathcal{X}$ there exist local coordinates (q,p,r,s) for which the system $(\mathcal{X},D,H)$ takes the form

$$\begin{aligned}
\dot{q} &= \frac{\partial H}{\partial p}(q,p,r,s),\\
\dot{p} &= -\frac{\partial H}{\partial q}(q,p,r,s),\\
\dot{r} &= 0,\\
0 &= \frac{\partial H}{\partial s}(q,p,r,s),
\end{aligned}$$

see [4]. Assuming that the system $(\mathcal{X},D,H)$ satisfies assumption 5 is equivalent in this case to assuming that the matrix $\frac{\partial^2 H}{\partial s^2}(q,p,r,s)$ is nonsingular. Hence by the Implicit Function Theorem we can locally express s in the coordinates q,p,r, that is, $s=s(q,p,r)$. Defining the constrained Hamiltonian $H_c(q,p,r)=H(q,p,r,s(q,p,r))$ the implicit Hamiltonian system $(\mathcal{X},D,H)$ becomes the *explicit* Hamiltonian system

$$\begin{aligned}
\dot{q} &= \frac{\partial H_c}{\partial p}(q,p,r),\\
\dot{p} &= -\frac{\partial H_c}{\partial q}(q,p,r),\\
\dot{r} &= 0.
\end{aligned}$$

3 Symmetries and first integrals

In this section we will investigate the notion of symmetry for implicit generalized Hamiltonian systems. We will recall some important results obtained in [14] and derive some new results. First, we will recall some mathematical definitions and results that we will use extensively. These can all be found e.g. in Abraham, Marsden and Ratiu [2], chapters 4 and 6. In the following all manifolds, maps, vector fields and k-forms are assumed to be smooth. M and N are manifolds. $\mathbb{X}(M)$, respectively $\mathbb{X}(N)$, is the space of vector fields on M, respectively on N. $\Omega^k(M)$, respectively $\Omega^k(N)$, is the space of k-forms on M, respectively on N.

Definition 5. *Let $\phi : M \to N$ be a diffeomorphism and $X \in \mathbb{X}(M)$ a vector field on M. Then the push-forward of X by ϕ is defined by $\phi_* X = T\phi \circ X \circ \phi^{-1} \in \mathbb{X}(N)$, where $T\phi$ is the tangent map of the map ϕ.*

If ϕ is not a diffeomorphism then the push-forward is not defined. In stead we can define the following

Definition 6. *Let $\phi : M \to N$ be a map. Two vector fields $X \in \mathbb{X}(M)$ and $Y \in \mathbb{X}(N)$ are said to be ϕ-related, denoted by $X \sim_\phi Y$, if $T\phi \circ X = Y \circ \phi$.*

The following proposition holds.

Proposition 7. *Consider a map $\phi : M \to N$, and let $X, Y \in \mathbb{X}(M)$, $\bar{X}, \bar{Y} \in \mathbb{X}(N)$ be vector fields such that $X \sim_\phi \bar{X}$ and $Y \sim_\phi \bar{Y}$. Then $[X,Y] \sim_\phi [\bar{X},\bar{Y}]$, where $[\cdot,\cdot]$ is the usual Lie bracket of vector fields.*

Now, suppose M is a submanifold of N with the corresponding inclusion map $\iota : M \to N$, and let $X \in \mathbb{X}(M)$ and $Y \in \mathbb{X}(N)$ be ι-related, $X \sim_\iota Y$, then from definition 6 it follows that at all points of $M \subset N$, Y must be tangent to M (i.e. $Y(x) = X(x)$ for all $x \in M$). Then X is called the restriction of Y to M.

Definition 7. *Let $\phi : M \to N$ be a map and $\omega \in \Omega^k(N)$ a k-form on N. Then the pull-back of ω by ϕ is defined by $\phi^*\omega = T\phi^* \circ \omega \circ \phi \in \Omega^k(M)$, where $T\phi^*$ is the adjoint of the map $T\phi$, i.e. $(\phi^*\omega)_x(v_1, \ldots, v_k) = \omega_{\phi(x)}(T_x\phi \cdot v_1, \ldots, T_x\phi \cdot v_k)$ where $v_1, \ldots, v_k \in T_xM$. Note that for the special case of a 0-form on N, i.e. a function $F : N \to \mathbb{R}$, the pull-back is defined as $\phi^*F = F \circ \phi$, which is a function on M.*

Note that ϕ need not to be a diffeomorphism for the pull-back to be defined.

Now, suppose M is a submanifold of N with the corresponding inclusion map $\iota : M \to N$, then a k-form $\omega \in \Omega^k(N)$ induces a k-form on M by $\omega_M = \iota^*\omega \in \Omega^k(M)$. We say that ω_M is the restriction of ω to M.
Now we will turn our attention to symmetries and first integrals of implicit generalized Hamiltonian systems. The notion of symmetry of a generalized Dirac structure was defined in [5].

Definition 8. *A vector field $f \in T\mathcal{X}$ is an (infinitesimal) symmetry of a generalized Dirac structure D on $\mathcal{X}$ if $(L_fX, L_f\alpha) \in D$ for all $(X, \alpha) \in D$.*

Analogously, a diffeomorphism $\phi : \mathcal{X} \to \mathcal{X}$ is called a symmetry of D if

$$(\phi_*X, (\phi^*)^{-1}\alpha) \in D \tag{9}$$

for all $(X, \alpha) \in D$ [14].

Example 5. *Consider the generalized Dirac structure given in example 1. Then $f \in T\mathcal{X}$ is a symmetry of D if and only if $L_f\omega = 0$ (see also [5], without proof).*

Proof. From the fact that ω is nondegenerate (so the matrix $\tilde{\omega}(x) : T_x\mathcal{X} \to T_x^*\mathcal{X}$, corresponding to ω, is nonsingular [1]) it follows that:

$$\mathsf{G}_0 = 0,\ \mathsf{G}_1 = T\mathcal{X},\ \mathsf{P}_0 = 0,\ \mathsf{P}_1 = T^*\mathcal{X}.$$

Since L_f is a derivation the following holds [2],p.363-364:

$$L_f(\omega(Y_1, Y_2)) = (L_f\omega)(Y_1, Y_2) + \omega(L_fY_1, Y_2) + \omega(Y_1, L_fY_2) \tag{10}$$

for all $Y_1, Y_2 \in T\mathcal{X}$.
Take $(X, \alpha) \in D$, i.e. $\alpha = \omega(X, \cdot)$, and suppose $L_f\omega = 0$. Then

$$\begin{aligned}
(L_f\alpha)(Y) &= i_YL_f\alpha = L_fi_Y\alpha - i_{[f,Y]}\alpha \\
&= L_f(\omega(X, Y)) - i_{[f,Y]}\alpha \\
&= \omega(L_fX, Y) + \omega(X, L_fY) - i_{[f,Y]}\alpha \\
&= \omega(L_fX, \cdot)(Y) + \alpha(L_fY) - i_{[f,Y]}\alpha \\
&= \omega(L_fX, \cdot)(Y)
\end{aligned}$$

for all $Y \in T\mathcal{X}$, which shows that

$$L_f\alpha = \omega(L_fX, \cdot)$$

and thus $(L_fX, L_f\alpha) \in D$. That means that f is a symmetry of D.
Conversely, let f be a symmetry of D. Then $(L_fX, L_f\alpha) \in D$, i.e. $L_f\alpha = \omega(L_fX, \cdot)$, for all $(X, \alpha) \in D$. Following the above derivation gives

$$L_f(\omega(X, Y)) = \omega(L_fX, Y) + \omega(X, L_fY)$$

for all $Y \in T\mathcal{X},\ (X, \alpha) \in D$. By (10) this implies

$$(L_f\omega)(X, Y) = 0$$

for all $Y \in T\mathcal{X},\ (X, \alpha) \in D$, that is, for all $Y \in T\mathcal{X},\ X \in \mathsf{G}_1 = T\mathcal{X}$. This implies $L_f\omega = 0$. □

Example 6. *Consider the generalized Dirac structure given in example 2. Then $f \in T\mathcal{X}$ is a symmetry of D if and only if f is canonical with respect to the Poisson bracket $\{\cdot,\cdot\}$, i.e.*

$$L_f\{H_1, H_2\} = \{L_f H_1, H_2\} + \{H_1, L_f H_2\},$$

for all $H_1, H_2 \in C^\infty(\mathcal{X})$.

Proof. We have

$$\mathsf{G}_0 = 0,\ \mathsf{G}_1(x) = \mathrm{Im}(J(x)),\ \mathsf{P}_0(x) = \ker(J(x)),\ \mathsf{P}_1 = T^*\mathcal{X}.$$

The matrix $J(x) : T_x^*\mathcal{X} \to T_x\mathcal{X},\ x \in \mathcal{X}$ defines a $(2,0)$-tensor $J : T^*\mathcal{X} \times T^*\mathcal{X} \to \mathbb{R}$. By definition

$$\{H_1, H_2\} = J(dH_1, dH_2)$$

for all $H_1, H_2 \in \mathcal{A}_D = C^\infty(\mathcal{X})$. Since L_f is a derivation [2],p.363-364:

$$\begin{aligned} L_f\{H_1, H_2\} &= L_f(J(dH_1, dH_2)) \\ &= (L_f J)(dH_1, dH_2) + J(L_f dH_1, dH_2) + J(dH_1, L_f dH_2) \\ &= (L_f J)(dH_1, dH_2) + \{L_f H_1, H_2\} + \{H_1, L_f H_2\} \end{aligned} \tag{11}$$

(use $L_f dH = dL_f H$), for all $H_1, H_2 \in C^\infty(\mathcal{X})$. Taking $H_1, H_2 = x_i, x_j$, where x_i, x_j are local coordinate functions on $\mathcal{X}$, shows that f is canonical with respect to $\{\cdot,\cdot\}$ if and only if $L_f J = 0$. Now we show that $L_f J = 0$ if and only if f is a symmetry of D.
Take $(X, \alpha) \in D$, i.e. $X = J(\alpha, \cdot)$, and suppose $L_f J = 0$. Then

$$\begin{aligned} (L_f X)[H] &= ([f, X])[H] = f[X[H]] - X[f[H]] \\ &= L_f(X[H]) - X[L_f H] \\ &= L_f(J(\alpha, dH)) - X[L_f H] \\ &= J(L_f\alpha, dH) + J(\alpha, L_f dH) - X[L_f H] \\ &= J(L_f\alpha, \cdot)[H] + J(\alpha, \cdot)[L_f H] - X[L_f H] \\ &= J(L_f\alpha, \cdot)[H] \end{aligned}$$

for all $H \in C^\infty(\mathcal{X})$, showing that

$$L_f X = J(L_f\alpha, \cdot)$$

So $(L_f X, L_f\alpha) \in D$ and that means that f is a symmetry of D.
Conversely, let f be a symmetry of D. Then $(L_f X, L_f\alpha) \in D$, i.e. $L_f X = J(L_f\alpha, \cdot)$, for all $(X, \alpha) \in D$. Following the derivation above gives

$$L_f(J(\alpha, dH)) = J(L_f\alpha, dH) + J(\alpha, L_f dH)$$

for all $H \in C^\infty(\mathcal{X}),\ (X, \alpha) \in D$. By (11) this implies

$$(L_f J)(\alpha, dH) = 0$$

for all $H \in C^\infty(\mathcal{X}),\ (X, \alpha) \in D$, that is, for all $H \in C^\infty(\mathcal{X}),\ \alpha \in \mathsf{P}_1 = T^*\mathcal{X}$. Taking $H = x_i,\ i = 1, \ldots, n$ shows that $L_f J = 0$. □

The following proposition immediately follows from the definition.

Proposition 8. *[14] Let f be a symmetry of a generalized Dirac structure D, then $L_f\mathsf{G}_i \subset \mathsf{G}_i$, $L_f\mathsf{P}_i \subset \mathsf{P}_i$, $i = 0, 1$.*

The next proposition gives necessary and sufficient conditions for a vector field f to be a symmetry of a generalized Dirac structure D.

Proposition 9. *If the vector field f is symmetry of a generalized Dirac structure D, then*

- *f is canonical with respect to $\{\cdot, \cdot\}_D$, i.e.*

$$L_f\{H_1, H_2\}_D = \{L_f H_1, H_2\}_D + \{H_1, L_f H_2\}_D, \quad \forall H_1, H_2 \in \mathcal{A}_D,$$

- *$L_f\mathsf{G}_i \subset \mathsf{G}_i$, $L_f\mathsf{P}_i \subset \mathsf{P}_i$, $i = 0, 1$.*

If P_1 is constant dimensional and involutive then the converse is also true.

Proof. Take arbitrary $(X_i, dH_i) \in D, i = 1, 2$. Because f is symmetry also $(L_f X_i, L_f dH_i) = (L_f X_i, dL_f H_i) \in D, i = 1, 2$. Now,

$$\begin{aligned} L_f\{H_1, H_2\}_D &= L_f\langle dH_1, X_2\rangle = \langle L_f dH_1, X_2\rangle + \langle dH_1, L_f X_2\rangle \\ &= \{L_f H_1, H_2\}_D + \{H_1, L_f H_2\}_D. \end{aligned}$$

Now, suppose P_1 is constant dimensional and involutive. Then ([11],p.66) $\mathsf{P}_1 = \mathrm{span}\{d\beta_i\}, \beta_i \in C^\infty(\mathcal{X})$. First we prove that

$$\text{if } (X, dH) \in D, \text{then } (L_f X, L_f dH) \in D, \ \forall H \in \mathcal{A}_D. \tag{12}$$

Take arbitrary $H_1, H_2 \in \mathcal{A}_D$ i.e. $(X_i, dH_i) \in D, i = 1, 2$. Since

$$\begin{aligned} L_f\{H_1, H_2\}_D &= L_f\langle dH_1, X_2\rangle = \langle dL_f H_1, X_2\rangle + \langle dH_1, L_f X_2\rangle \\ &= \{L_f H_1, H_2\}_D + \langle dH_1, L_f X_2\rangle \end{aligned}$$

and

$$\{L_f H_1, H_2\}_D + \{H_1, L_f H_2\}_D = \{L_f H_1, H_2\}_D + \langle dH_1, X_{L_f H_2}\rangle$$

(because $L_f\mathsf{P}_1 \subset \mathsf{P}_1$ we have $L_f dH_2 = dL_f H_2 \in \mathsf{P}_1$, i.e., $(X_{L_f H_2}, dL_f H_2) \in D$) it follows from f being canonical that

$$\langle dH_1, X_{L_f H_2} - L_f X_2\rangle = 0$$

for arbitrary $dH_1 \in \mathsf{P}_1$. Because P_1 is spanned by exact one-forms it follows that $X_{L_f H_2} = L_f X_2 + Z$ with $Z \in \ker \mathsf{P}_1 = \mathsf{G}_0$. Now $(X_{L_f H_2}, L_f dH_2) = (L_f X_2 + Z, L_f dH_2) \in D$ and $Z \in \mathsf{G}_0$, i.e. $(Z, 0) \in D$, imply $(L_f X_2, L_f dH_2) \in D$. Since H_2 was arbitrary we have proved (12). Now, because P_1 is spanned by exact one-forms from (12) it follows easily that $(X, \alpha) \in D$ implies $(L_f X, L_f \alpha) \in D$ and so f is a symmetry of D. □

An other version of proposition 9 is the following. Define $\{\alpha_1, \alpha_2\} = \langle \alpha_1, X_2\rangle = -\langle \alpha_2, X_1\rangle$ for $\alpha_1, \alpha_2 \in \mathsf{P}_1$, i.e. $(X_i, \alpha_i) \in D, i = 1, 2$. Then we have

Proposition 10. *f is a symmetry of D if and only if*

- *f is canonical with respect to $\{\cdot, \cdot\}$ i.e.*

$$L_f\{\alpha_1, \alpha_2\} = \{L_f\alpha_1, \alpha_2\} + \{\alpha_1, L_f\alpha_2\}$$

- $L_f \mathsf{G}_i \subset \mathsf{G}_i, L_f \mathsf{P}_i \subset \mathsf{P}_i, i = 0, 1$

Proof. Analogously to the proof of proposition 9. $\square$

The following proposition says that the set of symmetries of D is involutive.

Proposition 11. *Let f_1 and f_2 both be symmetries of a generalized Dirac structure D. Then the Lie bracket $[f_1, f_2]$ is also a symmetry of D.*

Proof. We have

$$L_{[f_1,f_2]}X = [[f_1, f_2], X] = [[f_1, X], f_2] - [[f_2, X], f_1] = L_{f_1}L_{f_2}X - L_{f_2}L_{f_1}X,$$

and

$$L_{[f_1,f_2]}\alpha = L_{f_1}L_{f_2}\alpha - L_{f_2}L_{f_1}\alpha,$$

see [2], and the result immediately follows from definition 8. $\square$

Notice that the set of symmetries of D is *not* a distribution, because if f is a symmetry then $Hf, H \in C^\infty(\mathcal{X})$ is not in general.
Now we will turn to the notion of symmetries, and correspondingly first integrals, of implicit (generalized) Hamiltonian systems.

Definition 9. *Consider the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$, with D a generalized Dirac structure on $\mathcal{X}$. We call a nontrivial function $P \in C^\infty(\mathcal{X})$ a first integral for $(\mathcal{X}, D, H)$ if*

$$\frac{dP}{dt}(x(t)) = \langle dP(x(t)), X_H(x(t))\rangle = 0, \forall t \in I, \tag{13}$$

for all solutions $x(t)$ of $(\mathcal{X}, D, H)$, i.e. with $X_H(x(t)) = \dot{x}(t)$.

Remark 2. Condition (13) can be difficult to check in practice. A sufficient condition for (13) to hold is that

$$\langle dP(x), X_H(x) + \mathsf{G}_0(x)\rangle = 0, \quad \forall x \in \mathcal{X}_c,$$

where $X_H(x)$ is arbitrary such that $(X_H(x), dH(x)) \in D(x)$, for every $x \in \mathcal{X}_c$.

We recall the following two results.

Proposition 12. *[14, 5, 3] Let D be a closed Dirac structure on $\mathcal{X}$ and $f \in T\mathcal{X}$ for which there exists a $F \in C^\infty(\mathcal{X})$ such that $(f, dF) \in D$. Then f is a symmetry of D.*

Proposition 13. *[14] Consider the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ with D a generalized Dirac structure on $\mathcal{X}$ satisfying assumption 5. Let $f \in T\mathcal{X}$ for which there exists a $F \in C^\infty(\mathcal{X})$ such that $(f(x), dF(x)) \in D(x), \forall x \in \mathcal{X}_c$. Furthermore, let f be a symmetry of H on $\mathcal{X}_c$, i.e. $L_f H(x) = 0, \forall x \in \mathcal{X}_c$. Then $L_{X_H}F = 0$ on $\mathcal{X}_c$, that is, F is a first integral.*

We have the following proposition.

Proposition 14. *Consider the implicit Hamiltonian system $(\mathcal{X}, D, H)$, i.e. with closed Dirac structure D. Let $P_1, P_2 \in C^\infty(\mathcal{X})$ be two first integrals such that $P_1, P_2 \in \mathcal{A}_D$. Then $\{P_1, P_2\}_D$ is also a first integral (with $\{P_1, P_2\}_D \in \mathcal{A}_D$).*

Proof. $P_1, P_2 \in \mathcal{A}_D$, so there exist vector fields X_{P_1}, X_{P_2} such that $(X_{P_1}, dP_1), (X_{P_2}, dP_2) \in D$. Because D is closed, it follows from theorem 3 that $([X_{P_1}, X_{P_2}], d\{P_1, P_2\}_D) \in D$. Now

$$\begin{aligned}
\langle d\{P_1,P_2\}_D(x(t)), X_H(x(t))\rangle &= -\langle dH(x(t)), [X_{P_1}, X_{P_2}](x(t))\rangle \\
&= -i_{[X_{P_1},X_{P_2}]} dH(x(t)) \\
&= -L_{X_{P_1}}(i_{X_{P_2}} dH)(x(t)) + i_{X_{P_2}}(L_{X_{P_1}} dH)(x(t)) \\
&= i_{X_{P_2}}(d(i_{X_{P_1}} dH) + i_{X_{P_1}} d(dH))(x(t)) \\
&= 0,
\end{aligned}$$

for all solutions $x(t)$ of $(\mathcal{X}, D, H)$, where we used the fact that $D = D^\perp$ and $i_{X_{P_k}} dH(x(t)) = \langle dH(x(t)), X_{P_k}(x(t))\rangle = 0, k = 1, 2$, because P_1, P_2 are first integrals. Thus, $\{P_1, P_2\}_D$ is also a first integral of $(\mathcal{X}, D, H)$. Note that we could also have used proposition 13 if assumption 5 is satisfied. □

Definition 10. *We will call a vector field $f \in T\mathcal{X}$ a symmetry of the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ if f is a symmetry of the generalized Dirac structure D (as in definition 8) and f is a symmetry of H, i.e. $L_f H(x(t)) = 0$ for all solutions $x(t)$ of $(\mathcal{X}, D, H)$, that is, f leaves H invariant (along solutions).*

Notice again that a sufficient condition for f to be a symmetry of H is that $L_f H(x) = 0$, $\forall x \in \mathcal{X}_c$.

We have the following proposition, corresponding to proposition 6.31 in [12].

Proposition 15. *Consider the implicit Hamiltonian system $(\mathcal{X}, D, H)$ and assume that D is closed. Let P be a first integral such that $P \in \mathcal{A}_D$, i.e. there exists a vector field X_P such that $(X_P, dP) \in D$. Then X_P is a symmetry of $(\mathcal{X}, D, H)$. Furthermore, X_P generates a one-parameter symmetry group of $(\mathcal{X}, D, H)$, i.e. the flow of X_P.*

Proof. We have $(X_P(x(t)), dP(x(t))), (X_H(x(t)), dH(x(t))) \in D$ for all solutions $x(t)$ of $(\mathcal{X}, D, H)$. from $D = D^\perp$ it follows that

$$\langle dH(x(t)), X_P(x(t))\rangle + \langle dP(x(t)), X_H(x(t))\rangle = 0. \tag{14}$$

Now because P is a first integral, from (13) it follows that $L_{X_P} H(x(t)) = \langle dH(x(t)), X_P(x(t))\rangle = 0$ for all solutions $x(t)$ of $(\mathcal{X}, D, H)$ so X_P is a symmetry of H. Because $(X_P, dP) \in D$ it follows from proposition 12 that X_P is a symmetry of D. Furthermore, from remark 14 [14] it is evident that the flow $\phi_t^{X_P}$ of X_P generates a one-parameter symmetry group of $(\mathcal{X}, D, H)$. □

From proposition 15 and proposition 13 we immediately get a generalization of theorem 6.33 in [12]. First we need the following.

Definition 11. *Consider a generalized Dirac structure D on $\mathcal{X}$. A nontrivial function $C \in C^\infty(\mathcal{X})$ is called a Casimir function if C is a first integral of $(\mathcal{X}, D, H)$, as in definition 9, for every $H \in C^\infty(\mathcal{X})$.*

Proposition 16. *Consider a generalized Dirac structure D on $\mathcal{X}$ and a function $C \in \mathcal{A}_D$, i.e. $(X_C, dC) \in D$. If $X_C \in \mathsf{G}_0$, or equivalently $dC \in \mathsf{P}_0$, then C is a Casimir function. If P_1 is constant dimensional and involutive, the converse is also true.*

Proof. Take arbitrary $H \in C^\infty(\mathcal{X})$. Like in (14) it follows that

$$\langle dH(x(t)), X_C(x(t))\rangle + \langle dC(x(t)), X_H(x(t))\rangle = 0, \tag{15}$$

for all solutions $x(t)$ of $(\mathcal{X}, D, H)$. Suppose $X_C \in G_0 = \ker \mathsf{P}_1$ then $\langle dH(x(t)), X_C(x(t))\rangle = 0$, and from (15) it follows that C is a first integral of $(\mathcal{X}, D, H)$. Conversely, suppose C is a Casimir function. Because P_1 is constant dimensional and involutive there exist local coordinates $(y, s) = (y_1, \ldots, y_{n-m}, s_1, \ldots, s_m)$ for $\mathcal{X}$ in which

$$\mathsf{P}_1 = \operatorname{span}\{dy_1, \ldots, dy_{n-m}\}.$$

C Casimir means that $\langle dC(x(t)), X_H(x(t))\rangle = 0$, for all solutions $x(t)$ of $(\mathcal{X}, D, H)$, for arbitrary $H \in C^\infty(\mathcal{X})$. Take $H_i(y, s) = y_i, i = 1, \ldots, n-m$, then $(\mathcal{X}_c)_{H=y_i} = \mathcal{X}$ because $H_i = y_i \in \mathcal{A}_D$, which implies that through each $x \in \mathcal{X}$ there goes a solution $x(t)$ of $(\mathcal{X}, D, y_i)$. It follows from (15) that $\langle dy_i, X_C\rangle = 0, i = 1, \ldots, n-m$, which implies that $X_C \in \ker \mathsf{P}_1 = \mathsf{G}_0$. □

Note that definition 11 and proposition 16 do not assume that the generalized Dirac structure is closed. In proposition 17 we will assume that the generalized Dirac structure D *is* closed, this implies that the codistribution P_1 is involutive [4].

Proposition 17. *Consider the implicit Hamiltonian system $(\mathcal{X}, D, H)$ and assume that D is closed. Furthermore, assume that assumption 5 is satisfied. If $P \in \mathcal{A}_D$ is a first integral then the corresponding vector field X_P is a symmetry of $(\mathcal{X}, D, H)$. Conversely, if $X_P \in T\mathcal{X}$ is a symmetry of $(\mathcal{X}, D, H)$ such that $(X_P, dP) \in D$ for some $P \in C^\infty(\mathcal{X})$, then P is a first integral. $\tilde{P} \in C^\infty(\mathcal{X})$ is a second function such that $(X_P, d\tilde{P}) \in D$ only if $\tilde{P} = P + C$ for some Casimir function C. If P_1 is constant dimensional then the converse is also true.*

Proof. The first two statements are proved in proposition 15 and 13 respectively. Now suppose $(X_P, dP), (X_P, d\tilde{P}) \in D$, then it follows that $(0, d(\tilde{P} - P)) \in D$ or $X_{\tilde{P}-P} = 0 \in \mathsf{G}_0$. Proposition 16 implies that $\tilde{P} - P = C$ is a Casimir function. Conversely, suppose that $\tilde{P} - P = C$ is a Casimir function, i.e. $X_{\tilde{P}-P} = Z \in \mathsf{G}_0$. Then $(Z, d(\tilde{P} - P)) \in D$. $Z \in \mathsf{G}_0$ implies $(Z, 0) \in D$ so it follows that $(0, d(\tilde{P} - P)) \in D$. Because also $(X_P, dP) \in D$ it follows that $(X_P, d\tilde{P})$ is also in D. □

Remark 3. In this section we derived some results about symmetries and first integrals of Dirac structures and implicit generalized Hamiltonian systems. For some converse results we assumed the constant dimensionality and involutivity of P_1. We want to remark that in the case of mechanical systems with kinematic constraints $A^T(q)\dot{q} = 0$, the codistribution P_1 is always constant dimensional and involutive [4].

4 Reduction

In this section we will derive some results on the reduction of generalized Dirac structures and correspondingly implicit generalized Hamiltonian systems.

4.1 Reduction of Dirac structures

Investigating reduction of implicit Hamiltonian systems we begin by looking at reduction of Dirac structures. Consider a manifold $\mathcal{X}$ and a generalized Dirac structure D on $\mathcal{X}$. Let $\bar{\mathcal{X}}$ be a submanifold of $\mathcal{X}$, then D induces a generalized Dirac structure $\bar{D}$ on $\bar{\mathcal{X}}$. This can be seen by the following. Assume that the distribution G_1, corresponding to D, is constant dimensional, then by theorem 4a there exists a skew-symmetric linear map $\omega(x) : \mathsf{G}_1(x) \to \mathsf{G}_1(x)^*$ such that the generalized Dirac structure D can be written as

$$D = \{(X, \alpha) \in T\mathcal{X} \oplus T^*\mathcal{X} \mid \alpha(x) - \omega(x)X(x) \in \operatorname{ann} \mathsf{G}_1(x),\ \forall x \in \mathcal{X},\ X \in \mathsf{G}_1\}. \tag{16}$$

The reduced generalized Dirac structure $\bar{D}$ on $\bar{\mathcal{X}}$ is now defined by restricting the map $\omega(x)$ to $\mathsf{G}_1(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}}$, $\bar{x} \in \bar{\mathcal{X}}$, giving the map $\bar{\omega}(\bar{x})$, i.e.

$$\begin{aligned} \bar{D} \;\; = \;\; & \{(\bar{X}, \bar{\alpha}) \in T\bar{\mathcal{X}} \oplus T^*\bar{\mathcal{X}} \mid \bar{\alpha}(\bar{x}) - \bar{\omega}(\bar{x})\bar{X}(\bar{x}) \in \text{ann}\,(\mathsf{G}_1(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}}), \\ & \bar{X}(\bar{x}) \in \mathsf{G}_1(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}}, \; \forall \bar{x} \in \bar{\mathcal{X}}\}, \end{aligned} \tag{17}$$

see also [3]. It follows from theorem 4a (assuming that $\mathsf{G}_1(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}}$ is constant dimensional) that $\bar{D}$ is a generalized Dirac structure on $\bar{\mathcal{X}}$. We will show that $\bar{D}$ can also be written in terms of the inclusion map $\iota : \bar{\mathcal{X}} \to \mathcal{X}$.

Proposition 18. *Consider a manifold $\mathcal{X}$ and a generalized Dirac structure D on $\mathcal{X}$ with G_1 constant dimensional. Let $\bar{\mathcal{X}}$ be a submanifold of $\mathcal{X}$, and assume that $\mathsf{G}_1(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}}$, $\bar{x} \in \bar{\mathcal{X}}$, is constant dimensional (on $\bar{\mathcal{X}}$). Then D induces a generalized Dirac structure $\bar{D}$ on $\bar{\mathcal{X}}$ given by*

$$\begin{aligned} \bar{D} \;\; = \;\; & \{(\bar{X}, \bar{\alpha}) \in T\bar{\mathcal{X}} \oplus T^*\bar{\mathcal{X}} \mid \exists\, X \textit{ such that } \bar{X} \sim_\iota X \textit{ and} \,\exists\, \alpha \textit{ such that} \\ & \bar{\alpha} = \iota^*\alpha \textit{ with } (X, \alpha) \in D\}. \end{aligned} \tag{18}$$

Furthermore, if D is closed then also $\bar{D}$ is closed.

Proof. Denote $\bar{D}$ in (17) by $\bar{D}_1$ and $\bar{D}$ in (18) by $\bar{D}_2$. We prove that $\bar{D}_1 = \bar{D}_2$.
$\bar{D}_2 \subset \bar{D}_1$: Let $(\bar{X}, \bar{\alpha}) \in \bar{D}_2$. There exists a vector field $X \in \mathsf{G}_1$ such that $\bar{X} \sim_\iota X$. This means that at points of $\bar{\mathcal{X}}$, X is tangent to $\bar{\mathcal{X}}$, so $\bar{X}(\bar{x}) = X(\bar{x}) \in \mathsf{G}_1(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}}$ for all $\bar{x} \in \bar{\mathcal{X}}$. Let $\bar{\alpha} = \iota^*\alpha$ where $\alpha(x) - \omega(x)X(x) \in \text{ann}\,\mathsf{G}_1(x)$, $\forall x \in \mathcal{X}$, i.e. $(X, \alpha) \in D$, then

$$(\iota^*\alpha)(\bar{x}) - \iota^*(\omega X)(\bar{x}) \in \iota^*(\text{ann}\,\mathsf{G}_1)(\bar{x}), \; \forall \bar{x} \in \bar{\mathcal{X}},$$

and so, because $\bar{X} \sim_\iota X$,

$$\bar{\alpha}(\bar{x}) - \bar{\omega}(\bar{x})\bar{X}(\bar{x}) \in \iota^*(\text{ann}\,\mathsf{G}_1)(\bar{x}), \; \forall \bar{x} \in \bar{\mathcal{X}}$$

Now, because $\iota^*(\text{ann}\,\mathsf{G}_1)(\bar{x}) \subset \text{ann}\,(\mathsf{G}_1(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}})$, $\forall \bar{x} \in \bar{\mathcal{X}}$, we get

$$\bar{\alpha}(\bar{x}) - \bar{\omega}(\bar{x})\bar{X}(\bar{x}) \in \text{ann}\,(\mathsf{G}_1(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}}), \; \forall \bar{x} \in \bar{\mathcal{X}},$$

which means that $(\bar{X}, \bar{\alpha}) \in \bar{D}_1$.
$\bar{D}_1 \subset \bar{D}_2$: Let $(\bar{X}, \bar{\alpha}) \in \bar{D}_1$. Then $\bar{X}(\bar{x}) \in \mathsf{G}_1(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}}$, $\forall \bar{x} \in \bar{\mathcal{X}}$. Because G_1 is a smooth subbundle of $T\mathcal{X}$ it follows that $\bar{X}$ can be extended to a vector field $X \in \mathsf{G}_1$ such that $\bar{X} \sim_\iota X$ (one can use the Smooth Tietze Extension Theorem ([2],theorem 5.5.9), note that X is not unique). There exists an α such that $(X, \alpha) \in D$, i.e. $\alpha(x) - \omega(x)X(x) \in \text{ann}\,G_1(x)$, $\forall x \in \mathcal{X}$. Then, by the above,

$$(\iota^*\alpha)(x) - \bar{\omega}(\bar{x})\bar{X}(\bar{x}) \in \text{ann}\,(\mathsf{G}_1(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}}), \; \forall \bar{x} \in \bar{\mathcal{X}}, \tag{19}$$

and so, by (17) and (19),

$$\bar{\alpha}(\bar{x}) - (\iota^*\alpha)(\bar{x}) \in \text{ann}\,(\mathsf{G}_1(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}}) \subset T^*_{\bar{x}}\bar{\mathcal{X}}, \; \forall \bar{x} \in \bar{\mathcal{X}}, \tag{20}$$

(note that the annihilation should be taken with respect to $T^*\bar{\mathcal{X}}$). However, because $\bar{\mathcal{X}}$ is a submanifold of $\mathcal{X}$ there exists (locally) a function $F \in C^\infty(\mathcal{X})$ such that $\bar{\mathcal{X}} = F^{-1}(0)$, i.e. a level set of F. $\mathsf{G}_1(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}}$ consists of all vector fields $X(\bar{x}) \in \mathsf{G}_1(\bar{x})$ which are tangent to $\bar{\mathcal{X}}$, so when we take the annihilator with respect to $T^*\mathcal{X}$

$$T^*_{\bar{x}}\mathcal{X} \supset \text{ann}\,(\mathsf{G}_1(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}}) = \text{span}_{C^\infty(\mathcal{X})}\{dF\}(\bar{x}) + \mathsf{P}_0(\bar{x}), \; \forall \bar{x} \in \bar{\mathcal{X}}.$$

Considered as an element of $T^*_{\bar{x}}\bar{\mathcal{X}}$, that is taking the annihilation with respect to $T^*\bar{\mathcal{X}}$, $dF(\bar{x})$ will be zero, i.e. $\iota^* dF(\bar{x}) = d\iota^* F(\bar{x}) = d0 = 0$, $\forall \bar{x} \in \bar{\mathcal{X}}$. Furthermore, the elements of P_0 will restrict to elements of $\iota^*\mathsf{P}_0 \subset T^*\bar{\mathcal{X}}$. Now (20) becomes

$$\bar{\alpha}(\bar{x}) - (\iota^*\alpha)(\bar{x}) \in \iota^*\mathsf{P}_0(\bar{x}),\ \forall \bar{x} \in \bar{\mathcal{X}}.$$

This means that $\bar{\alpha} = \iota^*\alpha + \iota^*\alpha_0$ for some $\alpha_0 \in \mathsf{P}_0$. Define $\beta = \alpha + \alpha_0$ then $\bar{\alpha} = \iota^*\beta$ and $(X, \beta) \in D$ (because $(X, \alpha) \in D$ and $(0, \alpha_0) \in D$). Therefore $(\bar{X}, \bar{\alpha}) \in \bar{D}_2$.

Now, assume that D is closed. Take arbitrary $(\bar{X}_k, \bar{\alpha}_k) \in \bar{D}$, $k = 1, 2, 3$, then $\bar{X}_k \sim_\iota X_k$ and $\bar{\alpha}_k = \iota^*\alpha_k$, with $(X_k, \alpha_k) \in D$ for some X_k and α_k, $k = 1, 2, 3$. Then,

$$\begin{aligned}
&\langle L_{\bar{X}_1}\bar{\alpha}_2, \bar{X}_3\rangle + \langle L_{\bar{X}_2}\bar{\alpha}_3, \bar{X}_1\rangle + \langle L_{\bar{X}_3}\bar{\alpha}_1, \bar{X}_2\rangle = \\
&\langle L_{\bar{X}_1}\iota^*\alpha_2, \bar{X}_3\rangle + \langle L_{\bar{X}_2}\iota^*\alpha_3, \bar{X}_1\rangle + \langle L_{\bar{X}_3}\iota^*\alpha_1, \bar{X}_2\rangle = \\
&\langle \iota^* L_{X_1}\alpha_2, \bar{X}_3\rangle + \langle \iota^* L_{X_2}\alpha_3, \bar{X}_1\rangle + \langle \iota^* L_{X_3}\alpha_1, \bar{X}_2\rangle = \\
&\langle L_{X_1}\alpha_2, X_3\rangle + \langle L_{X_2}\alpha_3, X_1\rangle + \langle L_{X_3}\alpha_1, X_2\rangle = 0,
\end{aligned}$$

because D is closed. This shows that also $\bar{D}$ is closed. □

There is also a direct proof of proposition 18, without having to involve (16,17). We assume that $\bar{\mathcal{X}}$ is a submanifold of $\mathcal{X}$ with $\dim \bar{\mathcal{X}} < \dim \mathcal{X}$, that is, $\bar{\mathcal{X}}$ is closed in $\mathcal{X}$. Define $\bar{D}$ as in (18). Because D is a linear space, that is $(X_i, \alpha_i) \in D$, $i = 1, 2$, implies $(X_1, \alpha_1) + (X_2, \alpha_2) = (X_1 + X_2, \alpha_1 + \alpha_2) \in D$ and $h(X_1, \alpha_1) = (hX_1, h\alpha_1) \in D$, $\forall h \in C^\infty(\mathcal{X})$, it easily follows that this also holds for $\bar{D}$. Thus, for every point $\bar{x} \in \bar{\mathcal{X}}$, $\bar{D}(\bar{x})$ is a linear subspace of $T_{\bar{x}}\bar{\mathcal{X}} \times T^*_{\bar{x}}\bar{\mathcal{X}}$. We make the assumption that $\dim(D(\bar{x}) \cap E_s(\bar{x})) = d$, $\forall \bar{x} \in \bar{\mathcal{X}}$, for some integer d (i.e. constant), where E_s is defined as the smooth bundle

$$E_s = \{(X, \alpha) \in T\mathcal{X} \oplus T^*\mathcal{X} \mid \bar{X} \sim_\iota X \text{ for some } \bar{X} \in T\bar{\mathcal{X}}\}, \tag{21}$$

(the subscript s stands for submanifold). This assumption equals the condition in [3]. Courant [3] calls $\bar{\mathcal{X}}$ under this assumption a *clean* submanifold of $\mathcal{X}$.

Proposition 19. *Assume that $D(\bar{x}) \cap E_s(\bar{x})$, $\bar{x} \in \bar{\mathcal{X}}$, is constant dimensional on $\bar{\mathcal{X}}$. Then $\bar{D}$ defined in (18) is a generalized Dirac structure on $\bar{\mathcal{X}}$.*

Proof. We begin by proving that $\bar{D} = \bar{D}^\perp$. The first inclusion, i.e. $\bar{D} \subset \bar{D}^\perp$, is easy. We prove the second inclusion, i.e. $\bar{D}^\perp \subset \bar{D}$. Take an arbitrary pair $(\bar{Y}, \bar{\beta}) \in \bar{D}^\perp$, that is

$$(\bar{Y}, \bar{\beta}) \in T\bar{\mathcal{X}} \oplus T^*\bar{\mathcal{X}} \text{ s.t. } \langle \bar{\beta}, \bar{X}\rangle + \langle \bar{\alpha}, \bar{Y}\rangle = 0,\ \forall (\bar{X}, \bar{\alpha}) \in \bar{D}.$$

There exist $Y \in T\mathcal{X}$ such that $\bar{Y} \sim_\iota Y$ and $\beta \in T^*\mathcal{X}$ such that $\bar{\beta} = \iota^*\beta$ (because ι^* is surjective). Notice that this only defines Y and β at points $\bar{x} \in \bar{\mathcal{X}} \subset \mathcal{X}$. Now,

$$0 = \langle \bar{\beta}, \bar{X}\rangle + \langle \bar{\alpha}, \bar{Y}\rangle = \langle \iota^*\beta, \bar{X}\rangle + \langle \iota^*\alpha, \bar{Y}\rangle = (\langle \beta, X\rangle + \langle \alpha, Y\rangle) \circ \iota,$$

which means that

$$\langle \beta, X\rangle(\bar{x}) + \langle \alpha, Y\rangle(\bar{x}) = 0$$

for all $\bar{x} \in \bar{\mathcal{X}} \subset \mathcal{X}$ and all pairs $(X, \alpha) \in D$ for which $\bar{X} \sim_\iota X$ for some $\bar{X} \in T\bar{\mathcal{X}}$. Therefore

$$(Y, \beta)(\bar{x}) \in [(D \cap E_s)(\bar{x})]^\perp = [D(\bar{x}) \cap E_s(\bar{x})]^\perp = D(\bar{x}) + [E_s(\bar{x})]^\perp = D(\bar{x}) + (0, \operatorname{ann} T_{\bar{x}}\bar{\mathcal{X}}) \tag{22}$$

for all $\bar{x} \in \bar{\mathcal{X}} \subset \mathcal{X}$, with E_s defined as in (21) (and where we used the assumption on constant dimensionality at the first equality, see e.g. [6]).

Consider

$$\tilde{E}_s = \{(0, \gamma) \in T\mathcal{X} \oplus T^*\mathcal{X} \mid \iota^*\gamma = 0\},$$

then $\tilde{E}_s$ is a smooth bundle. Indeed, $\tilde{E}_s(x)$, $x \notin \bar{\mathcal{X}}$, can locally (that is, in some neighborhood $U \subset \mathcal{X}$ of x, $U \cap \bar{\mathcal{X}} = \emptyset$) be written as

$$\tilde{E}_s(x) = \operatorname{span}_{C^\infty(\mathcal{X})}\{(0, dx_1), \ldots, (0, dx_n)\},$$

where $x_1, \ldots, x_n$ are local coordinates for $\mathcal{X}$ around x. Consider a point $\bar{x} \in \bar{\mathcal{X}}$. Because $\bar{\mathcal{X}}$ is a submanifold of $\mathcal{X}$ there exist local coordinates $x_1, \ldots, x_m, x_{m+1}, \ldots, x_n$ for $\mathcal{X}$ in some neighborhood U of $\bar{x}$ such that $x_1, \ldots, x_m$ are local coordinates for $\bar{X}$. Then $\tilde{E}_s(x)$ can be written as

$$\tilde{E}_s(x) = \operatorname{span}_{C^\infty(\mathcal{X})}\{f_1(x)(0, dx_1), \ldots, f_m(x)(0, dx_m), (0, dx_{m+1}), \ldots, (0, dx_n)\},$$

for all $x \in U$, with $f_1, \ldots, f_m \in C^\infty(U)$ such that $f_i(x) = 0 \Leftrightarrow x \in \bar{\mathcal{X}}$.

Notice that $\tilde{E}_s(\bar{x}) = (0, \operatorname{ann} T_{\bar{x}}\bar{\mathcal{X}})$ for all $\bar{x} \in \bar{\mathcal{X}}$. Then (22) becomes

$$(Y, \beta)(\bar{x}) \in D(\bar{x}) + \tilde{E}_s(\bar{x}), \tag{23}$$

for all $\bar{x} \in \bar{\mathcal{X}} \subset \mathcal{X}$. Because D is also a smooth bundle (by definition), around every point $x \in \mathcal{X}$ there exists a local basis $(X_i, \alpha_i) \in D$, $i = 1, \ldots, n$, – where X_i and α_i are locally (that is, around x) smooth vector fields, respectively one-forms – such that locally

$$D = \operatorname{span}_{C^\infty(\mathcal{X})}\{(X_i, \alpha_i)\}.$$

From (23) it follows that we can write

$$(Y, \beta)(\bar{x}) = \sum_{i=1}^{n} h_i(\bar{x})(X_i, \alpha_i)(\bar{x}) + \sum_{j=m+1}^{n} g_j(\bar{x})(0, dx_j) \tag{24}$$

for some functions $h_i, g_j \in C^\infty(U), i = 1, \ldots, n, j = m+1, \ldots, n$, $U \subset \mathcal{X}$ a neighborhood of $\bar{x}$. Define

$$\gamma(\bar{x}) = \sum_{j=m+1}^{n} g_j(\bar{x}) dx_j,$$

then from (24)

$$(Y, \beta - \gamma)(\bar{x}) \in D(\bar{x}),\ \forall \bar{x} \in \bar{\mathcal{X}} \subset \mathcal{X}.$$

Because of (24) $(Y, \beta - \gamma)$ can be locally, that is in some neighborhood $U \subset \mathcal{X}$ of every $\bar{x}$, extended to a smooth pair (Y_e, β_e) defined on U such that

$$Y_e(\bar{x}) = Y(\bar{x}),\ \beta_e(\bar{x}) = \beta(\bar{x}) - \gamma(\bar{x}),\ \forall \bar{x} \in U \cap \bar{\mathcal{X}},$$

and $(Y_e, \beta_e)(x) \in D(x)$, $\forall x \in U$. Indeed, take

$$(Y_e, \beta_e)(x) = \sum_{i=1}^{n} h_i(x)(X_i, \alpha_i)(x), \quad x \in U.$$

Then, by the Smooth Tietze Extension Theorem ([2],theorem 5.5.9), $(Y, \beta - \gamma)$ can be globally extended to a pair

$$(Y', \beta') \in D \tag{25}$$

such that

$$Y'(\bar{x}) = Y(\bar{x}),\ \beta'(\bar{x}) = \beta(\bar{x}) - \gamma(\bar{x}),\ \forall \bar{x} \in \bar{\mathcal{X}} \subset \mathcal{X}$$

(the proof relies on a partition of unity on $\mathcal{X}$). It follows that

$$Y' \sim_i \bar{Y} \tag{26}$$

and

$$\iota^*\beta' = \iota^*(\beta - \gamma) = \iota^*\beta - 0 = \bar{\beta}, \tag{27}$$

where we used that $\iota^*\beta'$ only depends on the definition of β' in the points $\bar{x} \in \bar{\mathcal{X}} \subset \mathcal{X}$. Now (25,26,27) imply that $(\bar{Y}, \bar{\beta}) \in \bar{D}$. So we have proved that $\bar{D}^\perp \subset \bar{D}$. So $\bar{D} = \bar{D}^\perp$. Smoothness of the pairs $(\bar{X}, \bar{\alpha}) \in \bar{D}$ comes from smoothness of D, and thus $\bar{D}$ is a generalized Dirac structure on $\bar{\mathcal{X}}$. □

Remark 4. With respect to the comparison of propositions 18 and 19 we remark that (i) G_1 and $\mathsf{G}(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}},\ \bar{x} \in \bar{\mathcal{X}}$, constant dimensional imply $D(\bar{x}) \cap E_s(\bar{x}),\ \bar{x} \in \bar{\mathcal{X}}$, constant dimensional, and (ii) G_1 and $D(\bar{x}) \cap E_s(\bar{x}),\ \bar{x} \in \bar{\mathcal{X}}$, constant dimensional imply $\mathsf{G}(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}},\ \bar{x} \in \bar{\mathcal{X}}$, constant dimensional.

Consider a manifold $\mathcal{X}$ and a generalized Dirac structure D on $\mathcal{X}$. Consider a symmetry Lie group G of D, that is, every $g \in G$ induces an action $\phi_g : \mathcal{X} \to \mathcal{X}$ on $\mathcal{X}$, which is a diffeomorphism, and ϕ_g is a symmetry of the generalized Dirac structure D. Equivalently, let $\mathcal{G}$ be the Lie algebra corresponding to G, then for every $\xi \in \mathcal{G}$ the infinitesimal generator $\xi_{\mathcal{X}}$, i.e. the vector field on $\mathcal{X}$ generated by $\xi \in \mathcal{G}$ (see for instance [7, 12]), is an (infinitesimal) symmetry of D as in definition 8. Then the generalized Dirac structure D on $\mathcal{X}$ induces a generalized Dirac structure $\hat{D}$ on the quotient space $\hat{\mathcal{X}} = \mathcal{X}/G$ of G-orbits on $\mathcal{X}$. Throughout we assume that $\hat{\mathcal{X}} = \mathcal{X}/G$ has a manifold structure. The usual assumption made is that G acts freely and properly on $\mathcal{X}$ (which is a sufficient condition, see [1]). Furthermore, in proposition 20 we need the following assumptions. Let V denote the distribution spanned by the infinitesimal generators of G. Assume that $V + \mathsf{G}_0$ is constant dimensional. Furthermore, define the smooth bundle

$$E_q = \{(X, \alpha) \in T\mathcal{X} \oplus T^*\mathcal{X} \mid \alpha = \pi^*\hat{\alpha} \text{ for some } \hat{\alpha} \in T^*\hat{\mathcal{X}}\}, \tag{28}$$

(the subscript q stands for quotient manifold). We assume that $D \cap E_q$ is constant dimensional (on $\mathcal{X}$).

Proposition 20. *[14] Consider a manifold $\mathcal{X}$ and a generalized Dirac structure D on $\mathcal{X}$. Let G be a symmetry Lie group of D and assume that $V + \mathsf{G}_0$ and $D \cap E_q$ are constant dimensional. Then D induces a generalized Dirac structure $\hat{D}$ on $\hat{\mathcal{X}} = \mathcal{X}/G$ given by*

$$\begin{aligned} \hat{D} &= \{(\hat{X}, \hat{\alpha}) \in T\hat{\mathcal{X}} \oplus T^*\hat{\mathcal{X}} \mid \exists\, X \textit{ such that } X \sim_\pi \hat{X} \textit{ and } (X, \alpha) \in D \\ &\quad\ \textit{where } \alpha = \pi^*\hat{\alpha}\}. \end{aligned}$$

Here, $\pi : \mathcal{X} \to \hat{\mathcal{X}} = \mathcal{X}/G$ is the projection map. Furthermore, if D is closed then also $\hat{D}$ is closed.

Proof. We show that $\hat{D}$ is a generalized Dirac structure. The first inclusion $\hat{D} \subset \hat{D}^{\perp}$ is easy. We prove the second inclusion, $\hat{D}^{\perp} \subset \hat{D}$. Take an arbitrary pair $(\hat{Y}, \hat{\beta}) \in \hat{D}^{\perp}$, that is

$$(\hat{Y}, \hat{\beta}) \in T\hat{\mathcal{X}} \oplus T^*\hat{\mathcal{X}} \text{ s.t. } \langle \hat{\beta}, \hat{X} \rangle + \langle \hat{\alpha}, \hat{Y} \rangle = 0, \ \forall (\hat{X}, \hat{\alpha}) \in \hat{D}. \tag{29}$$

Let $Y \in T\mathcal{X}$ be such that $Y \sim_{\pi} \hat{Y}$ and define $\beta = \pi^* \hat{\beta}$, then (29) becomes

$$0 = \langle \hat{\beta}, \hat{X} \rangle + \langle \hat{\alpha}, \hat{Y} \rangle = (\langle \beta, X \rangle + \langle \alpha, Y \rangle) \circ \pi,$$

which means that

$$\langle \beta, X \rangle + \langle \alpha, Y \rangle = 0 \tag{30}$$

for all $(X, \alpha) \in D$ for which $X \sim_{\pi} \hat{X}$ and $\alpha = \pi^* \hat{\alpha}$ for some $\hat{X} \in T\hat{\mathcal{X}}, \ \hat{\alpha} \in T^*\hat{\mathcal{X}}$. Now consider an arbitrary $(X, \alpha) \in D$ with $\alpha = \pi^* \hat{\alpha}$ for some $\hat{\alpha} \in T^*\hat{\mathcal{X}}$. Since G is a symmetry group

$$(L_{\xi_{\mathcal{X}}} X, L_{\xi_{\mathcal{X}}} \pi^* \hat{\alpha}) \in D,$$

for all infinitesimal generators $\xi_{\mathcal{X}}, \ \xi \in \mathcal{G}$. Since $L_{\xi_{\mathcal{X}}} \pi^* \hat{\alpha} = 0$, this yields

$$L_{\xi_{\mathcal{X}}} X \in \mathsf{G}_0, \quad \forall \xi_{\mathcal{X}}, \ \xi \in \mathcal{G}. \tag{31}$$

Furthermore, by proposition 8, $L_{\xi_{\mathcal{X}}} \mathsf{G}_0 \subset \mathsf{G}_0$. Take an arbitrary $v = \sum_i h_i (\xi_i)_{\mathcal{X}} \in V$, where $\{\xi_i\}_i$ is a basis of $\mathcal{G}$ and $h_i \in C^{\infty}(\mathcal{X})$, then by (31)

$$[X, v] = \sum_i h_i [X, (\xi_i)_{\mathcal{X}}] + \sum_i L_X h_i \ (\xi_i)_{\mathcal{X}} \in \mathsf{G}_0 + V,$$

so

$$[X, V] \subset V + \mathsf{G}_0.$$

Analogously, it follows that

$$[\mathsf{G}_0, V] \subset V + \mathsf{G}_0.$$

Now, since $V + \mathsf{G}_0$, is constant dimensional we have the following properties (see [6, 11] for the analog in controlled invariant distributions)

(a) there exist $Z_1, \ldots, Z_k$ which span G_0 such that $[Z_i, V] \subset V$, which implies that $Z_i \sim_{\pi} \hat{Z}_i$ for some $\hat{Z}_i \in T\hat{\mathcal{X}}$, $i = 1, \ldots, k$,

(b) there exists a $Z \in \mathsf{G}_0$ such that $[X + Z, V] \subset V$, which implies that $X + Z \sim_{\pi} \hat{X}$ for some $\hat{X} \in T\hat{\mathcal{X}}$.

Take an arbitrary $Z \in \mathsf{G}_0$ such that $Z \sim_{\pi} \hat{Z}$ for some $\hat{Z} \in T\hat{\mathcal{X}}$, then by (30) it follows that

$$\langle \pi^* \hat{\beta}, Z \rangle = 0.$$

Therefore

$$\langle \pi^* \hat{\beta}, Z_i \rangle = 0, \ i = 1, \ldots, k,$$

and since $Z_1, \ldots, Z_k$ span G_0

$$\langle \pi^* \hat{\beta}, \mathsf{G}_0 \rangle = 0. \tag{32}$$

Now take any pair $(X, \alpha) \in D$ for which there exists an $\hat{\alpha} \in T^*\hat{\mathcal{X}}$ such that $\alpha = \pi^*\hat{\alpha}$. Then by (b) there exists a $Z \in \mathsf{G}_0$ (so $(X + Z, \alpha) \in D$) such that $X + Z \sim_\pi \hat{X}$ for some $\hat{X} \in T\hat{\mathcal{X}}$, and so by (30)

$$\langle \beta, X + Z \rangle + \langle \alpha, Y \rangle = 0,$$

which by (32) and the fact that $\beta = \pi^*\hat{\beta}$ implies

$$\langle \beta, X \rangle + \langle \alpha, Y \rangle = 0. \tag{33}$$

Thus we have shown that (30), or (33), holds for all $(X, \alpha) \in D$ such that $\alpha = \pi^*\hat{\alpha}$ for some $\hat{\alpha} \in T^*\hat{\mathcal{X}}$. Hence

$$(Y, \beta) \in (D \cap E_q)^\perp = D + E_q^\perp, \tag{34}$$

where we used the constant dimensionality of $D \cap E_q$. We claim that

$$E_q^\perp = \{(\tilde{X}, 0) \in T\mathcal{X} \oplus T^*\mathcal{X} \mid \tilde{X} \sim_\pi 0\}. \tag{35}$$

Indeed, the inclusion $\supset$ is obvious, while for the reverse inclusion we note that if $(\tilde{X}, \tilde{\alpha})$ is such that

$$\langle \tilde{\alpha}, X \rangle + \langle \alpha, \tilde{X} \rangle = 0,$$

for all $(X, \alpha) \in E_q$, then (taking $X = 0$) $\langle \alpha, \tilde{X} \rangle = 0$ for all $\alpha = \pi^*\hat{\alpha},\ \hat{\alpha} \in T^*\hat{\mathcal{X}}$, and thus $\tilde{X} \sim_\pi 0$. Hence

$$0 = \langle \tilde{\alpha}, X \rangle + \langle \alpha, \tilde{X} \rangle = \langle \tilde{\alpha}, X \rangle,$$

for all $X \in T\mathcal{X}$, implying that $\tilde{\alpha} = 0$. This proves the claim. By (34,35) there exists a vector field $\tilde{Y} \in T\mathcal{X}$, with $\tilde{Y} \sim_\pi 0$, such that $(Y + \tilde{Y}, \beta) \in D$. Since $Y + \tilde{Y} \sim_\pi \hat{Y}$ this implies that $(\hat{Y}, \hat{\beta}) \in \hat{D}$. This shows that $\hat{D}^\perp \subset \hat{D}$. So $\hat{D} = \hat{D}^\perp$, which means that $\hat{D}$ is a generalized Dirac structure on $\hat{\mathcal{X}}$. For the proof that the closedness of D implies the closedness of $\hat{D}$ we refer to [14]. □

Remark 5. Take $\hat{F}_1, \hat{F}_2 \in \mathcal{A}_{\hat{D}}$, i.e. $(\hat{X}_1, d\hat{F}_1), (\hat{X}_2, d\hat{F}_2) \in \hat{D}$. Then (X_1, dF_1), $(X_2, dF_2) \in D$ for $X_j \sim_\pi \hat{X}_j$ and $F_j = \hat{F}_j \circ \pi,\ j = 1, 2$. So the bracket of admissible functions becomes

$$\{\hat{F}_1, \hat{F}_2\}_{\hat{D}}(\hat{x}) = \langle d\hat{F}_2, \hat{X}_1 \rangle(\hat{x}) = \langle dF_2, X_1 \rangle(x) = \{F_1, F_2\}_D(x),$$

where $\pi(x) = \hat{x}$. Equivalently

$$\{\hat{F}_1, \hat{F}_2\}_{\hat{D}} \circ \pi = \{\hat{F}_1 \circ \pi, \hat{F}_2 \circ \pi\}_D. \tag{36}$$

4.2 Reduction of implicit generalized Hamiltonian systems

In this section we will investigate the reduction possibilities of implicit Hamiltonian systems. We begin by stating the analogies of propositions 19, 20.

Consider an implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$. Let $P \in C^\infty(\mathcal{X})$ be a first integral of $(\mathcal{X}, D, H)$ as in definition 9, and consider the level set $\bar{\mathcal{X}} = \{x \in \mathcal{X} \mid P(x) = a\}$ for some $a \in \mathbb{R}$ such that $\bar{\mathcal{X}} \cap \mathcal{X}_c$ is nonempty. Then every solution of $(\mathcal{X}, D, H)$ starting in $\bar{\mathcal{X}}$ will remain in $\bar{\mathcal{X}}$. We can describe these solutions by using the induced Dirac structure on $\bar{\mathcal{X}}$.

Proposition 21. *Consider the assumptions described above. Let $D(\bar{x}) \cap E_s(\bar{x})$, $\bar{x} \in \bar{\mathcal{X}}$, be constant dimensional on $\bar{\mathcal{X}}$, where E_s is defined in (21). Then every solution of $(\mathcal{X}, D, H)$ lying in $\bar{\mathcal{X}}$ is a solution of the implicit generalized Hamiltonian system $(\bar{\mathcal{X}}, \bar{D}, \bar{H})$, where $\bar{D}$ is the generalized Dirac structure induced by D, see proposition 19, and $\bar{H} = \iota^* H$, i.e. the Hamiltonian H restricted to $\bar{\mathcal{X}}$.*

Proof. Let $x(t)$ be a solution of the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ contained in $\bar{\mathcal{X}}$, i.e.

$$(X_H, dH)(x(t)) \in D(x(t)), \quad \text{for all } t \in I,$$

where $X_H(x(t)) = \dot{x}(t)$ and $I \subset \mathbb{R}$ is the interval of existence of $x(t)$. Because P is a first integral it follows that $X_H(x(t))$ is tangent to $\bar{\mathcal{X}}$ at all times t, see also (13). Define $X_{\bar{H}}$ such that

$$T_{\bar{x}(t)}\iota \cdot X_{\bar{H}}(\bar{x}(t)) = X_H(x(t)), \ \forall t \in I, \tag{37}$$

where $\iota(\bar{x}(t)) = x(t)$. Take arbitrary $(\bar{Y}, \bar{\beta}) \in \bar{D}$. There exist $(Y, \beta) \in D$ such that $\bar{Y} \sim_\iota Y$ and $\bar{\beta} = \iota^*\beta$. Then

$$\left(\langle d\bar{H}, \bar{Y}\rangle + \langle \bar{\beta}, X_{\bar{H}}\rangle\right)(\bar{x}(t)) = \left(\langle \iota^* dH, \bar{Y}\rangle + \langle \iota^*\beta, X_{\bar{H}}\rangle\right)(\bar{x}(t)) = \left(\langle dH, Y\rangle + \langle \beta, X_H\rangle\right)(x(t)) = 0,$$

where in the last step we used that $(X_H, dH)(x(t)) \in D(x(t)) = D^\perp(x(t)) = [D(x(t))]^\perp$, $\forall t \in I$, by proposition 2. This shows that

$$(X_{\bar{H}}, d\bar{H})(\bar{x}(t)) \in [\bar{D}(\bar{x}(t))]^\perp = \bar{D}^\perp(\bar{x}(t)) = \bar{D}(\bar{x}(t)), \ \forall t \in I,$$

which implies that $\bar{x}(t)$ is a solution of $(\bar{\mathcal{X}}, \bar{D}, \bar{H})$. (Note that by (37) $\dot{\bar{x}}(t) = X_{\bar{H}}(\bar{x}(t))$.) □

This proposition can be easily extended to the case where we consider the level set $\bar{\mathcal{X}} = \{x \in \mathcal{X} \mid P_1(x) = a_1, \ldots, P_r(x) = a_r, \ (a_1, \ldots, a_r) \in \mathbb{R}^r\}$ of r independent first integrals $P_1, \ldots, P_r \in C^\infty(\mathcal{X})$ of $(\mathcal{X}, D, H)$.

Proposition 21 says that every solution of $(\mathcal{X}, D, H)$ lying in $\bar{\mathcal{X}}$ is a solution of $(\bar{\mathcal{X}}, \bar{D}, \bar{H})$. However, in general, $(\bar{\mathcal{X}}, \bar{D}, \bar{H})$ will generate more solutions, i.e. solutions that do not correspond to any solution of $(\mathcal{X}, D, H)$. This can be seen most easily in the classical case of reduction of a Hamiltonian system on a symplectic manifold N to a submanifold M of N. Consider a symplectic manifold N, i.e. a manifold N with a closed, nondegenerate, 2-form ω. On N the Hamiltonian system, with Hamiltonian function $H \in C^\infty(N)$, is given by the Hamiltonian vector field $X_H \in TN$ defined by

$$dH = \omega(X_H, \cdot) \tag{38}$$

[1, 7]. The Hamiltonian system on N projects to a Hamiltonian system on a submanifold $M \subset N$ by defining the new Hamiltonian to be $\bar{H} = \iota^* H$ and the corresponding 2-form to be $\bar{\omega} = \iota^*\omega$. Note that $\bar{\omega}$ is again a closed 2-form on M but in general it will be degenerate, meaning that it has a nontrivial kernel (we call M presymplectic). The 'new' Hamiltonian system on M is now given by the Hamiltonian vector field $X_{\bar{H}} \in TM$ defined by

$$d\bar{H} = \bar{\omega}(X_{\bar{H}}, \cdot). \tag{39}$$

Now, every solution of the Hamiltonian system defined by (38) lying in M (i.e. where $X_H \in TM$) is a solution of the Hamiltonian system defined by (39) (note that (38) and $X_H \in TM$ imply (39)). However, due to the fact that $\bar{\omega}$ has a nontrivial kernel, the Hamiltonian system defined by (39) generates more solutions than only those coming from solutions of (38). Indeed, if $X_H \in TM$ is a solution of (38) then $X_{\bar{H}} = X_H + Y$ is a solution of (39) for every $Y \in TM$ that lies in the kernel of $\bar{\omega}$.

An example where the above cannot happen, is when we restrict a Hamiltonian system on a Poisson manifold to a level set of a Casimir function. Then the solutions of the restricted system will all correspond to solutions of the original system. More generally we can say the following.

Proposition 22. *Consider an implicit generalized Hamiltonian system* $(\mathcal{X}, D, H)$*. Let* $C \in C^\infty(\mathcal{X})$ *be a Casimir function of* $(\mathcal{X}, D, H)$*, as in definition 11, and assume that* $dC \in \mathsf{P}_0$*. Consider the level set* $\bar{\mathcal{X}} = \{x \in \mathcal{X} \mid C(x) = a\}$ *for some* $a \in \mathbb{R}$ *such that* $\bar{\mathcal{X}} \cap \mathcal{X}_c$ *is nonempty. Then the solutions of* $(\mathcal{X}, D, H)$ *lying in* $\bar{\mathcal{X}}$ *are exactly the solutions of the implicit generalized Hamiltonian system* $(\bar{\mathcal{X}}, \bar{D}, \bar{H})$*, where* $\bar{D}$ *is the generalized Dirac structure induced by* D*, see proposition 19, and* $\bar{H} = \iota^* H$*, i.e. the Hamiltonian* H *restricted to* $\bar{\mathcal{X}}$*.*

Proof. First note that since $dC \in \mathsf{P}_0 = \operatorname{ann} \mathsf{G}_1$, $D(\bar{x}) \cap E_s(\bar{x}) = D(\bar{x})$, $\bar{x} \in \bar{\mathcal{X}}$, is constant dimensional on $\bar{\mathcal{X}}$. See the proof of proposition 21 to conclude that every solution $x(t)$ of $(\mathcal{X}, D, H)$ is a solution of $(\bar{\mathcal{X}}, \bar{D}, \bar{H})$.

Now, let $\bar{x}(t)$ be a solution of $(\bar{\mathcal{X}}, \bar{D}, \bar{H})$, i.e.

$$(X_{\bar{H}}, d\bar{H})(\bar{x}(t)) \in \bar{D}(\bar{x}(t)), \quad \text{for all } t \in I,$$

where $X_{\bar{H}}(\bar{x}(t)) = \dot{\bar{x}}(t)$. Define

$$X_H(x(t)) = T_{\bar{x}(t)}\iota \cdot X_{\bar{H}}(\bar{x}(t)), \ \forall t \in I, \tag{40}$$

where $x(t) = \iota(\bar{x}(t))$. Take arbitrary $(Y, \beta) \in D$. Because $dC \in \mathsf{P}_0$ it follows that $\langle dC, Y\rangle(x) = 0, \ \forall x \in \mathcal{X}$. This means that Y is tangent to $\bar{\mathcal{X}}$. Define $\bar{Y} \in T\bar{\mathcal{X}}$ such that

$$T_{\bar{x}}\iota \cdot \bar{Y}(\bar{x}) = Y(\iota(\bar{x})), \ \forall \bar{x} \in \bar{\mathcal{X}},$$

and $\bar{\beta} = \iota^*\beta$. Then

$$\begin{aligned}
&\left(\langle dH, Y\rangle + \langle \beta, X_H\rangle\right)(x(t)) = \\
&\langle dH(\iota(\bar{x}(t))), T_{\bar{x}(t)}\iota \cdot \bar{Y}(\bar{x}(t))\rangle + \langle \beta(\iota(\bar{x}(t))), T_{\bar{x}(t)}\iota \cdot X_{\bar{H}}(\bar{x}(t))\rangle = \\
&\left(\langle \iota^* dH, \bar{Y}\rangle + \langle \iota^*\beta, X_{\bar{H}}\rangle\right)(\bar{x}(t)) = \\
&\left(\langle d\bar{H}, \bar{Y}\rangle + \langle \bar{\beta}, X_{\bar{H}}\rangle\right)(\bar{x}(t)) = 0,
\end{aligned}$$

where in the last step we used that $(X_{\bar{H}}, d\bar{H})(\bar{x}(t)) \in \bar{D}(\bar{x}(t)) = \bar{D}^\perp(\bar{x}(t)) = [\bar{D}(\bar{x}(t))]^\perp, \ \forall t \in I$. This shows that

$$(X_H, dH)(x(t)) \in [D(x(t))]^\perp = D^\perp(x(t)) = D(x(t)), \ \forall t \in I.$$

Since $\dot{x}(t) = X_H(x(t))$ by (40) this means that $x(t)$ is a solution of $(\mathcal{X}, D, H)$.

□

Of course, this proposition can also be easily extended to the case of multiple independent Casimir functions.

To state the analog of proposition 20 we first need the following

Definition 12. *We will call a vector field* $f \in T\mathcal{X}$ *a strong symmetry of the implicit generalized Hamiltonian system* $(\mathcal{X}, D, H)$ *if* f *is a symmetry of the generalized Dirac structure* D *(as in definition 8) and* f *leaves* H *invariant everywhere, i.e.* $L_f H(x) = 0, \ \forall x \in \mathcal{X}$ *(note the difference with definition 10).* G *is called a strong symmetry Lie group of* $(\mathcal{X}, D, H)$ *if* G *is a symmetry Lie group of* D *(as in proposition 20) and every infinitesimal generator* $\xi_{\mathcal{X}}$*,* $\xi \in \mathcal{G}$*, leaves* H *invariant everywhere.*

Proposition 23. *[14] Consider an implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$. Let G be a strong symmetry Lie group of $(\mathcal{X}, D, H)$ and assume that $V + \mathsf{G}_0$ and $D \cap E_q$ are constant dimensional. Then $(\mathcal{X}, D, H)$ projects to the implicit generalized Hamiltonian system $(\hat{\mathcal{X}}, \hat{D}, \hat{H})$, where $\hat{\mathcal{X}} = \mathcal{X}/G$, $\hat{D}$ is the generalized Dirac structure induced by D, see proposition 20, and the Hamiltonian $\hat{H}$ is such that $H = \hat{H} \circ \pi$ (note that G leaves H invariant so $\hat{H}$ is well defined).*
More explicitly: Every solution $\hat{x}(t)$ of $(\hat{\mathcal{X}}, \hat{D}, \hat{H})$ is (locally) the projection under π of a solution $x(t)$ of $(\mathcal{X}, D, H)$. Conversely, let $x(t)$ be a solution of $(\mathcal{X}, D, H)$ along a projectable vector field X_H, that is, assume that there exists a vector field $X \in T\mathcal{X}$ such that $X \sim_\pi \hat{X}$ for some $\hat{X} \in T\hat{\mathcal{X}}$ and $X(x(t)) = X_H(x(t))$, $\forall t \in I$, then $x(t)$ can be projected to a solution $\hat{x}(t)$ of $(\hat{\mathcal{X}}, \hat{D}, \hat{H})$.

Proof. Let $\hat{x}(t)$ be a solution of $(\hat{\mathcal{X}}, \hat{D}, \hat{H})$, i.e.

$$(X_{\hat{H}}, d\hat{H})(\hat{x}(t)) \in \hat{D}(\hat{x}(t)), \quad \text{for all } t \in I,$$

where $X_{\hat{H}}(x(t)) = \dot{\hat{x}}(t)$ and $I \subset \mathbb{R}$ is the interval of existence of $\hat{x}(t)$. Define $\hat{A} = \{\hat{x}(t) \mid t \in I\}$, and assume that $\hat{A}$ is a closed subset of $\hat{X}$. If this is not the case, for instance if $\hat{x}(t)$ converges asymptotically to an equilibrium point, then by defining $\hat{A}$ on any closed interval $I' \subset I$ (i.e. by considering $\hat{x}(t)$ only "locally") $\hat{A}$ can be made into a closed subset of $\hat{X}$. Then it follows that

$$(X_{\hat{H}}, d\hat{H})(\hat{x}) \in \hat{D}(\hat{x}), \quad \forall \hat{x} \in \hat{A}.$$

Because $\hat{D}$ is a smooth bundle the pair $(X_{\hat{H}}, d\hat{H})$ can be locally extended to a pair in $\hat{D}$, and therefore also, by the Smooth Tietze Extension Theorem [2], globally extended to a pair $(\hat{X}, \hat{\alpha}) \in \hat{D}$. By definition of $\hat{D}$ there exists a pair $(X, \alpha) \in D$ where

$$X \sim_\rho \hat{X}, \ \alpha = \rho^* \hat{\alpha}. \tag{41}$$

Because

$$\rho^* \hat{\alpha}(x) = \hat{\alpha}(\rho(x))(T_x\rho\ \cdot) = d\hat{H}(\rho(x))(T_x\rho\ \cdot) = \rho^* d\hat{H}(x) = dH(x),$$

for all $x \in \mathcal{X}$ such that $\rho(x) = \hat{x} \in \hat{A}$, it follows that $(X, dH)(x) \in D(x)$ for all $x \in \mathcal{X}$ such that $\rho(x) = \hat{x} \in \hat{A}$. Equivalently, let $x(t)$ be such that $\dot{x}(t) = X(x(t))$, then $\rho(x(t)) = \hat{x}(t)$ (because of (41)), and

$$(X_H, dH)(x(t)) \in D(x(t)), \quad \text{for all } t \in I,$$

where we wrote X_H for X. This means that $x(t)$ is a solution of the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$.

Conversely, let $x(t)$ be a solution of $(\mathcal{X}, D, H)$, i.e.

$$(X_H, dH)(x(t)) \in D(x(t)), \quad \text{for all } t \in I,$$

where $X_H(x(t)) = \dot{x}(t)$ and $I \subset \mathbb{R}$ is the interval of existence of $x(t)$. Assume that $x(t)$ is the flow of a projectable vector field, that is, assume that there exists a vector field $X \in T\mathcal{X}$ such that $X \sim_\pi \hat{X}$ for some $\hat{X} \in T\hat{\mathcal{X}}$ and $X(x(t)) = X_H(x(t))$, $\forall t \in I$.
Take arbitrary $(\hat{Y}, \hat{\beta}) \in \hat{D}$. There exist $(Y, \beta) \in D$ such that $Y \sim_\pi \hat{Y}$ and $\beta = \pi^* \hat{\beta}$. Let $\hat{x}(t) = \pi(x(t))$, then

$$\begin{aligned}
&(\langle d\hat{H}, \hat{Y}\rangle + \langle \hat{\beta}, \hat{X}\rangle)\ (\hat{x}(t)) = \\
&\langle d\hat{H}(\hat{x}(t)), T_{x(t)}\pi \cdot Y(x(t))\rangle + \langle \hat{\beta}(\hat{x}(t)), T_{x(t)}\pi \cdot X(x(t))\rangle = \\
&(\langle dH, Y\rangle + \langle \beta, X\rangle)\ (x(t)) = \\
&(\langle dH, Y\rangle + \langle \beta, X_H\rangle)\ (x(t)) = 0,
\end{aligned}$$

where in the last step we used that $(X_H, dH)(x(t)) \in D(x(t)) = D^{\perp}(x(t)) = [D(x(t))]^{\perp}$, $\forall t \in I$. This shows that

$$(d\hat{H}, X_{\hat{H}})(\hat{x}(t)) \in [\hat{D}(\hat{x}(t))]^{\perp} = \hat{D}^{\perp}(\hat{x}(t)) = \hat{D}(\hat{x}(t)),\ \forall t \in I,$$

where we wrote $X_{\hat{H}}$ for $\hat{X}$. From the fact that $X \sim_{\pi} \hat{X}$ it follows that $\dot{\hat{x}}(t) = X_{\hat{H}}(\hat{x}(t))$, $\forall t \in I$, so $\hat{x}(t)$ is a solution of $(\hat{\mathcal{X}}, \hat{D}, \hat{H})$. □

Remark 6. When assumption 5 is satisfied, it can be shown that $x(t)$ always is the flow of a projectable vector field, see section 7.

Example 7. In proposition 23 we needed the assumption that a solution of $(\mathcal{X}, D, H)$ is a solution along a projectable vector field X_H in order to project to a solution of the reduced system $(\hat{\mathcal{X}}, \hat{D}, \hat{H})$. In general, not every solution of $(\mathcal{X}, D, H)$ projects to a solution of $(\hat{\mathcal{X}}, \hat{D}, \hat{H})$. This can be seen in this example. Consider the following generalized Dirac structure on $\mathcal{X} = \mathbb{R}^3$

$$D = \{(X, \alpha) \in T\mathcal{X} \oplus T^*\mathcal{X} \mid X \in \ker \mathsf{P}_1,\ \alpha \in \mathsf{P}_1\},$$

with

$$\mathsf{P}_1 = \mathrm{span}_{C^\infty(\mathcal{X})}\ \{dx_3\},\ \text{i.e. } \mathsf{G}_1 = \mathsf{G}_0 = \ker \mathsf{P}_1 = \mathrm{span}_{C^\infty(\mathcal{X})}\ \{\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}\}.$$

Take an arbitrary pair $(X, \alpha) = (h_1\frac{\partial}{\partial x_1} + h_2\frac{\partial}{\partial x_2}, h_3 dx_3) \in D$, with $h_1, h_2, h_3 \in C^\infty(\mathcal{X})$. Then it can easily be calculated that

$$(L_{\frac{\partial}{\partial x_1}} X, L_{\frac{\partial}{\partial x_1}} \alpha) = ((L_{\frac{\partial}{\partial x_1}} h_1)\frac{\partial}{\partial x_1} + (L_{\frac{\partial}{\partial x_1}} h_2)\frac{\partial}{\partial x_2}, (L_{\frac{\partial}{\partial x_1}} h_3)\, dx_3) \in D,$$

so $\frac{\partial}{\partial x_1}$ is a symmetry of D. Let the Hamiltonian function be of the form $H(x_1, x_2, x_3) = \hat{H}(x_3)$, then also $L_{\frac{\partial}{\partial x_1}} H = 0$, so $\frac{\partial}{\partial x_1}$ is a strong symmetry of $(\mathcal{X}, D, H)$. An arbitrary solution $x(t)$ of $(\mathcal{X}, D, H)$ is the flow along a vector field

$$X_H(x_1, x_2, x_3) = h_1(x_1, x_2, x_3)\frac{\partial}{\partial x_1} + h_2(x_1, x_2, x_3)\frac{\partial}{\partial x_2}, \tag{42}$$

with $h_1, h_2 \in C^\infty(\mathcal{X})$. However, the vector field X_H in (42) will only project to a vector field $\hat{X} = X_{\hat{H}}$ on $\mathcal{X}/\frac{\partial}{\partial x_1} \simeq \mathbb{R}^2$ if it has the form

$$X_H(x_1, x_2, x_3) = h_1(x_1, x_2, x_3)\frac{\partial}{\partial x_1} + h_2(x_2, x_3)\frac{\partial}{\partial x_2}$$

(note the difference, h_2 should not depend on x_1). In that case

$$X_{\hat{H}} = h_2(x_2, x_3)\frac{\partial}{\partial x_2}.$$

The reduced generalized Dirac structure is

$$\hat{D} = \{(\hat{X}, \hat{\alpha}) \in T\hat{\mathcal{X}} \oplus T^*\hat{\alpha} \mid \hat{X} \in \ker \hat{\mathsf{P}}_1,\ \hat{\alpha} \in \hat{\mathsf{P}}_1\},$$

with

$$\hat{\mathsf{P}}_1 = \mathrm{span}_{C^\infty(\hat{\mathcal{X}})}\{dx_3\},\ \text{i.e. } \hat{\mathsf{G}}_1 = \hat{\mathsf{G}}_0 = \ker \hat{\mathsf{P}}_1 = \mathrm{span}_{C^\infty(\hat{\mathcal{X}})}\ \{\frac{\partial}{\partial x_2}\},$$

and $(X_{\hat{H}}, d\hat{H}) \in \hat{D}$. With respect to remark 6, note that assumption 5 is not satisfied.

Now, after these preliminaries, we are ready to investigate what is going to be the main result of our work.

5 Main result

In this section we will derive our main result on reduction of implicit generalized Hamiltonian systems. This result will generalize the 'classical' reduction theorems of explicit Hamiltonian systems described in [1, 7, 9, 12].

Consider an implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ on an n-dimensional manifold $\mathcal{X}$, with generalized Dirac structure D and Hamiltonian function $H \in C^\infty(\mathcal{X})$. Suppose the system has r independent first integrals $P_i \in C^\infty(\mathcal{X}),\ i = 1, \ldots, r$, and suppose there exist corresponding independent vector fields $X_{P_i} \in T\mathcal{X}$, i.e. $(X_{P_i}, dP_i) \in D,\ i = 1, \ldots, r$, such that each X_{P_i} is a strong symmetry of $(\mathcal{X}, D, H)$. We assume that P_i and X_{P_i} satisfy the following conditions

$$\{P_i, P_j\}_D = \sum_{k=1}^{r} c_{ij}^k P_k, \tag{43}$$

and

$$[X_{P_i}, X_{P_j}] = \sum_{k=1}^{r} c_{ij}^k X_{P_k}, \tag{44}$$

where $c_{ij}^k \in \mathbb{R}$ are constants, $i, j = 1, \ldots, r$.

Remark 7. Note that in the case of a Poisson structure on $\mathcal{X}$, which satisfies the Jacobi identity (i.e. which is closed), (43) implies (44) (in the case of a symplectic structure on $\mathcal{X}$, (43) and (44) are equivalent). However, in the case of a Dirac structure (i.e. which is closed) on $\mathcal{X}$ (43) implies only $[X_{P_i}, X_{P_j}] = \sum_{k=1}^r c_{ij}^k X_{P_k} + Z_{ij}$ where $Z_{ij} \in \mathsf{G}_0$.

Because of condition (44) there exists an r-dimensional Lie group G with corresponding Lie algebra $\mathcal{G}$ for which the infinitesimal generators $(\xi_i)_\mathcal{X} = X_{P_i},\ i = 1, \ldots, r$, where $\{\xi_1, \ldots, \xi_r\}$ is a basis of $\mathcal{G}$ [12]. It follows that G is a strong symmetry Lie group of $(\mathcal{X}, D, H)$. Let $\{\mu_1, \ldots, \mu_r\}$ be a basis of $\mathcal{G}^*$. We define the following map from $\mathcal{X}$ to $\mathcal{G}^*$, also called *momentum map* [12, 1, 7],

$$P(x) = \sum_{i=1}^{r} P_i(x)\mu_i.$$

Proposition 24. *The momentum map P is Ad^*-equivariant, that is,*

$$P(\phi_g(x)) = Ad_g^*(P(x)),$$

for all $x \in \mathcal{X},\ g \in G$, where Ad^* is the coadjoint action corresponding to the Lie group G.

Proof. The proof equals the proof in [12], see also [1, 7], we only have to consider the bracket of admissible functions $\{\cdot, \cdot\}_D$ in stead of the Poisson bracket $\{\cdot, \cdot\}$. □

Now we will describe the reduction possibilities of the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ admitting the strong symmetry Lie group G corresponding to the the first integrals $P_1, \ldots, P_r$. There are two ways, which in a sense are dual, to reduce the Hamiltonian system. The first one is to begin by reducing the Hamiltonian system to a level set $P^{-1}(\mu)$ of the first integrals, using proposition 21. At this point the resulting implicit generalized Hamiltonian system will have some symmetry remaining from the symmetry group G, however, in general it will not be the whole group G but only a subgroup G_μ of G. Then we can use proposition 23 to further reduce the Hamiltonian system to an implicit generalized Hamiltonian system on the quotient manifold

$P^{-1}(\mu)/G_\mu$. The second way to reduce the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ is by beginning to reduce the Hamiltonian system to an implicit generalized Hamiltonian system on the quotient manifold $\mathcal{X}/G$, as in proposition 23. The resulting Hamiltonian system will have some first integrals (actually these will be Casimir functions) remaining from $P_1, \dots, P_r$ which we can use to further reduce the Hamiltonian system to a level set of these first integrals, proposition 21. The main result of our work will state that these two ways of reducing the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ will result in the same reduced implicit generalized Hamiltonian system (up to diffeomorphism). This is a generalization of the classical reduction theorems of [1, 7, 9, 12].

Reduction first using the first integrals, then a remaining symmetry group

Consider the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ with corresponding independent first integrals $P_1, \dots, P_r$ and strong symmetry Lie group G as described previously. Because $P_1, \dots, P_r$ are first integrals the solutions of $(\mathcal{X}, D, H)$ will live on some level set $\bar{\mathcal{X}} = \{x \in \mathcal{X} \mid P_1(x) = a_1, \dots, P_r(x) = a_r,\ (a_1, \dots, a_r) \in \mathbb{R}^r\}$, $\bar{\mathcal{X}} \cap \mathcal{X}_c$ nonempty. Note that by using the momentum map P we can denote this level set by $\bar{\mathcal{X}} = P^{-1}(\mu)$ for some $\mu \in \mathcal{G}^*$. Using proposition 21, assuming $D(\bar{x}) \cap E_s(\bar{x})$, $\bar{x} \in \bar{\mathcal{X}}$, is constant dimensional on $\bar{\mathcal{X}}$, we can reduce the Hamiltonian system to an implicit generalized Hamiltonian system $(P^{-1}(\mu), \bar{D}, \bar{H})$ on $P^{-1}(\mu)$, where $\bar{D}$ is the generalized Dirac structure induced by D, and $\bar{H} = \iota_1^* H$ is the Hamiltonian function on $P^{-1}(\mu)$, $\iota_1 : P^{-1}(\mu) \to \mathcal{X}$ being the inclusion map. Consider the subgroup

$$G_\mu = \{g \in G \mid Ad_g^*(\mu) = \mu\}, \tag{45}$$

or equivalently

$$G_\mu = \{g \in G \mid \phi_g(P^{-1}(\mu)) \subset P^{-1}(\mu)\}.$$

G_μ is a subgroup of G and therefore a Lie group itself.

Lemma 25. *G_μ is a strong symmetry Lie group of $(P^{-1}(\mu), \bar{D}, \bar{H})$.*

Proof. Consider $\bar{X} = (\xi_\mu)_{P^{-1}(\mu)}$ for some $\xi_\mu \in \mathcal{G}_\mu$. Then $\bar{X}$ is ι_1-related to $X = (\xi_\mu)_{\mathcal{X}}$. Now, let $(\bar{Y}, \bar{\beta}) \in \bar{D}$, then $\bar{Y} \sim_{\iota_1} Y$ and $\bar{\beta} = \iota_1^* \beta$, $(Y, \beta) \in D$, see (18). By proposition 7, $L_{\bar{X}} \bar{Y} = [\bar{Y}, \bar{X}] \sim_{\iota_1} [Y, X] = L_X Y$. Furthermore

$$L_{\bar{X}} \bar{\beta} = L_{\bar{X}} \iota_1^* \beta = \iota_1^* L_X \beta.$$

Now, X is a symmetry of D which means that $(L_X Y, L_X \beta) \in D$, and it follows that also $(L_{\bar{X}} \bar{Y}, L_{\bar{X}} \bar{\beta}) \in \bar{D}$, so $\bar{X}$ is a symmetry of $\bar{D}$. Because

$$L_{\bar{X}} \bar{H}_1 = L_{\bar{X}} \iota_1^* H = \iota_1^* L_X H = 0,$$

$\bar{X}$ is a strong symmetry of $(P^{-1}(\mu), \bar{D}, \bar{H})$. ◻

G_μ is called the residual symmetry group. Now we can use proposition 23 (in theorem 28 we will show that the assumptions of proposition 23 are satisfied) to further reduce the Hamiltonian system $(P^{-1}(\mu), \bar{D}, \bar{H})$ to an implicit generalized Hamiltonian system $(P^{-1}(\mu)/G_\mu, \hat{\bar{D}}, \hat{\bar{H}})$ on the quotient manifold $P^{-1}(\mu)/G_\mu$, where $\hat{\bar{D}}$ is the generalized Dirac structure induced by $\bar{D}$, and $\hat{\bar{H}}$ is the Hamiltonian function on $P^{-1}(\mu)/G_\mu$, with $\bar{H} = \hat{\bar{H}} \circ \pi_\mu$, where $\pi_\mu : P^{-1}(\mu) \to P^{-1}(\mu)/G_\mu$ is the projection map.

Reduction first using the symmetry group, then the remaining first integrals

Again, consider the same implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ with corresponding independent first integrals $P_1, \ldots, P_r$ and strong symmetry Lie group G as we started with in the reduction process described above. Contrary to starting with reduction to a level set of the first integrals, as we did above, we will now reduce the Hamiltonian system $(\mathcal{X}, D, H)$ by first reducing it to the quotient manifold $\mathcal{X}/G$. Assume that $V + \mathsf{G}_0$ and $D \cap E_q$ are constant dimensional. Using proposition 23 this gives us an implicit generalized Hamiltonian system $(\mathcal{X}/G, \hat{D}, \hat{H})$ on $\mathcal{X}/G$, where $\hat{D}$ is the generalized Dirac structure induced by D, and $\hat{H}$ is the Hamiltonian function on $\mathcal{X}/G$, with $H = \hat{H} \circ \pi$. Here, $\pi : \mathcal{X} \to \mathcal{X}/G$ is the projection map.
Consider the quotient manifold $\hat{\mathcal{G}}^* = \mathcal{G}^*/G$ of coadjoint orbits $\mathcal{O}_\mu$ in $\mathcal{G}^*$, along with the projection map $\varpi : \mathcal{G}^* \to \hat{\mathcal{G}}^*$. A coadjoint orbit is defined as

$$\mathcal{O}_\mu = \{Ad_g^*(\mu) \mid g \in G\}, \quad \mu \in \mathcal{G}^*. \tag{46}$$

Define the map $\hat{P} : \mathcal{X}/G \to \hat{\mathcal{G}}^*$ by [7]

$$\hat{P} \circ \pi = \varpi \circ P. \tag{47}$$

Then $\hat{P}$ is well defined, because take arbitrary $\hat{x} \in \mathcal{X}/G$ and $x_1, x_2 \in \mathcal{X}$ such that $\pi(x_1) = \pi(x_2) = \hat{x}$. This means that $x_1 = \phi_g(x_2)$ for some $g \in G$. Then

$$P(x_1) = P(\phi_g(x_2)) = Ad_g^*(P(x_2))$$

by Ad^*-equivariance of P (proposition 24). This implies that

$$P(x_1) \in \mathcal{O}_{P(x_2)},$$

and therefore that

$$\varpi(P(x_1)) = \varpi(P(x_2)),$$

so $\hat{P}(\hat{x})$ is well defined. Furthermore $\hat{P}$ is a conserved quantity along solutions of $(\mathcal{X}/G, \hat{D}, \hat{H})$. Indeed, let $\hat{x}(t)$ be a solution of $(\mathcal{X}/G, \hat{D}, \hat{H})$. Then there exists (locally) a solution $x(t)$ of $(\mathcal{X}, D, H)$ such that $\pi(x(t)) = \hat{x}(t)$, see proposition 23. The corresponding vector fields are related, i.e. $X_H \sim_\pi X_{\hat{H}}$. Then

$$\begin{aligned} &\langle d\hat{P}, X_{\hat{H}}\rangle(\hat{x}(t)) = \langle \pi^* d\hat{P}, X_H\rangle(x(t)) = \langle d(\hat{P} \circ \pi), X_H\rangle(x(t)) = \\ &= \langle d(\varpi \circ P), X_H\rangle(x(t)) = d\varpi(\langle dP, X_H\rangle)(x(t)) = 0, \end{aligned} \tag{48}$$

where the last step follows from the fact that P is a first integral of $(\mathcal{X}, D, H)$. Actually, $\hat{P}$ is a Casimir function, because take arbitrary $\hat{H} \in C^\infty(\mathcal{X}/G)$, then $\hat{H}$ corresponds to a G-invariant function $H \in C^\infty(\mathcal{X})$, by $H = \hat{H} \circ \pi$, for which again P will be a first integral, and so by (48) $\hat{P}$ will be conserved along solutions of $(\mathcal{X}/G, \hat{D}, \hat{H})$. In section 6 we will elaborate a bit more on the map $\hat{P}$. In particular we will show that "locally $d\hat{P} \in \hat{\mathsf{P}}_0$". Using proposition 22 (see also section 6) we can restrict the Hamiltonian system $(\mathcal{X}/G, \hat{D}, \hat{H})$ to an implicit generalized Hamiltonian system $(\hat{P}^{-1}(\hat{\mu}), \bar{\hat{D}}, \bar{\hat{H}})$ on a level set $\hat{P}^{-1}(\hat{\mu})$ of $\hat{P}$, for some $\hat{\mu} \in \hat{\mathcal{G}}^*$ (to be consistent with the procedure above we should take $\hat{\mu} = \varpi(\mu)$). Here, $\bar{\hat{D}}$ is the generalized Dirac structure induced by $\hat{D}$, $\bar{\hat{H}} = \iota_2^* \hat{H}$ is the Hamiltonian function on $\hat{P}^{-1}(\hat{\mu})$ and $\iota_2 : \hat{P}^{-1}(\hat{\mu}) \to \mathcal{X}/G$ is the inclusion map.

Consider the two reduction procedures described above.

Lemma 26. *There exists a diffeomorphism ψ from $P^{-1}(\mu)/G_\mu$ to $\hat{P}^{-1}(\hat{\mu})$, with $\hat{\mu} = \varpi(\mu)$, such that the following diagram commutes:*

$$\begin{array}{ccccc} & & \mathcal{X} & & \\ & \iota_1 \nearrow & & \searrow \pi & \\ P^{-1}(\mu) & & & & \mathcal{X}/G \\ \pi_\mu \downarrow & & & & \uparrow \iota_2 \\ P^{-1}(\mu)/G_\mu & & \psi \longrightarrow & & \hat{P}^{-1}(\hat{\mu}) \end{array} \tag{49}$$

Proof. The proof is based on [7]. First we prove that there exists a diffeomorphism $\psi : P^{-1}(\mu)/G_\mu \to \pi(P^{-1}(\mu))$. Note that $P^{-1}(\mu)$ is a submanifold of $\mathcal{X}$ so $\pi(P^{-1}(\mu))$ makes sense and is a subspace of $\mathcal{X}/G$ (formally, we should write $\pi(\iota_1(P^{-1}(\mu)))$ in stead of $\pi(P^{-1}(\mu))$).
Now, define $\psi : P^{-1}(\mu)/G_\mu \to \pi(P^{-1}(\mu))$ as follows: Let $\hat{\bar{x}} \in P^{-1}(\mu)/G_\mu$. There exists an $\bar{x} \in P^{-1}(\mu)$ such that $\pi_\mu(\bar{x}) = \hat{\bar{x}}$. Define $\psi(\hat{\bar{x}}) = \pi(\bar{x})$. To see that ψ is well defined, let $\bar{x}' \in P^{-1}(\mu)$ be another element such that $\pi_\mu(\bar{x}') = \hat{\bar{x}}$. Then there exists a $g \in G_\mu$ such that $\phi_g(\bar{x}) = \bar{x}'$ and it follows that $\pi(\bar{x}) = \pi(\bar{x}')$, so ψ is well defined. We have to proof that ψ is a diffeomorphism. The fact that ψ is surjective is trivial. Now, let $\hat{\bar{x}}_1, \hat{\bar{x}}_2 \in P^{-1}(\mu)/G_\mu$ be such that $\psi(\hat{\bar{x}}_1) = \psi(\hat{\bar{x}}_2)$. Then $\psi(\hat{\bar{x}}_1) = \pi(\bar{x}_1)$ and $\psi(\hat{\bar{x}}_2) = \pi(\bar{x}_2)$ for $\bar{x}_1, \bar{x}_2 \in P^{-1}(\mu)$ with $\pi_\mu(\bar{x}_1) = \hat{\bar{x}}_1$ and $\pi_\mu(\bar{x}_2) = \hat{\bar{x}}_2$. So $\pi(\bar{x}_1) = \pi(\bar{x}_2)$ and therefore there exists a $g \in G$ such that $\phi_g(\bar{x}_1) = \bar{x}_2$. From Ad^*-equivariance of P, proposition 24, it follows that $g \in G_\mu$. Indeed,

$$Ad_g^*(\mu) = Ad_g^*(P(\bar{x}_1)) = P(\phi_g(\bar{x}_1)) = P(\bar{x}_2) = \mu,$$

and comparing with (45) gives that $g \in G_\mu$. But $\phi_g(\bar{x}_1) = \bar{x}_2$ for some $g \in G_\mu$ implies that $\pi_\mu(\bar{x}_1) = \pi_\mu(\bar{x}_2)$ and so $\hat{\bar{x}}_1 = \hat{\bar{x}}_2$. That means that ψ is injective. So ψ is bijective and because we assume that all maps are smooth it follows that ψ is a diffeomorphism.
Secondly, we prove that $\pi(P^{-1}(\mu)) = \hat{P}^{-1}(\hat{\mu})$. $\pi(P^{-1}(\mu)) \subset \hat{P}^{-1}(\hat{\mu})$ is easy and follows directly from (47). We prove the converse inclusion. Take arbitrary $\hat{x} \in \hat{P}^{-1}(\hat{\mu}) \subset \mathcal{X}/G$ and let $x \in \mathcal{X}$ be such that $\pi(x) = \hat{x}$. Then by (47)

$$\varpi(\mu) = \hat{\mu} = \hat{P}(\pi(x)) = \varpi(P(x)),$$

which implies that $\mu \in \mathcal{O}_{P(x)}$, so there exists a $g \in G$ such that $Ad_g^*(P(x)) = \mu$ by (46). However, by Ad^*-equivariance of P this means that $P(\phi_g(x)) = \mu$, so $\phi_g(x) \in P^{-1}(\mu)$. Furthermore $\pi(\phi_g(x)) = \pi(x) = \hat{x}$. This proves the converse inclusion. □

Remark 8. A nice interpretation of $\hat{P}^{-1}(\hat{\mu})$ is given in the fact that it is equivalent to the quotient space $P^{-1}(\mathcal{O}_\mu)/G$, as can be easily seen. Lemma 26 then states that $P^{-1}(\mu)/G_\mu$ is diffeomorphic to $P^{-1}(\mathcal{O}_\mu)/G$, which is the famous *Orbit Reduction Theorem* [8].

Main result

Definition 13. *Let M and N be two manifolds, and let $\tau : M \to N$ be a diffeomorphism. Let D_M be a (generalized) Dirac structure on M and let D_N be a (generalized) Dirac structure on N. Then τ is called a Dirac isomorphism if*

$$(X, \alpha) \in D_M \Longleftrightarrow (\tau_* X, (\tau^*)^{-1}\alpha) \in D_N. \tag{50}$$

In this case we call D_M and D_N isomorphic, denoted by $D_M \cong D_N$.

Remark 9. Let D_M and D_N be isomorphic. It is very easy to prove that D_M is closed if and only if D_N is closed.

Note that by (9) every symmetry $\phi : \mathcal{X} \to \mathcal{X}$ of a generalized Dirac structure D is a Dirac isomorphism.

Recall the two possible reduction procedures described above. The first one starts with the reduction of $(\mathcal{X}, D, H)$ to a level set of the first integrals, and after factoring out the residual symmetry group results in the implicit generalized Hamiltonian system $(P^{-1}(\mu)/G_\mu, \hat{\bar{D}}, \hat{\bar{H}})$. The second one starts with the reduction of $(\mathcal{X}, D, H)$ by factoring out the symmetry group, and after restriction to the level set of the remaining Casimirs results in the implicit generalized Hamiltonian system $(\hat{P}^{-1}(\hat{\mu}), \bar{\hat{D}}, \bar{\hat{H}})$. In lemma 26 it is shown that there exists a diffeomorphism $\psi : P^{-1}(\mu)/G_\mu \to \hat{P}^{-1}(\hat{\mu})$.

Theorem 27. *ψ is a Dirac isomorphism. That is, $\hat{\bar{D}}$ and $\bar{\hat{D}}$ are isomorphic, $\hat{\bar{D}} \cong \bar{\hat{D}}$.*

Proof. First, notice that it is sufficient to prove that

$$(\hat{\bar{X}}, \hat{\bar{\alpha}}) \in \hat{\bar{D}} \Longrightarrow (\psi_* \hat{\bar{X}}, (\psi^*)^{-1} \hat{\bar{\alpha}}) \in \bar{\hat{D}}. \tag{51}$$

For assume that (51) holds. Being Dirac structures, $\hat{\bar{D}}$ and $\bar{\hat{D}}$ are (pointwise) linear spaces. Define

$$\psi(\hat{\bar{D}}) := \{(\psi_* \hat{\bar{X}}, (\psi^*)^{-1} \hat{\bar{\alpha}}) \mid (\hat{\bar{X}}, \hat{\bar{\alpha}}) \in \hat{\bar{D}}\}.$$

Since ψ_* and ψ^* are linear mappings, $\psi(\hat{\bar{D}})$ is also a linear space. By (51), $\psi(\hat{\bar{D}}) \subset \bar{\hat{D}}$. However, because ψ is a diffeomorphism the map $(\psi_* \cdot, (\psi^*)^{-1} \cdot)$ is a bijection. Therefore

$$\dim \psi(\hat{\bar{D}})(\bar{\hat{x}}) = \dim \hat{\bar{D}}(\hat{\bar{x}}) = \dim P^{-1}(\mu)/G_\mu = \dim \hat{P}^{-1}(\hat{\mu}) = \dim \bar{\hat{D}}(\bar{\hat{x}}),$$

$\forall\, \bar{\hat{x}} \in \hat{P}^{-1}(\hat{\mu})$, $\hat{\bar{x}} = \psi^{-1}(\bar{\hat{x}})$, and it follows that actually $\psi(\hat{\bar{D}}) = \bar{\hat{D}}$. (50) now follows immediately.

We prove (51). Suppose $(\hat{\bar{X}}, \hat{\bar{\alpha}}) \in \hat{\bar{D}}$, we prove that $(\psi_* \hat{\bar{X}}, (\psi^*)^{-1} \hat{\bar{\alpha}}) \in \bar{\hat{D}}^\perp = \bar{\hat{D}}$.

The pair $(\hat{\bar{X}}, \hat{\bar{\alpha}}) \in \hat{\bar{D}}$ corresponds to pairs

- $(\bar{X}, \bar{\alpha}) \in \bar{D}$ with $\bar{X} \sim_{\pi_\mu} \hat{\bar{X}}$, $\bar{\alpha} = \pi_\mu^* \hat{\bar{\alpha}}$,
- $(X, \alpha) \in D$ with $\bar{X} \sim_{\iota_1} X$, $\bar{\alpha} = \iota_1^* \alpha$.

Now, take an arbitrary pair $(\bar{\hat{Y}}, \bar{\hat{\beta}}) \in \bar{\hat{D}}$. This corresponds to pairs

- $(\hat{Y}, \hat{\beta}) \in \hat{D}$ with $\bar{\hat{Y}} \sim_{\iota_2} \hat{Y}$, $\bar{\hat{\beta}} = \iota_2^* \hat{\beta}$,
- $(Y, \beta) \in D$ with $Y \sim_\pi \hat{Y}$, $\beta = \pi^* \hat{\beta}$.

Well, for arbitrary $\bar{\hat{x}} \in \hat{J}^{-1}(\hat{\mu})$ we calculate

$$\langle (\psi^*)^{-1} \hat{\bar{\alpha}}, \bar{\hat{Y}} \rangle (\bar{\hat{x}}) + \langle \bar{\hat{\beta}}, \psi_* \hat{\bar{X}} \rangle (\bar{\hat{x}}) \tag{52}$$

First we work out the first term in the above equation. By definition

$$\langle (\psi^*)^{-1} \hat{\bar{\alpha}}, \bar{\hat{Y}} \rangle (\bar{\hat{x}}) = \langle \hat{\bar{\alpha}}(\hat{\bar{x}}), T_{\bar{\hat{x}}} \psi^{-1} \cdot \bar{\hat{Y}}(\bar{\hat{x}}) \rangle \tag{53}$$

where $\hat{\bar{x}} = \psi^{-1}(\bar{\hat{x}})$ (and note that $(\psi^*)^{-1} = (\psi^{-1})^*$). Now, $T_{\bar{\hat{x}}}\psi^{-1} \cdot \bar{\hat{Y}}(\bar{\hat{x}})$ is a tangent vector to $J^{-1}(\mu)/G_\mu$ at the point $\hat{\bar{x}}$. Because π_μ and therefore $T\pi_\mu$ is surjective, there exists a point $\bar{x} \in J^{-1}(\mu)$, such that $\pi_\mu(\bar{x}) = \hat{\bar{x}}$, and a tangent vector $\bar{Z}(\bar{x}) \in T_{\bar{x}}J^{-1}(\mu)$ such that

$$T_{\bar{\hat{x}}}\psi^{-1} \cdot \bar{\hat{Y}}(\bar{\hat{x}}) = T_{\bar{x}}\pi_\mu \cdot \bar{Z}(\bar{x}). \tag{54}$$

Then (53) becomes

$$\begin{aligned} \langle(\psi^*)^{-1}\hat{\bar{\alpha}}, \bar{\hat{Y}}\rangle(\bar{\hat{x}}) &= \langle\hat{\bar{\alpha}}(\hat{\bar{x}}), T_{\bar{x}}\pi_\mu \cdot \bar{Z}(\bar{x})\rangle \\ &= \langle\pi_\mu^*\hat{\bar{\alpha}}(\bar{x}), \bar{Z}(\bar{x})\rangle \\ &= \langle\bar{\alpha}(\bar{x}), \bar{Z}(\bar{x})\rangle \\ &= \langle\iota_1^*\alpha(\bar{x}), \bar{Z}(\bar{x})\rangle \\ &= \langle\alpha(x), T_{\bar{x}}\iota_1 \cdot \bar{Z}(\bar{x})\rangle, \end{aligned} \tag{55}$$

where $x = \iota_1(\bar{x})$.
Because ψ is a diffeomorphism $T\psi^{-1} = (T\psi)^{-1}$ is invertible. Then (54) becomes

$$\bar{\hat{Y}}(\bar{\hat{x}}) = T_{\hat{\bar{x}}}\psi \cdot T_{\bar{x}}\pi_\mu \cdot \bar{Z}(\bar{x}).$$

This implies

$$\begin{aligned} \hat{Y}(\iota_2(\bar{\hat{x}})) &= T_{\bar{\hat{x}}}\iota_2 \cdot \bar{\hat{Y}}(\bar{\hat{x}}) \\ &= T_{\bar{\hat{x}}}\iota_2 \cdot T_{\hat{\bar{x}}}\psi \cdot T_{\bar{x}}\pi_\mu \cdot \bar{Z}(\bar{x}) \\ &= T_x\pi \cdot T_{\bar{x}}\iota_1 \cdot \bar{Z}(\bar{x}), \end{aligned} \tag{56}$$

where we used the commutativity of diagram (49), $\iota_2 \circ \psi \circ \pi_\mu = \pi \circ \iota_1$, which implies $T\iota_2 \circ T\psi \circ T\pi_\mu = T\pi \circ T\iota_1$. Because also

$$\hat{Y}(\iota_2(\bar{\hat{x}})) = T_x\pi \cdot Y(x) \tag{57}$$

(note that again by commutativity $\iota_2(\bar{\hat{x}}) = \pi(x)$), we get the following from (56,57)

$$T_x\pi \cdot T_{\bar{x}}\iota_1 \cdot \bar{Z}(\bar{x}) = T_x\pi \cdot Y(x),$$

which implies that

$$T_{\bar{x}}\iota_1 \cdot \bar{Z}(\bar{x}) = Y(x) + Y_0(x), \tag{58}$$

where $Y_0(x) \in \ker T_x\pi$. Plugging (58) into (55) gives

$$\langle(\psi^*)^{-1}\hat{\bar{\alpha}}, \bar{\hat{Y}}\rangle(\bar{\hat{x}}) = \langle\alpha(x), Y(x) + Y_0(x)\rangle. \tag{59}$$

However, $\alpha(x)$ maps $\ker T_x\pi$ to zero. Indeed,

$$\ker T\pi = \mathrm{span}_{C^\infty(\mathcal{X})}\,\{X_{P_j}\},$$

i.e. the distribution spanned by the symmetry vector fields, and

$$\begin{aligned} \langle\alpha(x), \sum f_j(x)X_{P_j}(x)\rangle &= -\langle\sum f_j(x)dP_j(x), X(x)\rangle \\ &= -\sum f_j(x)\langle dP_j(x), X(x)\rangle \\ &= 0, \end{aligned} \tag{60}$$

where we used that $(X, \alpha) \in D$ and $(X_{P_j}, dP_j) \in D$ so

$$\langle \alpha, X_{P_j} \rangle + \langle dP_j, X \rangle = 0,$$

$(D = D^{\perp})$, and $\bar{X} \sim_{\iota_1} X$ which gives that

$$\langle dP_j(x), X(x) \rangle = 0$$

(because $\bar{X}$ is tangent at $J^{-1}(\mu)$, i.e. the common level set of $P_j, j = 1, \ldots, r$). From (60) it follows that

$$\langle \alpha(x), Y_0(x) \rangle = 0,$$

so (59) becomes

$$\langle (\psi^*)^{-1} \hat{\bar{\alpha}}, \bar{\hat{Y}} \rangle (\bar{\hat{x}}) = \langle \alpha(x), Y(x) \rangle. \tag{61}$$

Now we will work out the second term of (52), which is a bit easier.

$$\begin{aligned}
\langle \bar{\hat{\beta}}, \psi_* \hat{\bar{X}} \rangle (\bar{\hat{x}}) &= \langle \iota_2^* \hat{\beta}, \psi_* \hat{\bar{X}} \rangle (\bar{\hat{x}}) \\
&= \langle \hat{\beta}(\iota_2(\bar{\hat{x}})), T_{\bar{\hat{x}}} \iota_2 \cdot \psi_* \hat{\bar{X}}(\bar{\hat{x}}) \rangle \\
&= \langle \hat{\beta}(\iota_2(\bar{\hat{x}})), T_{\bar{\hat{x}}} \iota_2 \cdot T_{\hat{\bar{x}}} \psi \cdot \hat{\bar{X}}(\hat{\bar{x}}) \rangle \\
&= \langle \hat{\beta}(\iota_2(\bar{\hat{x}})), T_{\bar{\hat{x}}} \iota_2 \cdot T_{\hat{\bar{x}}} \psi \cdot T_{\bar{x}} \pi_\mu \cdot \bar{X}(\bar{x}) \rangle \\
&\quad \text{and now using commutativity gives} \\
&= \langle \hat{\beta}(\iota_2(\bar{\hat{x}})), T_x \pi \cdot T_{\bar{x}} \iota_1 \cdot \bar{X}(\bar{x}) \rangle \\
&= \langle \pi^* \hat{\beta}(x), T_{\bar{x}} \iota_1 \cdot \bar{X}(\bar{x}) \rangle \\
&= \langle \beta(x), X(x) \rangle.
\end{aligned} \tag{62}$$

Using (61,62) our original equation (52) becomes

$$\langle (\psi^*)^{-1} \hat{\bar{\alpha}}, \bar{\hat{Y}} \rangle (\bar{\hat{x}}) + \langle \bar{\hat{\beta}}, \psi_* \hat{\bar{X}} \rangle (\bar{\hat{x}}) = \langle \alpha(x), Y(x) \rangle + \langle \beta(x), X(x) \rangle = 0, \tag{63}$$

because $(X, \alpha), (Y, \beta) \in D$ which by $D = D^{\perp}$ implies that

$$\langle \alpha, Y \rangle + \langle \beta, X \rangle = 0.$$

Note that $(\bar{\hat{Y}}, \bar{\hat{\beta}}) \in \bar{\hat{D}}$ and $\bar{\hat{x}} \in \hat{J}^{-1}(\hat{\mu})$ where arbitrarily chosen, so (63) proves that $(\psi_* \hat{\bar{X}}, (\psi^*)^{-1} \hat{\bar{\alpha}}) \in \bar{\hat{D}}^{\perp} = \bar{\hat{D}}$. This ends the proof. □

Theorem 28. *Consider the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$. Suppose the system has r independent first integrals $P_1, \ldots, P_r$, satisfying (43), and corresponding independent vector fields $X_{P_1}, \ldots, X_{P_r}$, satisfying (44), which generate a strong symmetry Lie group G of $(\mathcal{X}, D, H)$. Assume that $D(\bar{x}) \cap E_s(\bar{x})$, $\bar{x} \in P^{-1}(\mu)$, is constant dimensional on $P^{-1}(\mu)$, and that $V + \mathsf{G}_0$ and $D \cap E_q$ are constant dimensional on $\mathcal{X}$. Then, using the two reduction procedures described above, the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ reduces to implicit generalized Hamiltonian systems on the manifolds $P^{-1}(\mu)$, $P^{-1}(\mu)/G_\mu$, $\mathcal{X}/G$ and $\hat{P}^{-1}(\hat{\mu})$ in diagram (49). The solutions of $(P^{-1}(\mu)/G_\mu, \hat{\bar{D}}, \hat{\bar{H}})$, respectively $(\hat{P}^{-1}(\hat{\mu}), \bar{\hat{D}}, \bar{\hat{H}})$, are just the projections under π_μ, respectively π, of certain solutions of $(\mathcal{X}, D, H)$ (i.e., the solutions along projectable vector fields, see proposition 23). Furthermore, the reduced Hamiltonians satisfy $\hat{\bar{H}} = \bar{\hat{H}} \circ \psi$ and the solutions of the two systems are diffeomorphic, that is, $\hat{\bar{x}}(t)$ is a solution of $(P^{-1}(\mu)/G_\mu, \hat{\bar{D}}, \hat{\bar{H}})$ if and only if $\psi(\hat{\bar{x}}(t))$ is a solution of $(\hat{P}^{-1}(\hat{\mu}), \bar{\hat{D}}, \bar{\hat{H}})$.*

Proof. Because $D(\bar{x}) \cap E_s(\bar{x}),\ \bar{x} \in P^{-1}(\mu)$, is constant dimensional on $P^{-1}(\mu)$, the system $(\mathcal{X}, D, H)$ can be reduced to the implicit generalized Hamiltonian system $(P^{-1}(\mu), \bar{D}, \bar{H})$, using proposition 19. Let $\bar{V}_\mu$ denote the distribution on $P^{-1}(\mu)$ spanned by the infinitesimal generators of G_μ. Let $\bar{\mathsf{G}}_0$ be the distribution as defined in section 2 corresponding to the generalized Dirac structure $\bar{D}$. Finally, let $\bar{E}_q$ be the bundle as defined in (28) corresponding to $P^{-1}(\mu)$. We show that constant dimensionality of $V + \mathsf{G}_0$ and $D \cap E_q$ on $\mathcal{X}$ implies constant dimensionality of $\bar{V}_\mu + \bar{\mathsf{G}}_0$ and $\bar{D} \cap \bar{E}_q$ on $P^{-1}(\mu)$.
First note that $\bar{V}_\mu \subset \bar{\mathsf{G}}_0$, because take arbitrary $\bar{X} \in \bar{V}_\mu$, then $\bar{X} \sim_{\iota_1} X = \sum_i h_i X_{P_i},\ h_i \in C^\infty(\mathcal{X})$, for some $X \in V$ (because G_μ is the Lie subgroup of symmetries of G that leave the level set $P^{-1}(\mu)$ invariant, i.e. that are tangent to this level set). Because $(X, \sum_i h_i dP_i) \in D$ this implies that

$$(\bar{X}, \iota_1^* \sum_i h_i dP_i) = (\bar{X}, \sum_i (h_i \circ \iota_1)\ d(P_i \circ \iota_1)) = (\bar{X}, \sum_i (h_i \circ \iota_1) \cdot 0) = (\bar{X}, 0) \in \bar{D},$$

and so $\bar{X} \in \bar{\mathsf{G}}_0$. Furthermore, by definition of $\bar{D}$, $\bar{\mathsf{G}}_0$ consists of all $\bar{X} \in TP^{-1}(\mu)$ such that $\bar{X} \sim_{\iota_1} X \in \mathsf{G}_1$ with $(X, \alpha) \in D$ such that $\iota_1^* \alpha = 0$. This means $\bar{X} \sim_{\iota_1} X \in \mathsf{G}_0$ or $\bar{X} \sim_{\iota_1} X \in V$ (note that if $X \in \mathsf{G}_0$ then $\langle dP_i, X \rangle = 0,\ i = 1, \ldots, r$, so X is tangent to the level set $P^{-1}(\mu)$). Concluding we get that

$$\bar{V}_\mu + \bar{\mathsf{G}}_0 = \bar{\mathsf{G}}_0 = \bar{V}_\mu + \mathsf{G}_0|_{P^{-1}(\mu)},$$

where $\mathsf{G}_0|_{P^{-1}(\mu)}$ denotes the set of all vector fields in G_0 restricted to $P^{-1}(\mu)$. Now, since $V + \mathsf{G}_0$ is constant dimensional on $\mathcal{X}$, it follows that $\bar{V}_\mu + \mathsf{G}_0|_{P^{-1}(\mu)}$ is constant dimensional on $P^{-1}(\mu)$ (since the only elements in $V + \mathsf{G}_0$ that do not lie in $TP^{-1}(\mu)$ are the elements of the $(r-m)$-dimensional distribution $\bar{V}_\mu^\circ \in T\mathcal{X}$, where $V(\bar{x}) = \bar{V}_\mu(\bar{x}) \oplus \bar{V}_\mu^\circ(\bar{x}),\ \forall \bar{x} \in P^{-1}(\mu)$, and $m = \dim G_\mu$). Thus, $\bar{\mathsf{G}}_0 = \bar{V}_\mu + \bar{\mathsf{G}}_0$ is constant dimensional on $P^{-1}(\mu)$. Since $\bar{\mathsf{G}}_0$ and $\bar{V}_\mu + \bar{\mathsf{G}}_0$ are constant dimensional it follows that also $\operatorname{ann}(\bar{V}_\mu) \cap \bar{\mathsf{P}}_1$ is constant dimensional on $P^{-1}(\mu)$, where $\bar{\mathsf{P}}_1$ is the co-distribution corresponding to $\bar{D}$ as defined in section 2. From $\bar{\mathsf{G}}_0$ and $\operatorname{ann}(\bar{V}_\mu) \cap \bar{\mathsf{P}}_1$ constant dimensional it immediately follows that also $\bar{D} \cap \bar{E}_q$ is constant dimensional on $P^{-1}(\mu)$. So the assumptions of proposition 23 are satisfied and we can reduce the system $(P^{-1}(\mu), \bar{D}, \bar{H})$ further to the implicit generalized Hamiltonian system $(P^{-1}(\mu)/G_\mu, \hat{\bar{D}}, \hat{\bar{H}})$. This proves the first part of the theorem.

For the second part, lemma 26 states that there exists a diffeomorphism ψ which makes the diagram (49) commuting, that is $\pi \circ \iota_1 = \iota_2 \circ \psi \circ \pi_\mu$. Take arbitrary $\hat{\bar{x}} \in P^{-1}(\mu)/G_\mu$ and let $\bar{x} \in P^{-1}(\mu)$ be such that $\pi_\mu(\bar{x}) = \hat{\bar{x}}$, then

$$\begin{aligned}
\hat{\bar{H}}(\hat{\bar{x}}) &= \bar{H}(\bar{x}) = \iota_1^* H(\bar{x}) = \iota_1^*(\pi^* \hat{H})(\bar{x}) = \hat{H} \circ \pi \circ \iota_1(\bar{x}) \\
&= \hat{H} \circ \iota_2 \circ \psi \circ \pi_\mu(\bar{x}) = \hat{H} \circ \iota_2 \circ \psi(\hat{\bar{x}}) = \iota_2^* \hat{H} \circ \psi(\hat{\bar{x}}) \\
&= \bar{\hat{H}} \circ \psi(\hat{\bar{x}}),
\end{aligned}$$

proving that $\hat{\bar{H}} = \bar{\hat{H}} \circ \psi$.
For proving that the solutions of the two reduced systems are diffeomorphic we use that ψ is a Dirac isomorphism and the fact that $\hat{\bar{H}} = \psi^* \bar{\hat{H}}$. First notice that ψ is a Dirac isomorphism implies that ψ is pointwise an isomorphism between the two linear spaces $\hat{\bar{D}}(\hat{\bar{x}})$ and $\bar{\hat{D}}(\psi(\hat{\bar{x}}))$. Now, let $\hat{\bar{x}}(t)$ be a solution of $(P^{-1}(\mu)/G_\mu, \hat{\bar{D}}, \hat{\bar{H}})$, i.e.

$$(X_{\hat{\bar{H}}}, d\hat{\bar{H}})(\hat{\bar{x}}(t)) \in \hat{\bar{D}}(\hat{\bar{x}}(t)), \quad \text{for all } t \in I,$$

where $X_{\hat{\bar{H}}}(\hat{\bar{x}}(t)) = \dot{\hat{\bar{x}}}(t)$ and I is the interval of existence of $\hat{\bar{x}}(t)$. Because ψ is pointwise an isomorphism it follows that

$$(X_{\bar{\hat{H}}}, d\bar{\hat{H}})(\psi(\hat{\bar{x}}(t))) \in \bar{\hat{D}}(\psi(\hat{\bar{x}}(t))), \quad \text{for all } t \in I,$$

where we defined

$$X_{\bar{\hat{H}}}(\psi(\hat{\bar{x}}(t))) = T_{\hat{\bar{x}}(t)}\psi \cdot X_{\hat{\bar{H}}}(\hat{\bar{x}}(t)). \tag{64}$$

Because of (64) it follows that $\frac{d}{dt}\psi(\hat{\bar{x}}(t)) = X_{\bar{\hat{H}}}(\psi(\hat{\bar{x}}(t)))$, which implies that $\psi(\hat{\bar{x}}(t))$ is a solution of $(\hat{P}^{-1}(\hat{\mu}), \bar{\hat{D}}, \bar{\hat{H}})$. The converse statement is proven in the same way. □

Example 8. Consider the Dirac structure given in example 1 (with D closed), and the Hamiltonian system $(\mathcal{X}, D, H)$ corresponding to a function $H \in C^\infty(\mathcal{X})$. Assuming the conditions in theorem 28 are satisfied, the system reduces to Hamiltonian systems on $P^{-1}(\mu)/G_\mu$ and $\hat{P}^{-1}(\hat{\mu})$. The corresponding Dirac structure $\hat{\bar{D}}$, respectively $\bar{\hat{D}}$, is again a symplectic structure on $P^{-1}(\mu)/G_\mu$, respectively $\hat{P}^{-1}(\hat{\mu})$.

Proof. The Dirac structure D is given by

$$D = \{(X, \alpha) \in T\mathcal{X} \oplus T^*\mathcal{X} \mid \alpha = \omega(X, \cdot)\}.$$

Now, the induced Dirac structure $\bar{D}$ on $P^{-1}(\mu)$ is defined as (where we use some shorthand notation)

$$\begin{aligned} \bar{D} &= \{(\bar{X}, \bar{\alpha}) \mid \bar{X} \sim_{\iota_1} X,\ \bar{\alpha} = \iota_1^*\alpha \text{ for some } (X, \alpha) \in D\} \\ &= \{(\bar{X}, \bar{\alpha}) \mid \bar{X} \sim_{\iota_1} X,\ \bar{\alpha} = \iota_1^*\alpha \text{ for some } X, \alpha \text{ s.t. } \alpha = \omega(X, \cdot)\}. \end{aligned} \tag{65}$$

Now, $\alpha = \omega(X, \cdot)$ implies

$$\bar{\alpha} = \iota_1^*\alpha = \iota_1^*(\omega(X, \cdot)) = (\iota_1^*\omega)(\bar{X}, \cdot),$$

where in the last step we used that $\bar{X} \sim_{\iota_1} X$. So (65) becomes

$$\bar{D} = \{(\bar{X}, \bar{\alpha}) \mid \bar{\alpha} = (\iota_1^*\omega)(\bar{X}, \cdot)\}.$$

Notice that this indeed defines a *presymplectic* structure on $P^{-1}(\mu)$ (because $\iota_1^*\omega$ has a nontrivial kernel given by the distribution spanned by G_μ, [1]). Reduction by using the residual symmetry group G_μ gives the Dirac structure

$$\begin{aligned} \hat{\bar{D}} &= \{(\hat{\bar{X}}, \hat{\bar{\alpha}}) \mid \exists \bar{X} \text{ s.t. } \bar{X} \sim_{\pi_\mu} \hat{\bar{X}} \text{ and } (\bar{X}, \pi_\mu^*\hat{\bar{\alpha}}) \in \bar{D}\} \\ &= \{(\hat{\bar{X}}, \hat{\bar{\alpha}}) \mid \exists \bar{X} \text{ s.t. } \bar{X} \sim_{\pi_\mu} \hat{\bar{X}} \text{ and } \pi_\mu^*\hat{\bar{\alpha}} = (\iota_1^*\omega)(\bar{X}, \cdot)\}. \end{aligned} \tag{66}$$

Now, the fact that G_μ spans the kernel of $\iota_1^*\omega$ implies that there exists a 2-form ω_μ on $P^{-1}(\mu)/G_\mu$ such that $\iota_1^*\omega = \pi_\mu^*\omega_\mu$, [1]. Then (66) becomes

$$\begin{aligned} \hat{\bar{D}} &= \{(\hat{\bar{X}}, \hat{\bar{\alpha}}) \mid \exists \bar{X} \text{ s.t. } \bar{X} \sim_{\pi_\mu} \hat{\bar{X}} \text{ and } \pi_\mu^*\hat{\bar{\alpha}} = (\pi_\mu^*\omega_\mu)(\bar{X}, \cdot)\} \\ &= \{(\hat{\bar{X}}, \hat{\bar{\alpha}}) \mid \pi_\mu^*\hat{\bar{\alpha}} = \pi_\mu^*(\omega_\mu(\hat{\bar{X}}, \cdot))\} \\ &= \{(\hat{\bar{X}}, \hat{\bar{\alpha}}) \mid \hat{\bar{\alpha}} = \omega_\mu(\hat{\bar{X}}, \cdot)\}, \end{aligned}$$

where in the second step we used that $\bar{X} \sim_{\pi_\mu} \hat{\bar{X}}$ and in the third step the fact that π_μ^* is injective. Furthermore, the fact that ω_μ is nondegenerate follows from ω being nondegenerate, [1]. Therefore $\hat{\bar{D}}$ defines a symplectic structure on $P^{-1}(\mu)/G_\mu$. Because $\hat{\bar{D}}$ and $\bar{\hat{D}}$ are isomorphic, $\bar{\hat{D}}$ also defines a symplectic structure on $\hat{P}^{-1}(\hat{\mu})$. □

This example shows that theorem 28 is a generalization of the classical (symplectic) reduction theorems described in [1, 7].

Example 9. Consider the Dirac structure given in example 2 (with D closed), and the Hamiltonian system $(\mathcal{X}, D, H)$ corresponding to a function $H \in C^\infty(\mathcal{X})$. Assuming the conditions in theorem 28 are satisfied, the system reduces to Hamiltonian systems on $P^{-1}(\mu)/G_\mu$ and $\hat{P}^{-1}(\hat{\mu})$. The corresponding Dirac structure $\hat{\bar{D}}$, respectively $\bar{\hat{D}}$, is again a Poisson structure on $P^{-1}(\mu)/G_\mu$, respectively $\hat{P}^{-1}(\hat{\mu})$.

Proof. The Dirac structure D is given by

$$D = \{(X, \alpha) \in T\mathcal{X} \oplus T^*\mathcal{X} \mid X = J(\alpha, \cdot)\}.$$

J is the structure matrix of the corresponding Poisson bracket $\{\cdot,\cdot\} = \{\cdot,\cdot\}_D$. The reduced Dirac structure $\hat{D}$ on $\mathcal{X}/G$ is defined as (where we use some shorthand notation)

$$\begin{aligned} \hat{D} &= \{(\hat{X}, \hat{\alpha}) \mid \exists X \text{ s.t. } X \sim_\pi \hat{X} \text{ and } (X, \pi^*\hat{\alpha}) \in D\} \\ &= \{(\hat{X}, \hat{\alpha}) \mid \exists X \text{ s.t. } X \sim_\pi \hat{X} \text{ and } X = J(\pi^*\hat{\alpha}, \cdot)\}. \end{aligned} \tag{67}$$

By (36) the bracket on $\mathcal{X}/G$ is defined by

$$\{\hat{F}_1, \hat{F}_2\}_{\hat{D}} \circ \pi = \{\hat{F}_1 \circ \pi, \hat{F}_2 \circ \pi\}_D,$$

for all $\hat{F}_1, \hat{F}_2 \in \mathcal{A}_{\hat{D}_2}$. Now, take arbitrary $\hat{F} \in C^\infty(\mathcal{X}/G)$, then $(X, \pi^* d\hat{F}) = (J(\pi^* d\hat{F}, \cdot), \pi^* d\hat{F}) \in D$. Let $(\xi)_\mathcal{X}$, $\xi \in \mathcal{G}$, be any symmetry vector field generated by G, then

$$(L_{(\xi)_\mathcal{X}}(J(\pi^* d\hat{F}, \cdot)), L_{(\xi)_\mathcal{X}}(\pi^* d\hat{F})) = ([J(\pi^* d\hat{F}, \cdot), (\xi)_\mathcal{X}], 0) \in D,$$

so $[J(\pi^* d\hat{F}, \cdot), (\xi)_\mathcal{X}] \in \mathsf{G}_0 = 0$ which implies that $J(\pi^* d\hat{F}, \cdot) \sim_\pi \hat{X}$ for some $\hat{X} \in T(\mathcal{X}/G)$. From this we can conclude that $\mathcal{A}_{\hat{D}} = C^\infty(\mathcal{X}/G)$.
Then $\{\cdot,\cdot\}_{\hat{D}}$ defines a skew-symmetric matrix $\hat{J}(\hat{x}) : T^*_{\hat{x}}(\mathcal{X}/G) \to T_{\hat{x}}(\mathcal{X}/G)$, $\hat{x} \in \mathcal{X}/G$, by taking as the (i,j)-th element $\hat{J}_{ij} = \{\hat{x}_i, \hat{x}_j\}_{\hat{D}}$, with $\hat{x}_i, \hat{x}_j$ the coordinate functions on $\mathcal{X}/G$. Denote the corresponding (2,0)-tensor by $\hat{J} : T^*(\mathcal{X}/G) \times T^*(\mathcal{X}/G) \to C^\infty(\mathcal{X}/G)$.

Now, let X be such that $X = J(\pi^*\hat{\alpha}, \cdot)$, then $X \sim_\pi \hat{X}$ for some $\hat{X} \in T(\mathcal{X}/G)$ (see the argument above, $L_{(\xi)_\mathcal{X}}(\pi^*\hat{\alpha}) = 0$ for an arbitrary symmetry $(\xi)_\mathcal{X}$, $\xi \in \mathcal{G}$, generated by G). We prove that $\hat{X} = \hat{J}(\hat{\alpha}, \cdot)$. Indeed, take an arbitrary G-invariant function $F \in C^\infty(\mathcal{X})$, then $F = \hat{F} \circ \pi$ for some $\hat{F} \in C^\infty(\mathcal{X}/G)$ and

$$\begin{aligned} L_X F = X[F] &= J(\pi^*\hat{\alpha}, dF) = J(\pi^* \sum_i \hat{h}_i d\hat{x}_i, dF) \\ &= \sum_i (\hat{h}_i \circ \pi)\; J(d(\hat{x}_i \circ \pi), d(\hat{F} \circ \pi)) \\ &= \sum_i (\hat{h}_i \circ \pi)\; \{\hat{x}_i \circ \pi, \hat{F} \circ \pi\}_D \\ &= \sum_i (\hat{h}_i \circ \pi)\; \{\hat{x}_i, d\hat{F}\}_{\hat{D}} \circ \pi \\ &= \sum_i (\hat{h}_i \circ \pi)\; \hat{J}(d\hat{x}_i, d\hat{F}) \circ \pi \\ &= \hat{J}(\sum_i \hat{h}_i d\hat{x}_i, d\hat{F}) \circ \pi \\ &= \hat{J}(\hat{\alpha}, d\hat{F}) \circ \pi \\ &= \hat{X}'[\hat{F}] \circ \pi = L_{\hat{X}'} \hat{F} \circ \pi, \end{aligned}$$

where we defined $\hat{X}' = \hat{J}(\hat{\alpha}, \cdot)$. So

$$L_X(\hat{F} \circ \pi) = (L_{\hat{X}'}\hat{F}) \circ \pi \tag{68}$$

for arbitrary $\hat{F} \in C^\infty(\mathcal{X}/G)$. This implies that $X \sim_\pi \hat{X}'$, see for instance ([2],proof of theorem 4.2.8). Because also $X \sim_\pi \hat{X}$ it follows that $\hat{X} = \hat{X}'$, and so $\hat{X} = \hat{J}(\hat{\alpha}, \cdot)$. Now (67) becomes

$$\hat{D} = \{(\hat{X}, \hat{\alpha}) \mid \hat{X} = \hat{J}(\hat{\alpha}, \cdot)\}.$$

This defines a Poisson structure on $\mathcal{X}/G$. Reducing the system to a level set $\hat{P}^{-1}(\hat{\mu})$ of the Casimir function $\hat{P}$ gives the Dirac structure

$$\begin{aligned}\bar{\hat{D}} &= \{(\bar{\hat{X}}, \bar{\hat{\alpha}}) \mid \bar{\hat{X}} \sim_{\iota_2} \hat{X}, \bar{\hat{\alpha}} = \iota_2^*\hat{\alpha} \text{ for some } (\hat{X}, \hat{\alpha}) \in \hat{D}\} \\ &= \{(\bar{\hat{X}}, \bar{\hat{\alpha}}) \mid \bar{\hat{X}} \sim_{\iota_2} \hat{X}, \bar{\hat{\alpha}} = \iota_2^*\hat{\alpha} \text{ for some } \hat{X}, \hat{\alpha} \text{ s.t. } \hat{X} = \hat{J}(\hat{\alpha}, \cdot)\}.\end{aligned} \tag{69}$$

Define a Poisson bracket $\widetilde{\{\cdot,\cdot\}}$ on $\hat{P}^{-1}(\hat{\mu})$ as follows. Take arbitrary $\bar{\hat{F}}_1, \bar{\hat{F}}_2 \in C^\infty(\hat{P}^{-1}(\hat{\mu}))$ and let $\hat{F}_1, \hat{F}_2 \in C^\infty(\mathcal{X}/G)$ be such that $\bar{\hat{F}}_j = \hat{F}_j \circ \iota_2,\ j = 1, 2$. Then define

$$\widetilde{\{\bar{\hat{F}}_1, \bar{\hat{F}}_2\}} = \{\hat{F}_1, \hat{F}_2\}_{\hat{D}} \circ \iota_2. \tag{70}$$

To see that $\widetilde{\{\cdot,\cdot\}}$ is well defined note that

$$\begin{aligned}\{\hat{F}_1, \hat{F}_2\}_{\hat{D}} \circ \iota_2 &= \hat{J}(d\hat{F}_1, d\hat{F}_2) \circ \iota_2 \\ &= \hat{X}_1[\hat{F}_2] \circ \iota_2 \\ &= (L_{\hat{X}_1}\hat{F}_2) \circ \iota_2 \\ &= \iota_2^*(L_{\hat{X}_1}\hat{F}_2) \\ &= L_{\bar{\hat{X}}_1}(\iota_2^*\hat{F}_2) \\ &= L_{\bar{\hat{X}}_1}\bar{\hat{F}}_2,\end{aligned} \tag{71}$$

where $\hat{X}_1 = \hat{J}(d\hat{F}_1, \cdot)$, and where we used that $\bar{\hat{X}}_1 \sim_{\iota_2} \hat{X}_1$ for some $\bar{\hat{X}}_1$, because every $\hat{X} \in \hat{\mathsf{G}}_1 \subset T(\mathcal{X}/G)$ is tangent to the level set $\hat{P}^{-1}(\hat{\mu})$ of the *Casimir* function $\hat{P}$. Now, let $\hat{F}_3 \in C^\infty(\mathcal{X}/G)$ be another function such that $\bar{\hat{F}}_1 = \hat{F}_3 \circ \iota_2$. Then $\iota_2^* d\hat{F}_1 = \iota_2^* d\hat{F}_3$, so $\hat{\gamma} := d\hat{F}_3 - d\hat{F}_1 \in \ker \iota_2^*$. Because

$$(\ker \iota_2^*)(\bar{\hat{x}}) = (\mathrm{span}_{C^\infty(\mathcal{X}/G)}\ \{d\hat{P}\})(\bar{\hat{x}}), \quad \forall\, \bar{\hat{x}} \in \hat{P}^{-1}(\hat{\mu}),$$

it follows that

$$\hat{X}_3(\bar{\hat{x}}) = \hat{J}(d\hat{F}_3, \cdot)(\bar{\hat{x}}) = \hat{J}(\bar{\hat{x}})(d\hat{F}_1(\bar{\hat{x}}) + \hat{\gamma}(\bar{\hat{x}}), \cdot) = \hat{J}(d\hat{F}_1, \cdot)(\bar{\hat{x}}) = \hat{X}_1(\bar{\hat{x}}), \tag{72}$$

for all $\bar{\hat{x}} \in \hat{P}^{-1}(\hat{\mu})$, where

$$\hat{J}(\bar{\hat{x}})(\hat{\gamma}(\bar{\hat{x}}), \cdot) = \hat{J}(\bar{\hat{x}})(h(\bar{\hat{x}})d\hat{P}(\bar{\hat{x}}), \cdot) = h(\bar{\hat{x}})\hat{X}_{\hat{P}}(\bar{\hat{x}}) = 0,$$

because $\hat{X}_{\hat{P}} \in \hat{\mathsf{G}}_0 = 0$. (72) implies that $\bar{\hat{X}}_1$ does not depend on the choice of extending $\bar{\hat{F}}_1$. Then (71) clearly shows that $\widetilde{\{\bar{\hat{F}}_1, \bar{\hat{F}}_2\}}$ does not depend on the choice of $\hat{F}_1$ and $\hat{F}_2$, so $\widetilde{\{\cdot,\cdot\}}$ is well defined.

$\widetilde{\{\cdot,\cdot\}}$ defines a skew-symmetric matrix $\bar{\hat{J}}(\bar{\hat{x}}) : T^*_{\bar{\hat{x}}}\hat{P}^{-1}(\hat{\mu}) \to T_{\bar{\hat{x}}}\hat{P}^{-1}(\hat{\mu})$, $\bar{\hat{x}} \in \hat{P}^{-1}(\hat{\mu})$, by taking as the (i,j)-th element $\bar{\hat{J}}_{ij} = \widetilde{\{\bar{\hat{x}}_i, \bar{\hat{x}}_j\}}$, with $\bar{\hat{x}}_i, \bar{\hat{x}}_j$ the coordinate functions on $\hat{P}^{-1}(\hat{\mu})$. Denote the corresponding (2,0)-tensor by $\bar{\hat{J}} : T^*\hat{P}^{-1}(\hat{\mu}) \times T^*\hat{P}^{-1}(\hat{\mu}) \to C^\infty(\hat{P}^{-1}(\hat{\mu}))$.

Now let $\hat{X} = \hat{J}(\hat{\alpha}, \cdot)$, then $\bar{\hat{X}} \sim_{\iota_2} \hat{X}$ for some $\bar{\hat{X}} \in T\hat{P}^{-1}(\hat{\mu})$. We prove that $\bar{\hat{X}} = \bar{\hat{J}}(\iota_2^*\hat{\alpha}, \cdot) = \bar{\hat{J}}(\bar{\hat{\alpha}}, \cdot)$. Using the same derivation that led to (68), we can prove that

$$L_{\bar{\hat{X}}'}(\hat{F} \circ \iota_2) = (L_{\hat{X}}\hat{F}) \circ \iota_2,$$

for all $\hat{F} \in C^\infty(\mathcal{X}/G)$, where we defined $\bar{\hat{X}}' = \bar{\hat{J}}(\bar{\hat{\alpha}}, \cdot)$. This implies that $\bar{\hat{X}}' \sim_{\iota_2} \hat{X}$. Because also $\bar{\hat{X}} \sim_{\iota_2} \hat{X}$, it follows that $\bar{\hat{X}} = \bar{\hat{X}}'$, and so $\bar{\hat{X}} = \bar{\hat{J}}(\bar{\hat{\alpha}}, \cdot)$. Then (69) becomes

$$\bar{\hat{D}} = \{(\bar{\hat{X}}, \bar{\hat{\alpha}}) \mid \bar{\hat{X}} = \bar{\hat{J}}(\bar{\hat{\alpha}}, \cdot)\}. \tag{73}$$

We see that $\bar{\hat{D}}$ defines a Poisson structure on $\hat{P}^{-1}(\hat{\mu})$. Because $\hat{\bar{D}}$ and $\bar{\hat{D}}$ are isomorphic, $\hat{\bar{D}}$ also defines a Poisson structure on $P^{-1}(\mu)/G_\mu$. □

Remark 10. From (73) it is immediately clear that $\mathcal{A}_{\bar{\hat{D}}} = C^\infty(\hat{P}^{-1}(\hat{\mu}))$. Then from (71) and (73) it follows that the bracket defined in (70) equals the bracket $\{\cdot,\cdot\}_{\bar{\hat{D}}}$ induced by $\bar{\hat{D}}$.

This example shows that theorem 28 is a generalization of the classical (Poisson) reduction theorems described in [9, 12].

Finally, note that the reduced system on $P^{-1}(\mu)$ does not represent a classical Poisson system, but it is described by an implicit generalized Hamiltonian system, with a Dirac structure as the underlying geometric structure. This was already noticed in [3].

6 The Casimir function $\hat{P}$

In this section we will take a closer look at the map $\hat{P}$ introduced in the second reduction procedure in the previous section. In particular we will show that "locally $d\hat{P} \in \mathsf{P}_0$", which allows us to moderate the proof of proposition 22 a little bit such that the result still holds in case $\bar{\mathcal{X}} = \hat{P}^{-1}(\hat{\mu})$ (as is the case in the reduction procedure in theorem 28).

Recall that the momentum map was defined as $P : \mathcal{X} \to \mathcal{G}^*$

$$P(x) = \sum_{i=1}^{r} P_i(x)\mu_i, \tag{74}$$

where $\{\mu_1, \ldots, \mu_r\}$ is a basis of $\mathcal{G}^*$, and $P_1, \ldots, P_r$ are the first integrals of the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$. Define the quotient manifold $\hat{\mathcal{G}}^* = \mathcal{G}^*/G$ of coadjoint orbits $\mathcal{O}_\mu$ in $\mathcal{G}^*$, and the corresponding projection map $\varpi : \mathcal{G}^* \to \hat{\mathcal{G}}^*$. Define the map $\hat{P} : \mathcal{X}/G \to \hat{\mathcal{G}}^*$ by

$$\hat{P} \circ \pi = \varpi \circ P, \tag{75}$$

where $\pi : \mathcal{X} \to \mathcal{X}/G$ is the projection map. It was shown in section 5 that $\hat{P}$ is well defined.

Example 10. Consider an abelian r-dimensional strong symmetry Lie group G of the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$, with corresponding first integrals $P_1, \ldots, P_r$ which are *in involution*, i.e.

$$\{P_i, P_j\}_D = 0, \; i, j = 1, \ldots, r. \tag{76}$$

Since G is *abelian*, i.e. $g_1 \cdot g_2 = g_2 \cdot g_1$ for all $g_1, g_2 \in G$ where $\cdot$ is the group multiplication in G, it follows that the map $a_g : G \to G, \; g \in G$, defined by

$$a_g = R_{g^{-1}} \circ L_g = L_g \circ R_{g^{-1}}, \; i.e. \; a_g(g_1) = g \cdot g_1 \cdot g^{-1} \; \forall g_1 \in G,$$

is the identity, i.e. $a_g(g_1) = g_1, \; \forall g_1 \in G$. Therefore also the adjoint map $Ad_g : \mathcal{G} = T_eG \to \mathcal{G} = T_eG$ defined by $Ad_g = T_e a_g$ (i.e. the tangent of the map a_g at the identity element $e \in G$) is the identity map, and consequently the coadjoint map (action) $Ad_g^* : \mathcal{G}^* \to \mathcal{G}^*$, which is the dual of the map Ad_g, also equals the identity map. So $Ad_g^* = I : \mathcal{G}^* \to \mathcal{G}^*, \; I(\mu) = \mu \; \forall \mu \in \mathcal{G}^*$, for every $g \in G$. This implies that the coadjoint orbits $\mathcal{O}_\mu$ in $\mathcal{G}^*$, defined by (46), are just the points in $\mathcal{G}^*$, i.e. $\mathcal{O}_\mu = \{\mu\}, \; \forall \mu \in \mathcal{G}^*$. Therefore in the case of an abelian symmetry Lie group G, $\hat{\mathcal{G}}^* = \mathcal{G}^*$ (after identification of the set $\{\mu\}$ with the point μ), so the projection ϖ is the identity map.
The momentum map $P : \mathcal{X} \to \mathcal{G}^*$ is defined as in (74). Then by (75), $\hat{P} : \mathcal{X}/G \to \mathcal{G}^*$ satisfies

$$\hat{P} \circ \pi = \sum_{i=1}^{r} P_i \, \mu_i,$$

which implies that

$$\hat{P} = \sum_{i=1}^{r} \hat{P}_i \, \mu_i,$$

with $\hat{P}_i \in C^\infty(\mathcal{X}/G)$ such that $\hat{P}_i \circ \pi = P_i, \; i = 1, \ldots, r$ (note that by (76) it follows that every P_i is invariant under the action of G). Since $(X_{P_i}, dP_i) \in D$ and $X_{P_i} \sim_\pi 0$ it follows that $(0, d\hat{P}_i) \in \hat{D}$, so $d\hat{P}_i \in \mathsf{P}_0, \; i = 1, \ldots, r$. This implies that $\hat{P}_1, \ldots, \hat{P}_r$ are Casimirs functions of $(\mathcal{X}/G, \hat{D}, \hat{H})$, and we can use proposition 22 to further reduce the system to an implicit generalized Hamiltonian system $(\hat{P}^{-1}(\hat{\mu}), \bar{\hat{D}}, \bar{\hat{H}})$ on the level set $\hat{P}^{-1}(\hat{\mu})$, where $\hat{\mu} = \varpi(\mu) = \mu$ (note that this is exactly a level set of $\hat{P}_1, \ldots, \hat{P}_r$).

Now we return to the general case. Let $\hat{P}$ be defined as in (75). Because $\mathcal{G}^*$ is the dual of the Lie algebra $\mathcal{G} = T_eG$, which is a vector space (over $\mathbb{R}$), also $\mathcal{G}^*$ is a vector space (over $\mathbb{R}$). Therefore $\mathcal{G}^*$ is globally isomorphic to $\mathbb{R}^r$ via some isomorphism $\varphi : \mathcal{G}^* \to \mathbb{R}^r$. Since $\hat{\mathcal{G}}^* = \mathcal{G}^*/G$ is a manifold (under the appropriate assumptions on G) it is locally diffeomorphic to $\mathbb{R}^m$, where m is the dimension of $\hat{\mathcal{G}}^*$, via some diffeomorphism $\hat{\varphi}_U : U \subset \hat{\mathcal{G}}^* \to \mathbb{R}^m$. Consider a local chart $(U, \hat{\varphi}_U)$ of $\hat{\mathcal{G}}^*$, then (75) implies

$$\hat{\varphi}_U \circ \hat{P} \circ \pi(x) = \hat{\varphi}_U \circ \varpi \circ \varphi^{-1} \circ \varphi \circ P(x), \quad \forall x \in W \subset \mathcal{X}, \tag{77}$$

where W is such that $\hat{P} \circ \pi(W) \subset U$. Now, since $\hat{\varphi}_U \circ \varpi \circ \varphi^{-1} : \mathbb{R}^r \to \mathbb{R}^m$ is a projection, it is a linear map and therefore it can be described by a matrix $[Proj] \in \mathbb{R}^{m \times r}$. Note that $\varphi \circ P$ is exactly the r-vector of first integrals, i.e.

$$\varphi \circ P(x) = [P_1(x), \ldots, P_r(x)]^T.$$

Then (77) becomes

$$\begin{aligned} \hat{\varphi}_U \circ \hat{P} \circ \pi(x) &= [Proj][P_1(x), \ldots, P_r(x)]^T \\ &= \begin{bmatrix} c_{11}P_1(x) + \cdots + c_{1r}P_r(x) \\ \vdots \\ c_{m1}P_1(x) + \cdots + c_{mr}P_r(x) \end{bmatrix}, \end{aligned} \tag{78}$$

for some constants $c_{ij} \in \mathbb{R},\ i = 1, \ldots, m,\ j = 1, \ldots, r$. Now, $\hat{\varphi}_U \circ \hat{P}$ defines the m-vector

$$\hat{\varphi}_U \circ \hat{P}(\hat{x}) = [\hat{P}_1(\hat{x}), \ldots, \hat{P}_m(\hat{x})]^T,$$

where $\hat{P}_i \in C^\infty(W/G),\ i = 1, \ldots, m$. By (78) it follows that

$$\pi^* d\hat{P}_i = c_{i1} dP_1 + \cdots + c_{ir} dP_r, \quad i = 1, \ldots, m.$$

Now, take an arbitrary pair $(\hat{Y}, \hat{\beta}) \in \hat{D}$. Then

$$\langle d\hat{P}_i, \hat{Y}\rangle(\hat{x}) = \langle \pi^* d\hat{P}_i, Y\rangle(x) = \langle \sum_{j=1}^{r} c_{ij} dP_j, Y\rangle(x) = -\langle \beta, \sum_{j=1}^{r} c_{ij} X_{P_j}\rangle(x) = -\langle \hat{\beta}, 0\rangle(\hat{x}) = 0, \tag{79}$$

$\forall \hat{x} \in W/G$, where $x \in W,\ \pi(x) = \hat{x},\ i = 1, \ldots, m$, since $Y \sim_\pi \hat{Y},\ \beta = \pi^*\hat{\beta}$, with $(Y, \beta) \in D$, and $\sum_j c_{ij} X_{P_j} \sim_\pi 0$. So locally $d\hat{P}_i \in \text{ann}\ \hat{\mathsf{G}}_1 = \hat{\mathsf{P}}_0,\ i = 1, \ldots, m$. (79) is what we meant saying that "locally $d\hat{P} \in \hat{\mathsf{P}}_0$".

Now consider the implicit generalized Hamiltonian system $(\mathcal{X}/G, \hat{D}, \hat{H})$ in the reduction procedure of theorem 28. The map $\hat{P}$ is a Casimir function by (48) (or more correctly, by (79)). As in proposition 22 we want to conclude that the solutions of $(\mathcal{X}/G, \hat{D}, \hat{H})$ lying in $\hat{P}^{-1}(\hat{\mu})$ are exactly the solutions of the reduced system $(\hat{P}^{-1}(\hat{\mu}), \bar{\hat{D}}, \bar{\hat{H}})$. Since it is in general not true that $d\hat{P} \in \hat{\mathsf{P}}_0$ we cannot use proposition 22 directly. However, since (79) holds, and since the level set $\hat{P}^{-1}(\hat{\mu})$ is locally given by the level set of $\hat{P}_1, \ldots, \hat{P}_m$, we can conclude that for every pair $(\hat{Y}, \hat{\beta}) \in \hat{D}$ it holds that $\hat{Y}$ is tangent to $\hat{P}^{-1}(\hat{\mu})$. Then we can copy the rest of the proof of proposition 22 to conclude that the solutions of $(\mathcal{X}/G, \hat{D}, \hat{H})$ lying in $\hat{P}^{-1}(\hat{\mu})$ are exactly the solutions of the reduced system $(\hat{P}^{-1}(\hat{\mu}), \bar{\hat{D}}, \bar{\hat{H}})$.

7 The explicit generalized Hamiltonian system

In this section we will take a closer look at the reduction procedure in theorem 28 in case the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ satisfies assumption 5. The motivation for this is as follows. Considering the reduction procedure in theorem 28 notice that we made some assumptions.

(i) To define the generalized Dirac structure $\bar{D}$ on the submanifold $P^{-1}(\mu)$, we needed the assumption that $D(\bar{x}) \cap E_s(\bar{x}),\ \bar{x} \in P^{-1}(\mu)$, is constant dimensional on $P^{-1}(\mu)$.

(ii) To define the generalized Dirac structure $\hat{D}$ on the quotient manifold $\mathcal{X}/G$, we needed the assumption that $V + \mathsf{G}_0$ and $D \cap E_q$ are constant dimensional on $\mathcal{X}$.

(iii) Finally, concerning proposition 23 about reduction of an implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ to an implicit generalized Hamiltonian $(\hat{\mathcal{X}}, \hat{D}, \hat{H})$ on a quotient manifold $\hat{X}$, we needed the assumption of projectability of a solution $x(t)$ to show that it reduces to a solution $\hat{x}(t)$ of $(\hat{\mathcal{X}}, \hat{D}, \hat{H})$.

These three assumptions are new with respect to the assumptions made in the classical reduction theorems of [1, 7, 9, 12]. Indeed, considering the reduction of classical explicit Hamiltonian systems like in examples 8 and 9, these three assumptions are void. For take an explicit Hamiltonian system defined with respect to a symplectic structure as in example 8. Because $\mathsf{G}_1 = T\mathcal{X}$ is constant dimensional, see example 5, $\mathsf{G}_1(\bar{x}) \cap T_{\bar{x}}\bar{\mathcal{X}} = T_{\bar{x}}\bar{\mathcal{X}}$, $\bar{x} \in \bar{\mathcal{X}}$, is constant dimensional on $\bar{\mathcal{X}}$, which implies that $D(\bar{x}) \cap E_s(\bar{x})$, $\bar{x} \in \bar{\mathcal{X}}$, is constant dimensional on $\bar{\mathcal{X}}$, see remark 4. Also, since $\mathsf{G}_0 = 0$, $V + \mathsf{G}_0 = V$ is constant dimensional (with $\dim V = r = \dim G$). Furthermore, since $\mathsf{P}_1 = T^*\mathcal{X}$, $\mathrm{ann}(V) \cap \mathsf{P}_1 = \mathrm{ann}(V)$ is constant dimensional, and together with G_0 constant dimensional this implies that $D \cap E_q$ is constant dimensional on $\mathcal{X}$. Finally, the vector field $X_{\bar{H}} \in TP^{-1}(\mu)$, corresponding to a solution $\bar{x}(t)$ of $(P^{-1}(\mu), \bar{D}, \bar{H})$ coming from a solution $x(t)$ of $(\mathcal{X}, D, H)$, is projectable to a vector field on $P^{-1}(\mu)/G_\mu$ [1, 7]. Note that the reduced Hamiltonian system $(\mathcal{X}/G, \hat{D}, \hat{H})$ on $\mathcal{X}/G$ is not a symplectic system anymore, so the reduction procedures in [1, 7] do not include the system $(\mathcal{X}/G, \hat{D}, \hat{H})$. However, $(\mathcal{X}/G, \hat{D}, \hat{H})$ is a Poisson system, and in [9, 12] it is proved that every solution of $(\mathcal{X}, D, H)$ projects to a solution of $(\mathcal{X}/G, \hat{D}, \hat{H})$.
With respect to the second classical example, consider an explicit Hamiltonian system defined on a Poisson structure as in example 9. Just as in the symplectic case $\mathsf{G}_0 = 0$ and $\mathsf{P}_1 = T^*\mathcal{X}$ (see example 6) imply that $V + \mathsf{G}_0$ and $D \cap E_q$ are constant dimensional on $\mathcal{X}$. Furthermore, in [9, 12] it is shown that every solution $x(t)$ of $(\mathcal{X}, D, H)$ projects to a solution $\hat{x}(t)$ of $(\mathcal{X}/G, \hat{D}, \hat{H})$. Again note that the reduced Hamiltonian system $(P^{-1}(\mu), \bar{D}, \bar{H})$ on $P^{-1}(\mu)$ is not a Poisson system anymore, and therefore is not included in the reduction procedures in [9, 12]. Under assumption (i), the reduced system on $P^{-1}(\mu)$ can be described as an implicit generalized Hamiltonian system on $P^{-1}(\mu)$. In [3] it is shown that assumption (i) is equivalent to the condition that every point $\bar{x} \in P^{-1}(\mu)$ lies on an orbit (of the group action on $\mathcal{X}$ corresponding to G) of *principal type.*

We saw in proposition 6 that, assuming the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ satisfies assumption 5, the system can be reduced to an explicit generalized Hamiltonian system $(\mathcal{X}_c, D_c, H_c)$ given by (8) (where the generalized Dirac structure D_c is defined by the structure matrix J_c). Then considering the examples above we would expect that the assumptions (ii) and (iii) are again automatically satisfied (because $(\mathcal{X}, D, H)$ is in essence the explicit system $(\mathcal{X}_c, D_c, H_c)$). Note that we already saw in the Poisson case that we cannot expect assumption (i) to be satisfied in general. Here we will investigate the contents of assumptions (ii) and (iii) if the system $(\mathcal{X}, D, H)$ satisfies assumption 5.

Consider the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ and assume that assumption 5 is satisfied. Assumption (ii) says that $V + \mathsf{G}_0$ and $D \cap E_q$ should be constant dimensional. Since V is constant dimensional and by assumption 5 also $\mathsf{G}_0 = \ker \mathsf{P}_1$ is constant dimensional $V + \mathsf{G}_0$ will be constant dimensional as well if and only if $V \cap \mathsf{G}_0$ is constant dimensional. Consider a strong symmetry X_{P_i} of $(\mathcal{X}, D, H)$, then by ([14], proposition 17) X_{P_i} will be tangent to $\mathcal{X}_c$, so $X_{P_i}(x_c) \in T_{x_c}\mathcal{X}_c$, $\forall x_c \in \mathcal{X}_c$. Furthermore, by assumption 5 it follows that $\mathsf{G}_0(x_c) \cap T_{x_c}\mathcal{X}_c = 0$, $\forall x_c \in \mathcal{X}_c$, see also ([14], proposition 17). Because V is the distribution spanned by the symmetries X_{P_i}, $i = 1, \ldots, r$, which generate the Lie group G, we have that

$$V(x_c) \cap \mathsf{G}_0(x_c) = 0, \quad \forall x_c \in \mathcal{X}_c,$$

which implies that $V + \mathsf{G}_0$ is constant dimensional *on* $\mathcal{X}_c$. Secondly, since P_1 is constant dimensional by assumption 5, $\mathrm{ann}(V + \mathsf{G}_0) = \mathrm{ann}(V) \cap \mathsf{P}_1$. Now, $V + \mathsf{G}_0$ constant dimensional on $\mathcal{X}_c$ implies $\mathrm{ann}(V) \cap \mathsf{P}_1$ constant dimensional on $\mathcal{X}_c$ and it follows that also $D \cap E_q$ is constant dimensional *on* $\mathcal{X}_c$.

Assumption (iii) says that a solution $x(t)$ of $(\mathcal{X}, D, H)$ should be projectable in order to reduce to a solution $\hat{x}(t)$ of $(\mathcal{X}/G, \hat{D}, \hat{H})$. Take an arbitrary solution $x(t)$ of $(\mathcal{X}, D, H)$, i.e. $\dot{x}(t) = X_H(x(t))$

where $X_H(x_c) \in T_{x_c}\mathcal{X}_c$, $\forall x_c \in \mathcal{X}_c$, is the *unique* vector field on $\mathcal{X}_c$ (by assumption 5, i.e. the vector field corresponding to the explicit system (8)) corresponding to H. By ([14],proposition 17)

$$[(\xi)_{\mathcal{X}}, X_H](x_c) = 0, \ \forall x_c \in \mathcal{X}_c, \tag{80}$$

for all symmetries $(\xi)_{\mathcal{X}}$, where $\xi \in \mathcal{G}$. This implies that

$$[V, X_H](x_c) \in V(x_c), \ \forall x_c \in \mathcal{X}_c,$$

which implies that X_H is projectable on $\mathcal{X}_c$ to a vector field $\hat{X}$ on $\mathcal{X}_c/G$. Using the Smooth Tietze Extension Theorem we can extend X_H to a vector field $X \in T\mathcal{X}$ which is projectable to a vector field on $\mathcal{X}/G$.
Furthermore, a solution $\bar{x}(t)$ of $(P^{-1}(\mu), \bar{D}, \bar{H})$, coming from a solution $x(t)$ of $(\mathcal{X}, D, H)$, should be projectable in order to reduce to a solution $\hat{\bar{x}}(t)$ of $(P^{-1}(\mu)/G_\mu, \hat{\bar{D}}, \hat{\bar{H}})$. Consider an arbitrary solution $x(t)$ of $(\mathcal{X}, D, H)$ in $P^{-1}(\mu)$, i.e. $\dot{x}(t) = X_H(x(t))$. By proposition 21, $X_{\bar{H}} \sim_{\iota_1} X_H$. Consider an arbitrary symmetry $(\xi)_{\bar{\mathcal{X}}} \in \bar{V}_\mu$, where $\xi \in \mathcal{G}_\mu$ (note that $\mathcal{G}_\mu$ is a Lie subalgebra of $\mathcal{G}$), then $(\xi)_{\bar{\mathcal{X}}} \sim_{\iota_1} (\xi)_{\mathcal{X}}$. Then by proposition 7 and (80) it follows that

$$[(\xi)_{\bar{\mathcal{X}}}, X_{\bar{H}}](\bar{x}_c) = 0, \ \forall \bar{x}_c \in \mathcal{X}_c \cap P^{-1}(\mu).$$

This implies that

$$[\bar{V}_\mu, X_{\bar{H}}](\bar{x}_c) \in \bar{V}_\mu(\bar{x}_c), \ \forall \bar{x}_c \in \mathcal{X}_c \cap P^{-1}(\mu),$$

which implies that $X_{\bar{H}}$ is projectable on $\mathcal{X}_c \cap P^{-1}(\mu)$ to a vector field $\hat{\bar{X}}$ on $(\mathcal{X}_c \cap P^{-1}(\mu))/G_\mu$. Using the Smooth Tietze Extension Theorem we can extend $X_{\bar{H}}$ to a vector field on $P^{-1}(\mu)$ which is projectable to a vector field on $P^{-1}(\mu)/G_\mu$. We conclude that the solutions of $(\mathcal{X}, D, H)$ and $(P^{-1}(\mu), \bar{D}, \bar{H})$ all satisfy the projectability assumption.

What we concluded above is quite interesting and also quite understandable. Consider the implicit generalized Hamiltonian system $(\mathcal{X}, D, H)$ and assume that assumption 5 is satisfied. Then the system reduces to the *explicit* generalized Hamiltonian system $(\mathcal{X}_c, D_c, H_c)$ given by (8). The solutions of the implicit system $(\mathcal{X}, D, H)$ are *exactly* the solutions of the explicit system $(\mathcal{X}_c, D_c, H_c)$, so, like in the classical cases in examples 8 and 9, they should *always* be projectable to solutions on the reduced systems. As we showed above, this is indeed the case (assumption (iii) is always satisfied). On the other hand however, we could *not* show that assumption (i) and (ii) are always satisfied. Indeed, even in the classical case of a Poisson structure on $\mathcal{X}$, we need assumption (i) to describe the reduced system on $P^{-1}(\mu)$ as an implicit generalized Hamiltonian system. Although for the explicit Hamiltonian system $(\mathcal{X}_c, D_c, H_c)$, so for the *reduced* generalized Dirac structure D_c, assumption (ii) *is* always satisfied, like in examples 8 and 9, this is in general *not* the case for the *original* generalized Dirac structure D. We could only show that $V + \mathsf{G}_0$ and $D \cap E_q$ are constant dimensional *on* $\mathcal{X}_c$.

8 Conclusions

In this paper we studied the notion of symmetry for implicit generalized Hamiltonian systems, as defined in [5, 14]. We derived some results concerning symmetries, first integrals and Casimir functions. Furthermore, we investigated the possibilities of reducing an implicit generalized Hamiltonian system on a manifold $\mathcal{X}$ to a system on a submanifold of $\mathcal{X}$, e.g. a level set of first integrals, or a quotient manifold $\mathcal{X}/G$, where G is a strong symmetry Lie group of the implicit generalized Hamiltonian system. We proved that, under some assumptions, these reduced systems are again implicit generalized Hamiltonian systems. Finally, we investigated the reduction of implicit generalized

Hamiltonian systems, having a strong symmetry Lie group generated by first integrals. It turns out that first reducing the system to a level set of the first integrals and then factoring out the remaining symmetries is equivalent to first factoring out the symmetry group and then reducing the system to a level set of the remaining first integrals (which are now Casimir functions). This is our main result, and is a generalization of the classical reduction theorems for (explicit) symplectic and Poisson Hamiltonian systems described in [1, 7, 9, 12]. The general setting, using the geometric notion of a Dirac structure and correspondingly implicit generalized Hamiltonian systems, makes the theory applicable to mechanical systems with nonholonomic constraints, and in general to interconnected multibody systems, as well as electromechanical systems which can be described in this Hamiltonian format.

References

[1] R. Abraham and J.E. Marsden. *Foundations of Mechanics*. Benjamin/Cummings Publishing Company, second edition, 1978.

[2] R. Abraham, J.E. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis, and Applications*. Springer-Verlag, second edition, 1988.

[3] T. Courant. Dirac manifolds. *Trans. American Math. Soc.*, 319:631–661, 1990.

[4] M. Dalsmo and A.J. van der Schaft. On representations and integrability of mathematical structures in energy-conserving physical systems. *SIAM J. Cont. Opt.*, 37(1):54–91, 1999.

[5] I. Dorfman. *Dirac Structures and Integrability of Nonlinear Evolution Equations*. Chichester: John Wiley, 1993.

[6] A. Isidori. *Nonlinear Control Systems*. Springer-Verlag, third edition, 1995.

[7] P. Libermann and C.-M. Marle. *Symplectic Geometry and Analytical Mechanics*. Reidel, Dordrecht, 1987.

[8] J.E. Marsden. *Lectures on Mechanics*. Cambridge University Press, 1992.

[9] J.E. Marsden and T. Ratiu. Reduction of Poisson manifolds. *Letters in Mathematical Physics*, 11:161–169, 1986.

[10] B.M. Maschke and A.J. van der Schaft. Interconnected mechanical systems, part II: the dynamics of spatial mechanical networks. In A. Astolfi, D.J.N Limebeer, C. Melchiorri, A. Tornambè, and R.B. Vinter, editors, *Modelling and Control of Mechanical Systems*, pages 17–30. Imperial College Press, 1997.

[11] H. Nijmeijer and A.J. van der Schaft. *Nonlinear Dynamical Control Systems*. Springer-Verlag, 1990.

[12] P.J. Olver. *Applications of Lie Groups to Differential Equations*. Springer-Verlag, second edition, 1993.

[13] S. Smale. Topology and mechanics (I,II). *Inv. Math.*, 10,11:305–331,45–64, 1970.

[14] A.J. van der Schaft. Implicit Hamiltonian Systems with Symmetry. *Rep. Math. Phys.*, 41:203–221, 1998.

[15] A.J. van der Schaft and B.M Maschke. The Hamiltonian formulation of energy conserving physical systems with external ports. *Archiv für Elektronik und Übertragungstechnik*, 49:362–371, 1995.

[16] A.J. van der Schaft and B.M Maschke. Interconnected mechanical systems, part I: geometry of interconnection and implicit Hamiltonian systems. In A. Astolfi, D.J.N Limebeer, C. Melchiorri, A. Tornambè, and R.B. Vinter, editors, *Modelling and Control of Mechanical Systems*, pages 1–15. Imperial College Press, 1997.