Resenhas IME-USP 2003, Vol. 6, No. 1, 1 - 71.

On the Birkhoff approach to classical mechanics

Marcelo H Kobayashi and Waldyr M Oliva

Abstract: The notion of Birkhoffian allows the introduction of the intrinsic concepts of reversibility, reciprocity, regularity, affine structure in the accelerations, conservation of energy, as well as the Principle of Gibbs-Appell on the energy of the acceleration. The Birkhoff systems generalize the Lagrangian systems either regular or not, being possible to provide an external characterization for the latter within this framework: the so called inverse problem of Lagrangian mechanics. It is also possible to introduce, in a natural manner, the notion of constrained Birkhoff systems beyond the classical affine constraints. Symmetry and reduction as well as conservation of volume are analyzed for constrained and unconstrained Birkhoff systems.

1 Introduction

Let us consider a Lagrangian system of equations (see Abraham & Marsden (1978))

$$Q_i^L(q,\dot{q},\ddot{q}) := \frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{q}^i} - \frac{\partial \mathrm{L}}{\partial q^i} = 0, \quad i = 1,\ldots,n, \tag{1}$$

corresponding to an autonomous C^2-Lagrangian function $\mathrm{L} = \mathrm{L}(q,\dot{q})$. We point out the following properties of the left hand side of such equations.

i) They define a second order system of ordinary differential equations (may be in implicit quasi-linear form)

$$Q_i^L(q,\dot{q},\ddot{q}) := \frac{\partial^2 \mathrm{L}}{\partial \dot{q}^j \partial \dot{q}^i}\ddot{q}^j + \frac{\partial^2 \mathrm{L}}{\partial q^j \partial \dot{q}^i}\dot{q}^j - \frac{\partial \mathrm{L}}{\partial q^i} = 0, \quad i = 1,\ldots,n, \tag{2}$$

ii) Under the so called Legendre condition, that is,

$$\det\left(\frac{\partial^2 \mathrm{L}}{\partial \dot{q}^j \partial \dot{q}^i}\right) \neq 0 \text{ everywhere}, \tag{3}$$

the Lagrangian L is regular (Abraham & Marsden (1978)) and the system in i) can be solved for $\ddot{q}$; so one obtains, in the neighborhood of each point $(q_0,\dot{q}_0,\ddot{q}_0)$, a system in explicit form

$$\ddot{q}^i = h^i(q,\dot{q}), \quad i = 1,\ldots,n, \tag{4}$$

Remark that regularity means $\det\left(\frac{\partial Q_i^L}{\partial \ddot{q}^j}\right) \neq 0$, everywhere.

iii) The left hand sides $Q_i^L(q,\dot q,\ddot q)$, $i=1,\ldots,n$, have covariant character. In fact, let $\bar q=(\bar q^i)$ be another system of position coordinates and construct, analogously, the functions

$$\bar Q_l^L=\bar Q_l^L(\bar q,\dot{\bar q},\ddot{\bar q}),\quad l=1,\ldots,n.$$

After a C^2 change of coordinates given by $q^k=q^k(\bar q^1,\ldots,\bar q^n)$, $k=1,\ldots,n$, one obtains

$$\bar Q_l^L=Q_i^L\frac{\partial q^i}{\partial \bar q^l},\quad i,l=1,\ldots,n, \tag{5}$$

with the right hand sides being functions of $(\bar q,\dot{\bar q},\ddot{\bar q})$ through the "natural" transformation:

$$\left\{\begin{array}{l} q^k=q^k(\bar q^1,\ldots,\bar q^n)\\ \dot q^k=\frac{\partial q^k}{\partial \bar q^j}(\bar q)\dot{\bar q}^j\\ \ddot q^k=\frac{\partial q^k}{\partial \bar q^j\partial \bar q^r}\dot{\bar q}^r\dot{\bar q}^j+\frac{\partial q^k}{\partial \bar q^j}\ddot{\bar q}^j,\quad j,k,r=1,\ldots,n\end{array}\right. \tag{6}$$

iv) The Lagrangian system $Q_i^L(q,\dot q,\ddot q):=\frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot q^i}-\frac{\partial \mathrm{L}}{\partial q^i}=0$, $i=1,\ldots,n$, is "conservative" in the sense that the "work" $\left(\int_A^C Q_j^L \mathrm{d}q^j\right)$ done along a piecewise C^1 path $AC=(q(t),\dot q(t))$ by the functions $Q_i^L(q,\dot q,\ddot q)$ is independent of the path taken, and so it depends only upon the values of $(q^1,\ldots,q^n,\dot q^1,\ldots,\dot q^n)$ at A and C. In other words, the Lagrangian system (1) satisfies the "Principle of the conservation of energy".

The quantity

$$E_L(q,\dot q):=\frac{\partial \mathrm{L}}{\partial \dot q^j}\dot q^j-\mathrm{L} \tag{7}$$

is the "energy" of L and is conserved along any path defined by a solution $q(t)$ of (1).

In 1927, G.D. Birkhoff, is his celebrated book *Dynamical Systems* (see Birkhoff (1927)) extended the properties i) to iii) to a (not necessarily Lagrangian) system of second differential order equations

$$Q_i(q,\dot q,\ddot q)=0,\quad i=1,\ldots,n, \tag{8}$$

with a physical interpretation for the n functions Q_i as "generalized forces" and $Q_i(q,\dot q,\ddot q)\mathrm{d}q^i$ as an "elementary work" done by these forces. Following Birkhoff, the generalized forces are assumed to have the covariant character, that is,

$$\bar Q_l=Q_i\frac{\partial q^i}{\partial \bar q^l},\quad i,l=1,\ldots,n, \tag{9}$$

for any natural change of coordinates of type (6). Regularity means, as above, $\det\left(\frac{\partial Q_i}{\partial \ddot{q}^j}\right) \neq 0$, everywhere. If, moreover, these forces satisfy the Principle of the conservation of energy (iv), in the sense that the work $\left(\int_A^C Q_j \mathrm{d}q^j\right)$ done, along a piecewise C^1 path, by the generalized forces $Q_i(q, \dot{q}, \ddot{q})$, is independent of the path, then (8) is called a conservative system.[1] Birkhoff gave also another characterization for the principle of conservation of the energy through a fundamental identity (see Birkhoff (1927), p.16, equation (4)).

In Birkhoff (1927) it is also introduced the notions of reversibility

$$Q_i(q, \dot{q}, \ddot{q}) = Q_i(q, -\dot{q}, \ddot{q}),$$

and "affine" generalized forces, that is,

$$Q_i(q, \dot{q}, \ddot{q}) = a_{ij}(q, \dot{q})\ddot{q}^j + b_i(q, \dot{q}), \quad i, j = 1, \ldots, n; \tag{10}$$

it is defined the "Principle of reciprocity" characterized by the symmetry condition

$$a_{ij}(q, \dot{q}) = a_{ji}(q, \dot{q}), \quad i, j = 1, \ldots, n, \tag{11}$$

and it is considered from (10) those forces Q_i for which

$$a_{ij} = a_{ij}(q), \quad i, j = 1, \ldots, n. \tag{12}$$

Moreover, he proved that most of these notions do not depend on the "natural coordinates" $(q, \dot{q}, \ddot{q})$ used in their definitions, that is, they are invariant under the transformations given in (6).

Finally, Birkhoff considered the so called "inverse problem of Lagrangian mechanics" by giving an external characterization for regular and non-singular (in his terminology) Lagrangian systems, that is, systems defined by the nowadays called "classical Lagrangians":

$$\mathrm{L} = \frac{1}{2} g_{ij}(q)\dot{q}^i\dot{q}^j - U(q) \tag{13}$$

with $\det(g_{ij}(q)) \neq 0$ everywhere (see Theorem 3 in the present paper).

After the seminal work of Smale in Smale (1970a) and Smale (1970b) and the publication of a series of books like those by Godbillon (1969), Abraham & Marsden (1978), Arnold (1989), Gallavotti (1983), Marsden (1992), Marsden & Ratiu (1999) and Oliva (2002) among others, the subject of Geometric Mechanics became more and more well posed and popular. In the present paper we come back to the Birkhoff approach to classical mechanics presenting all his ideas from the

[1] Remark made by Birkhoff in Birkhoff (1927), p. 4: "I presented the results here obtained at the Chicago Colloquium in 1920. The following treatment of the Principle of the conservation of energy differs essentially from any other which I have seen".

viewpoint of differential geometry and proving new results which will be discussed in the sequel.

As we will see in §2, given a smooth, connected n-dimensional manifold M, one starts by considering the tangent bundles $(\mathrm{TM}, \mathrm{M}, \tau_{\mathrm{M}})$ and $(\mathrm{T(TM)}, \mathrm{TM}, \tau_{\mathrm{TM}})$. The 2-jets manifold $J^2(\mathrm{M})$ is a $3n$-dimensional submanifold of $\mathrm{T(TM)}$ defined as

$$J^2(\mathrm{M}) = \{z \in \mathrm{T(TM)} \mid \tau_{\mathrm{TM}}(z) = \mathrm{T}\tau_{\mathrm{M}}(z)\},$$

$\mathrm{T}\tau_{\mathrm{M}}$ being the tangent map of τ_{M}.

Call $\tau_J := \tau_{\mathrm{TM}}|_{J^2(\mathrm{M})} = \mathrm{T}\tau_{\mathrm{M}}|_{J^2(\mathrm{M})}$. The second order vector fields Y on TM are such that $\mathrm{T}\tau_{\mathrm{M}}(Y_v) = v$, for all $v \in \mathrm{TM}$, so any such Y has values on $J^2(\mathrm{M})$ and we can identify any second order vector field with a cross section X of the bundle $\tau_J : J^2(\mathrm{M}) \to \mathrm{TM}$ because $Y = i \circ X$, where $i : J^2(\mathrm{M}) \to \mathrm{T(TM)}$ is the canonical inclusion.

A local system of coordinates (U, q), $q = (q^1, \dots, q^n)$, for M induces natural local coordinates $((\tau_{\mathrm{M}} \circ \tau)_J^{-1}(U); q, \dot{q}, \ddot{q})$ for $J^2(\mathrm{M})$.

A *Birkhoffian* (see §3) is a special Pfaffian form ω on $J^2(\mathrm{M})$ such that in the natural local coordinates is written as $\omega = Q_i(q, \dot{q}, \ddot{q})\mathrm{d}q^i$. One can show that $(J^2(\mathrm{M}), \mathrm{TM}, \tau_J)$ is an affine bundle and a Birkhoffian ω induces a morphism $\hat{\omega}$ from $(J^2(\mathrm{M}), \mathrm{TM}, \tau_J)$ into the cotangent bundle $(\mathrm{T}^*\mathrm{M}, \mathrm{M}, \tau_{\mathrm{M}}^*)$ (see Proposition 2). When $\hat{\omega}$ restricted to each fiber is an affine map we say that the Birkhoffian ω is *affine*. Any Birkhoffian ω induces also a map $\bar{\omega}$ (see Proposition 4) such that to each $z \in J^2(\mathrm{M})$, $\bar{\omega}(z)$ is a bilinear map defined on $\mathrm{T}_p\mathrm{M}$ in the following way:

$$\bar{\omega}(z)(u_p, w_p) := \frac{d}{dt}\left[\hat{\omega}(z + t\lambda_v u_p)w_p\right]_{t=0}$$

where $v = \tau_J(z)$, $p = \tau_{\mathrm{M}}(v)$, $u_p, w_p \in \mathrm{T}_p\mathrm{M}$ and $\lambda_v u_p$ is the vertical lifting of $u_p \in \mathrm{T}_p\mathrm{M}$ to $\mathrm{T}_v(\mathrm{TM})$. In natural local coordinates, if $\omega = Q_i(q, \dot{q}, \ddot{q})\mathrm{d}q^i$ and $u_p = u^i \frac{\partial}{\partial q^i}(p)$, $w_p = w^i \frac{\partial}{\partial q^i}(p)$, we have

$$\bar{\omega}(z)(u_p, w_p) = \frac{\partial Q_i}{\partial \ddot{q}^j}(q, \dot{q}, \ddot{q})u^j w^i.$$

When $\bar{\omega}(z)$ is symmetric for all $z \in J^2(\mathrm{M})$, ω is said to satisfy the *Principle of reciprocity*, that is, $\frac{\partial Q_i}{\partial \ddot{q}^j} = \frac{\partial Q_j}{\partial \ddot{q}^i}$, $\forall i, j = 1, \dots, n$, in natural local coordinates.

When $\bar{\omega}(z)$ is non-degenerate (as a bilinear form) for all $z \in J^2(\mathrm{M})$, that is, $\det\left(\frac{\partial Q_i}{\partial \ddot{q}^j}\right) \neq 0$ everywhere in natural local coordinates, ω is said to be *regular* provided that for each $v \in \mathrm{TM}$, there exists $\bar{z} \in \tau_J^{-1}(v)$ such that $\hat{\omega}(\bar{z}) = 0$.

The *Principle of the conservation of energy* is introduced for a Birkhoffian ω (see Definition 13) and it is shown that ω is conservative if and only if there exists a smooth function $E_\omega : \mathrm{TM} \to \mathbb{R}$ (the energy of ω) such that

$$(X^*\omega)Y = \mathrm{d}E_\omega(Y)$$

for all second order vector field $Y = i \circ X$ on TM. That is equivalent, in natural local coordinates, to the identity

$$Q_i(q, \dot{q}, \ddot{q})\dot{q}^i = \frac{\partial E_\omega}{\partial q^i}\,\dot{q}^i + \frac{\partial E_\omega}{\partial \dot{q}^i}\,\ddot{q}^i \quad \text{(see Proposition 5).}$$

Any smooth Lagrangian function $\mathrm{L} : \mathrm{TM} \to \mathbb{R}$ defines a Birkhoffian ω_L that satisfies the principle of reciprocity, is affine and is the unique Birkhoffian such that

$$X^*\omega_L = \mathrm{i}_Y \mathsf{d}(\mathsf{d}_v \mathrm{L}) + \mathsf{d}(Z(\mathrm{L}) - \mathrm{L})$$

for all second order vector field $Y = i \circ X$ on TM. Here $Z(\mathrm{L}) - \mathrm{L}$ is the energy of ω_L where Z is the Liouville vector field, defined by $Z(v) = \lambda_v(v) \in \mathrm{T}_v(\mathrm{TM})$ for all $v \in \mathrm{TM}$. In natural local coordinates one can write

$$\omega_L = \left(\frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{q}^i} - \frac{\partial \mathrm{L}}{\partial q^i}\right)\mathrm{d}q^i$$

so ω_L is regular if, and only if, L is a regular Lagrangian.

The paper follows with a concrete Example 1 corresponding to the equations modelling the flight of rigid airplane. It is not conservative, in general, so does not fit into the Lagrangian formalism. Also, it has an implicit character and cannot be included into the general Godbillon approach (see Remark 5). Thus, in a sense, Birkhoff formalism extends not only the Lagrangian and Godbillon approaches but also includes many mathematical-physical models (conservative or not) involving implicit second order equations.

The notion of normal Birkoffian (see Definition 9 and Example 2) includes classical Lagrangian systems (kinetic energy minus potential energy).

The Gibbs and Gibbs-Appell functions as well as the Gibbs-Appell principle are also introduced in the context of Birkhoff systems (see Definition 12). These concepts are usually associated with constrained mechanical systems, where it provides a simple form for deriving the equations of motion. Their derivation goes back to the works of Gibbs (Gibbs (1879)) and Appell (Appell (1900*b*), Appell (1900*a*)). It can be found in classical texts as Pars (1965) and more recently they have been generalized to regular Lagrangian systems within a geometric framework in Lewis (1996) (see also the references therein).

The principle of the conservation of energy is completely described and characterized locally as well as globally. The Birkhoffian $\omega = \omega_L$ defined by a smooth Lagrangian has its main properties showed in Proposition 7. The Principle of determinism holds for any regular Birkhoffian (see Theorem 1).

The inverse problem of the Lagrangian mechanics is the search for necessary and sufficient conditions for a Birkhoffian ω to be equal to a Birkhoffian ω_L corresponding to a Lagrangian function L. This problem has known a long history, dating back to Helmholtz. The reader is referred to Anderson & Thompson (1992) and the references therein for a throughout treatment on the subject. The main ingredient of the modern treatment is the so called variational bicomplex of differential forms on the infinite jet bundle of some fibred manifold $\pi : E \to \mathrm{M}$. Roughly speaking, one of the differentials in the variational bicomplex may be identified with the Euler-Lagrange operator, while the other differential provides the intrinsic characterization of the Helmholtz conditions (see also Massa & Pagani (1994)). However, following Birkhoff's ideas we propose to give a simpler external characterization of Lagrangian systems. The case of a classical Lagrangian is developed in Theorems 2 and 3 and the general Lagrangian case is considered in Theorem 5. The first case is related with the property of an associated Gibbs-Appell function and the more general case with the classical Helmholtz conditions (see Definition 14) whose geometrical (intrinsic) meaning is also described (see Theorem 4). We also deal with the extension of the classical Liouville's Theorem to the non-conservative case (see Remark 10).

In section 4, we extend the notion of Birkhoff systems to constrained Birkhoff systems. The constraint is defined as an affine sub-bundle of the affine bundle $J^2(\mathrm{M})$. The Principle of d'Alembert-Birkhoff is introduced (see Definition 19) and the existence and uniqueness for regular constrained systems (see Definition 18) is proved in Theorem 6. The Gauss Principle of least constraint is also introduced in the context of constrained Birkhoff systems (see Definition 21) and the equivalence for affine and symmetric Birkhoffians between this principle and the d'Alembert-Birkhoff principle is proved in Proposition 12. The notion of *reaction field* is also introduced (see Definition 20). The constrained inverse problem is proved in Corollary 2. We close this section with a proof of the Liouville theorem for classical affine constraints and classical Lagrangian systems.

The last part of the paper concerns the study of symmetry and reduction of Birkhoff systems. The action of a Lie group G, the momentum mapping and the implications for the symmetry and reduction of Birkhoff systems are also considered. A version of the Noether theorem is proved (see Proposition 13) for Birkhoff systems admitting a momentum mapping, generalizing the Lagrangian situation and can be effected even in the non-conservative case (see Proposition 16 and the concrete example that follows, which considers the rigid body model of an artificial satellite orbiting around the Earth under the influence of the gravity and drag, only). We also give a characterization for the concept of Lagrangian function invariant under the tangent action (see Theorem 10). The paper ends with a generalization to constrained Birkhoff systems of the symmetry and reduction procedures developed for the unconstrained Birkhoff systems.

2 Preliminaries

Let M be a smooth finite-dimensional differentiable connected manifold, $\dim \mathrm{M} = n$, and $\tau_{\mathrm{M}} : \mathrm{TM} \to \mathrm{M}$, $\tau^*_{\mathrm{M}} : \mathrm{T}^*\mathrm{M} \to \mathrm{M}$ denote the tangent and cotangent bundles of M, respectively. If $(U, q^1, \dots, q^n)$ is a local coordinate system in $U \subset \mathrm{M}$, then $\dot{q}^i : \mathrm{T}U = \tau^{-1}(U) \to \mathbb{R}$ defined as $\dot{q}^i(v) = \mathrm{d}q^i(v)$ for all $v \in \mathrm{T}U$ enable us to construct natural local coordinates $(\mathrm{T}U, q^1 \circ \tau_{\mathrm{M}}, \dots, q^n \circ \tau_{\mathrm{M}}, \dot{q}^1, \dots, \dot{q}^n)$. For simplicity we write $(\mathrm{T}U, q, \dot{q})$ with the identification $q^i \sim q^i \circ \tau_{\mathrm{M}}$. Analogously, using $\tau_{\mathrm{TM}} : \mathrm{T(TM)} :\longrightarrow \mathrm{TM}$, one can construct natural local coordinates for $\mathrm{T(TM)}$: $(\mathrm{T}(\mathrm{T}U) = \tau_{\mathrm{TM}}^{-1} \circ \tau_{\mathrm{M}}^{-1}(U); q \circ \tau_{\mathrm{TM}}, \dot{q} \circ \tau_{\mathrm{TM}}, \mathrm{d}q, \mathrm{d}\dot{q})$, where $\mathrm{d}q^i$ means $\mathrm{d}_{\mathrm{TM}} q^i$ and $\mathrm{d}\dot{q}^i$ means $\mathrm{d}_{\mathrm{TM}} \dot{q}^i$. Also, for simplicity, we keep denoting this last natural system of coordinates by $(\mathrm{T}(\mathrm{T}U); q, \dot{q}, \mathrm{d}q, \mathrm{d}\dot{q})$.

Denoting by $\mathrm{T}\tau_{\mathrm{M}}$ the tangent map of τ_{M} we obtain the following "rhombic" commutative diagram:

$$
\begin{tikzcd}
 & \mathrm{T(TM)} \arrow[dl, "\tau_{\mathrm{TM}}"'] \arrow[dr, "\mathrm{T}\tau_{\mathrm{M}}"] & \\
\mathrm{TM} \arrow[dr, "\tau_{\mathrm{M}}"'] & & \mathrm{TM} \arrow[dl, "\tau_{\mathrm{M}}"] \\
 & \mathrm{M} &
\end{tikzcd}
$$

In natural local coordinates we set

$$
\begin{array}{lcllcl}
\mathrm{T(TM)} & \xrightarrow{\mathrm{T}\tau_{\mathrm{M}}} & \mathrm{TM} & \text{or} & (q, \dot{q}, \mathrm{d}q, \mathrm{d}\dot{q}) & \longrightarrow & (q, \mathrm{d}q) \\
\mathrm{T(TM)} & \xrightarrow{\tau_{\mathrm{TM}}} & \mathrm{TM} & \text{or} & (q, \dot{q}, \mathrm{d}q, \mathrm{d}\dot{q}) & \longrightarrow & (q, \dot{q})
\end{array}
$$

2.1 The vertical vector bundle

For each $v \in \mathrm{TM}$, define the vector space

$$
\mathrm{V}_v(\mathrm{M}) = \{\xi \in \mathrm{T}_v(\mathrm{TM}) \mid (\mathrm{T}\tau_{\mathrm{M}}(v))\, \xi = 0\},
$$

to be the set of tangent vertical vectors at $v \in \mathrm{TM}$. So, it is well defined the vector bundle $\mathrm{V(M)} \xrightarrow{\rho} \mathrm{TM}$ where $\rho = \tau_{\mathrm{TM}}|_{\mathrm{V(M)}}$ and $\rho^{-1}(v) = \mathrm{V}_v(\mathrm{M})$ for each $v \in \mathrm{TM}$.

In natural coordinates $(q, \dot{q}, \mathrm{d}q, \mathrm{d}\dot{q})$ of $\mathrm{T(TM)}$, $\mathrm{V(M)}$ is defined by $\mathrm{d}q = 0$, that is, the elements of $\mathrm{V(M)}$ are given, locally, in natural coordinates, by $(q, \dot{q}, 0, \mathrm{d}\dot{q})$. The elements of the n-dimensional vector space $\mathrm{V}_v(\mathrm{M})$ are called the *tangent vertical vectors* at $v \in \mathrm{TM}$.

2.1.1 The Liouville vector-field Z on TM

The *Liouville vector-field* is the vector field Z on TM defined by

$$
Z(v) = \lambda_v v \in \mathrm{T}_v(\mathrm{TM})
$$

where $\lambda_v v$ is the vertical lifting (see, for example, Marle (1995)) of $v \in \mathrm{T}_{\tau_{\mathrm{M}}(v)}\mathrm{M}$ to $\mathrm{T}_v(\mathrm{TM})$. The (natural) local expression of the Liouville vector field at $v \in \mathrm{TM}$ is $Z(v) = \dot{q}^i(v) \frac{\partial}{\partial \dot{q}^i}(v)$, where, unless otherwise stated, we use the summation convention throughout. On the other hand the flow of Z through $(q,\dot{q}) \in \mathrm{TM}$ satisfies the ODE: $\frac{dq}{dt} = 0$, $\frac{d\dot{q}}{dt} = \dot{q}$, with initial conditions $q(0) = q_0$ and $\dot{q}(0) = \dot{q}_0$. So, $q(t) = q_0$ and $\dot{q}(t) = \dot{q}_0 e^t$. In other words, Z generates the one parameter group of homotheties of TM and vanishes on the zero section of TM and only there.

2.2 The 2-jets manifold $J^2(\mathrm{M})$

Let us introduce, now the following $3n$-dimensional manifold

$$J^2(\mathrm{M}) = \{z \in \mathrm{T(TM)} \mid \tau_{\mathrm{TM}}(z) = \mathrm{T}\tau_{\mathrm{M}}(z)\}$$

Restricting τ_{TM} and $\mathrm{T}\tau_{\mathrm{M}}$ to $J^2(\mathrm{M})$ one obtains, $\tau_J := \tau_{\mathrm{TM}}|_{J^2(\mathrm{M})} = \mathrm{T}\tau_{\mathrm{M}}|_{J^2(\mathrm{M})}$. That is, we have the following commutative diagram:

$$\begin{tikzcd}
 & J^2(\mathrm{M}) \arrow[dl, "\tau_{\mathrm{TM}}|_{J^2(\mathrm{M})}"'] \arrow[dr, "\mathrm{T}\tau_{\mathrm{M}}|_{J^2(\mathrm{M})}"] & \\
\mathrm{TM} \arrow[rr, "\mathrm{id}"] \arrow[dr, "\tau_{\mathrm{M}}"'] & & \mathrm{TM} \arrow[dl, "\tau_{\mathrm{M}}"] \\
 & \mathrm{M} &
\end{tikzcd}$$

We also define $\beta = \beta_{\mathrm{M}} : J^2(\mathrm{M}) \to \mathrm{M}$ by $\beta(z) := (\tau_{\mathrm{M}} \circ \tau_{\mathrm{TM}})(z)$ for all $z \in J^2(\mathrm{M})$.

In natural local coordinates the elements of $J^2(\mathrm{M})$ are of the form $(q, \dot{q}, \dot{q}, \mathrm{d}\dot{q})$, because $(q, \mathrm{d}q) = (q, \dot{q})$. We denote by $\ddot{q} := \mathrm{d}\dot{q}|_{J^2(\mathrm{M})}$. The cross sections (resp. local cross sections) of $\tau_J : J^2(\mathrm{M}) \to \mathrm{TM}$ are identified with the so called *second order vector fields* (resp. *local second order vector fields*). Recall that any vector field Y on TM is a second order vector field if and only if $\mathrm{T}\tau_{\mathrm{M}}(Y_v) = v$ for all $v \in \mathrm{TM}$. Also, for any cross section X of $\tau_J : J^2(\mathrm{M}) \to \mathrm{TM}$ there is a unique second order vector field Y on TM which satisfies $Y = i \circ X$, where $i : J^2(\mathrm{M}) \to \mathrm{T(TM)}$ is the canonical embedding. In local natural coordinates a second order vector field can be represented as

$$Y = \dot{q}^i \frac{\partial}{\partial q^i} + \ddot{q}^i(q,\dot{q}) \frac{\partial}{\partial \dot{q}^i}\,.$$

For the sake of simplicity, we may also write

$$Y = \dot{q}^i \frac{\partial}{\partial q^i} + \ddot{q}^i \frac{\partial}{\partial \dot{q}^i}\,.$$

The vertical space $\mathrm{V}_v(\mathrm{M})$ acts transitively on $\tau_J^{-1}(v)$. Indeed, if $\xi \in \mathrm{V}_v(\mathrm{M})$ and $z \in J_v^2(\mathrm{M})$, one uses the sum of $\mathrm{T}_v(\mathrm{TM})$ to construct $\xi + z \in J_v^2(\mathrm{M})$. In fact, in local natural coordinates we have

$$(q(v), \dot{q}(v), 0, \mathrm{d}\dot{q}(\xi)) + (q(v), \dot{q}(v), \dot{q}(z), \mathrm{d}\dot{q}(z)) = (q(v), \dot{q}(v), \dot{q}(z), \mathrm{d}\dot{q}(z+\xi)).$$

It is easy to check the axioms of an affine space:

1. $\xi_1 + (\xi_2 + z) = (\xi_1 + \xi_2) + z$, $\xi_1, \xi_2 \in \mathrm{V}_v(\mathrm{M})$, $z \in J^2_v(\mathrm{M})$;
2. $0 + z = z$, for all $z \in J^2_v(\mathrm{M})$;
3. given $z, \bar{z} \in J^2_v(\mathrm{M})$, there exists a unique $\xi \in \mathrm{V}_v(\mathrm{M})$ such that $z + \xi = \bar{z}$.

Proposition 1. *$\tau_J : J^2(\mathrm{M}) \to \mathrm{TM}$ is an affine bundle, modelled by the vertical vector bundle $\rho : \mathrm{V}(\mathrm{M}) \to \mathrm{TM}$ (see Goldschmidt (1967) and Lewis (1996)).*

Definition 1. *Let $f : N \to \bar{N}$ be a smooth map of manifolds; we define the affine bundle morphism $J^2 f : J^2(N) \to J^2(\bar{N})$ by:*

$$J^2 f(z) = T^2 f(z)$$

for all $z \in J^2(N)$.

3 Birkhoff systems

Looking at the diagram

$$J^2(\mathrm{M}) \xrightarrow{\tau_J} \mathrm{TM} \xrightarrow{\tau_\mathrm{M}} \mathrm{M}$$

the mapping $\beta = \tau_\mathrm{M} \circ \tau_J : J^2(\mathrm{M}) \to \mathrm{M}$ is a smooth fibration.

Definition 2. *A* Birkhoffian *of the* configuration space M *is a smooth Pfaffian form ω on $J^2(\mathrm{M})$ such that, for any $p \in \mathrm{M}$, we have $i_p^* \omega = 0$, where $i_p : \beta^{-1}(p) \to J^2(\mathrm{M})$ is the embedding of the submanifold $\beta^{-1}(p)$ into $J^2(\mathrm{M})$. The pair (M, ω) is said to be a* Birkhoff system.

From the last definition it follows that, in natural local coordinates $(q, \dot{q}, \ddot{q})$ of $J^2(\mathrm{M})$ ($\ddot{q}$ represents $\mathrm{d}\dot{q}|_{J^2(\mathrm{M})}$), a Birkhoffian ω is given by $\omega = Q_i(q, \dot{q}, \ddot{q}) \mathrm{d}q^i$.

Proposition 2. *A Birkhoffian ω of the configuration space* M *induces a smooth morphism $\hat{\omega}$ between affine bundles such that the following diagram is commutative:*

$$\begin{tikzcd}
J^2(\mathrm{M}) \arrow[r, "\hat{\omega}"] \arrow[d, "\tau_J"'] & \mathrm{T}^*\mathrm{M} \arrow[d, "\tau^*_\mathrm{M}"] \\
\mathrm{TM} \arrow[r, "\tau_\mathrm{M}"'] & \mathrm{M}
\end{tikzcd}$$

Proof. Let $z \in J^2(\mathrm{M})$, $v = \tau_J(z)$ and $p = \tau_\mathrm{M}(v)$. For any $w_p \in \mathrm{T}_p\mathrm{M}$ one defines $\hat{\omega}_p(w_p) := \omega_z(\mu_z)$ where $\mu_z \in \mathrm{T}_z J^2(\mathrm{M})$ is any vector such that $\beta_*(z)(\mu_z) = w_p$. $\hat{\omega}_p$ is well defined because, if $\bar{\mu}_z \in \mathrm{T}_z J^2(\mathrm{M})$ also satisfies $\beta_*(z)(\bar{\mu}_z) = w_p = \beta_*(z)(\mu_z)$, one sees that $\beta_*(z)(\bar{\mu}_z - \mu_z) = 0$, that is, $\bar{\mu}_z - \mu_z$ is tangent to $\beta^{-1}(p)$ at $z \in \beta^{-1}(p)$. So $\omega_z(\bar{\mu}_z - \mu_z) = 0$ and then $\omega_z(\bar{\mu}_z) = \omega_z(\mu_z)$. □

Remark 1. Given a Birkhoff system (M, ω), then, for each cross section X of $J^2(\mathrm{M})$, we have $(X^*\omega)\xi = \hat{\omega}(X)(\mathrm{T}\tau_{\mathrm{M}}\xi)$, for all vector fields ξ on TM. Indeed, in local natural coordinates $(q, \dot{q}, \ddot{q})$, if $X(q, \dot{q}) = (q, \dot{q}, \ddot{q}_X)$, then

$$(X^*\omega)\xi = Q_i(q, \dot{q}, \ddot{q}_X)\xi^i = \hat{\omega}(X)(\mathrm{T}\tau_{\mathrm{M}}\xi)$$

where $\xi = \xi^i \frac{\partial}{\partial q^i} + \dot{\xi}^i \frac{\partial}{\partial \dot{q}^i}$.

Definition 3. *The* differential system $D(\omega)$ *associated to a Birkhoffian* ω *of* M *(see Oliva (1970)) is the set (may be empty) given by*

$$D(\omega) := \left\{ z \in J^2(\mathrm{M}) \mid \omega(z) = 0 \right\}.$$

Let I *be an open interval of* $\mathbb{R}$. *Any smooth curve* $\phi : t \in I \mapsto \phi(t) \in \mathrm{M}$ *defines a lifting* $\frac{T}{dt}\left(\frac{T\phi}{dt}\right) : I \to J^2(\mathrm{M})$.

A motion of ω *is a smooth curve* ϕ *whose lifting to* $J^2(\mathrm{M})$ *has image contained in* $D(\omega)$.

Locally, in natural coordinates, if $\omega = Q_i \mathrm{d}q^i$, the differential system $D(\omega)$ is characterized by the following implicit system of second order ODE:

$$Q_i(q = \phi(t), \dot{q} = \dot{\phi}(t), \ddot{q} = \ddot{\phi}(t)) = 0$$

for all $i = 1, \ldots, n$, and motions of ω are the solutions $q = \phi(t)$ of the implicit system of ODE above.

Definition 4. *A* Birkhoff vector field *associated to a Birkhoffian* ω *of the configuration space* M *is a smooth second order vector field* $Y = i \circ X$, $X : \mathrm{TM} \to J^2(\mathrm{M})$ *such that* $\mathrm{Im}\ X \subset D(\omega)$, *that is, whose base curves on* M *are motions of* ω, *or equivalently, if* $X^*\omega = 0$. *Analogously, one defines the notion of* local Birkhoff vector-fields.

Definition 5. *A Birkhoffian* ω *is* regular *if* $\bar{\omega}(z)$ *is non-degenerate for all* $z \in J^2(\mathrm{M})$ *and for each* $v \in \mathrm{TM}$, *there exists* $\bar{z} \in \tau_J^{-1}(v)$ *such that* $\hat{\omega}(\bar{z}) = 0$.

Locally, this means that

$$\det\left[\frac{\partial Q_i}{\partial \ddot{q}^j}(q, \dot{q}, \ddot{q})\right]_{i,j=1,\ldots,n} \neq 0$$

and for each $(q, \dot{q})$, there exists $\ddot{q}$ such that $Q_i(q, \dot{q}, \ddot{q}) = 0$, $i = 1, \ldots, n$. Note that the Birkhoffian $\omega = e^{\ddot{x}}\mathrm{d}x$ of $\mathbb{R}$ is such that $\det\left[\frac{\partial Q}{\partial \ddot{x}}\right] = e^{\ddot{x}} \neq 0$, but is not regular, since there is no $(x, \dot{x}, \ddot{x}) \in J^2(\mathbb{R})$ such that $Q(x, \dot{x}, \ddot{x}) = 0$. In other words, we can not dispense with the second requirement in the previous definition.

Lemma 1. *Let* $(\mathsf{E}, \|\|)$ *be a Banach space and* $f : U \to \mathsf{E}$ *a* C^1*-mapping defined on an open convex set* $U \subset \mathsf{E}$. *If for all* $v \in U$ *the derivative* $\mathrm{D}f(v) \in \mathrm{End}(\mathsf{E})$ *is injective, then there exists at most one* $v \in \mathsf{E}$ *such that* $f(v) = 0$.

Proof. Suppose, by contradiction, that there are two distinct points $v_0, v_1 \in \mathsf{E}$, such that $f(v_0) = f(v_1) = 0$. Given a continuous linear functional $\varphi \in \mathsf{E}'$, we define the smooth function $g_\varphi : [0,1] \to \mathbb{R}$ as follows

$$g_\varphi : t \mapsto \varphi(f(v_1 t + (1-t)v_0)).$$

By Rolle's Theorem and a corollary of the Hanh-Banach Theorem (see Corollary 2, p. 108 in Yosida (1980)), there is a sequence (φ_i), $i \in \mathbb{N}$ of continuous linear functionals and a sequence of vectors (w_i), $i \in \mathbb{N}$ satisfying

$$\|\varphi_i\| = 1 \quad \& \quad \varphi_i(w_i) = 0, i = 0, \ldots,$$

$$\varphi_0(w) = \|w\| \quad \& \quad \varphi_i(w_{i-1}) = \|w_{i-1}\|, i = 1, \ldots,$$

and

$$w_i = \mathrm{D}f(v_1 c_i + (1-c_i)v_0) \cdot w.$$

with, $c_i \in]0,1[$, for all $i \in \mathbb{N}$. By compactness of the unit interval, there is a converging subsequence $c_{i_k} \in]0,1[$, $k \in \mathbb{N}$, say $c_{i_k} \to c \in [0,1]$; smoothness of f yields

$$w_{i_k} = \mathrm{D}f(v_1 c_{i_k} + (1-c_{i_k})v_0) \cdot w \to \bar{w} := \mathrm{D}f(v_1 c + (1-c)v_0) \cdot w,$$

whence

$$0 = \varphi_{i_k}(w_{i_k}) = \varphi_{i_k}(w_{i_k} - w_{i_k-1}) + \varphi_{i_k}(w_{i_k-1}) \to \|\bar{w}\|.$$

Thus,

$$\bar{w} = \mathrm{D}f(v_1 c + (1-c)v_0) \cdot w = 0$$

which contradicts the hypothesis of $\mathrm{D}f$ being injective everywhere on U and the proof is complete. □

Theorem 1 (The principle of determinism). *If a Birkhoffian ω of* M *is regular, then it satisfies the principle of determinism, that is, there exists a unique Birkhoff vector field $Y = i \circ X$ associated to ω such that* $\operatorname{Im} X = D(\omega)$.

Proof. Applying Lemma 1 to the local representative of ω restricted to each fiber of $\tau_J : J^2(\mathrm{M}) \to \mathrm{TM}$, $(Q_1(q,\dot{q},\cdot), \ldots, Q_n(q,\dot{q},\cdot)) : \mathbb{R}^n \to \mathbb{R}^n$, yields a unique $\bar{z}$ in Definition 5; then, by using the Implicit Function Theorem locally in natural coordinates to the system of ODE $Q_i(q,\dot{q},\ddot{q}) = 0$, $i = 1, \ldots, n$, we see that $D(\omega)$ is locally equivalent to a smooth second order system of ODE in normal form, that is, of the type

$$\ddot{q}^i = f^i(q,\dot{q})$$

for all $i = 1, \ldots, n$. The existence and uniqueness of the solution as well as the differentiable dependence of solutions with respect to initial conditions, show the existence and uniqueness of a Birkhoff vector field $Y = i \circ X$ such that $\operatorname{Im} X = D(\omega)$. □

Let $\rho : \mathrm{TM} \to \mathrm{TM}$ be the mapping $\mathrm{TM} \ni v \mapsto -v \in \mathrm{TM}$. Besides the mapping between affine bundles $\hat{\omega}$ defined above, we may also define the mapping $\tilde{\omega} : J^2(\mathrm{M}) \to \mathrm{T}^*\mathrm{M}$, between affine bundles, making the following diagram commutative:

$$\begin{tikzcd} J^2(\mathrm{M}) \arrow[r, "\hat{\omega}"] \arrow[d, "\rho \circ \tau_J"'] & \mathrm{T}^*\mathrm{M} \arrow[d, "\tau^*_{\mathrm{M}}"] \\ \mathrm{TM} \arrow[r, "\tau_{\mathrm{M}}"'] & \mathrm{M} \end{tikzcd}$$

Definition 6. *A Birkhoffian ω of the configuration space* M *is said to be a* reversible *or to satisfy the* principle of reversibility *if $\hat{\omega} = \tilde{\omega}$.*

In local coordinates this condition is equivalent to $Q_i(q, \dot{q}, \ddot{q}) = Q_i(q, -\dot{q}, \ddot{q})$, and from this we obtain,

Proposition 3. *Let (M, ω) be a Birkhoff system with a reversible Birkhoffian ω and let $c(t)$ be an integral curve of a Birkhoff vector field Y. Then, $\rho \circ c(-t)$ is also an integral curve of Y. Thus, $F_{-t}(x) = \rho(F_t(\rho(x)))$, where F_t is the flow of Y.*

Proposition 4. *A Birkhoffian ω of the configuration space* M *induces a smooth morphism $\bar{\omega}$ between fibres bundles such that the following diagram is commutative*

$$\begin{tikzcd} J^2(\mathrm{M}) \arrow[r, "\bar{\omega}"] \arrow[d, "\tau_J"'] & \mathrm{T}^0_2(\mathrm{M}) \arrow[d, "(\tau^0_2)_{\mathrm{M}}"] \\ \mathrm{TM} \arrow[r, "\tau_{\mathrm{M}}"'] & \mathrm{M} \end{tikzcd}$$

where $\mathrm{T}^0_2(\mathrm{M}) = \bigcup_{p \in \mathrm{M}} \mathrm{T}^0_2(\mathrm{T}_p\mathrm{M})$ is the smooth vector bundle of bilinear forms on TM.

Proof. Let $z \in J^2(\mathrm{M})$, $v = \tau_J(z)$ and $p = \tau_{\mathrm{M}}(v)$. For any pair of vectors $u_p, w_p \in \mathrm{T}_p\mathrm{M}$ one defines

$$\bar{\omega}(z)(u_p, w_p) := \frac{d}{dt} \left\{ [\hat{\omega}\,(z + t\lambda_v u_p)]\, w_p \right\}\Big|_{t=0}$$

where $\lambda_v u_p$ is the vertical lifting of u_p to $\mathrm{T}_v(\mathrm{TM})$. It is easy to check that $\bar{\omega}(z)$ is bilinear. $\square$

In natural coordinates, taking $u_p = u^i \frac{\partial}{\partial q^i}\Big|_p$ and $w_p = w^i \frac{\partial}{\partial q^i}\Big|_p$ we can write for $\omega = Q_i \mathrm{d}q^i$:

$$\bar{\omega}(z)(u_p, w_p) = \frac{\partial Q_i}{\partial \ddot{q}^j}\Bigg|_{z\equiv(q(z),\dot{q}(z),\ddot{q}(z))} u^j w^i$$

Definition 7 (The principle of reciprocity). *A Birkhoffian ω satisfies* the principle of reciprocity *if $\bar{\omega}(z)$ is symmetric for all $z \in J^2(\mathrm{M})$.*

Locally, this means that

$$\frac{\partial Q_i}{\partial \ddot{q}^j} = \frac{\partial Q_j}{\partial \ddot{q}^i}$$

for all $i, j = 1, \ldots, n$.

Definition 8. *A Birkhoffian ω is said to be* affine with respect to accelerations *if the induced morphism $\hat{\omega}$ restricted to each fibre of $\tau_J : J^2(\mathrm{M}) \to \mathrm{TM}$ is an affine map.*

To write the local expression of a Birkhoffian affine with respect to accelerations, we start by fixing a second order vector field $S : \mathrm{TM} \to J^2(\mathrm{M})$. To say that $\hat{\omega}$ is an affine map when restricted to each fibre of $\tau_J : J^2(\mathrm{M}) \to \mathrm{TM}$, means that there exists a smooth map

$$v \in \mathrm{TM} \mapsto \tilde{S}_v \in \mathrm{Hom}\left(\mathrm{V}_v(\mathrm{M}), \mathrm{T}^*_{\tau_{\mathrm{M}}(v)}\mathrm{M}\right)$$

such that for each $z \in \tau_J^{-1}(v)$ we have

$$\hat{\omega}(z) = \hat{\omega}(S(v)) + \tilde{S}_v(z - S(v)). \tag{14}$$

In local natural coordinates, if $z = (q, \dot{q}, \ddot{q})$, $S(v) = (q, \dot{q}, \ddot{\bar{q}})$ and $a_{ij}(q, \dot{q})$ is the matrix of $\tilde{S}_v$ with respect to the basis $(\partial/\partial\dot{q}^i)_v$ and $(\mathrm{d}q^i)_{\tau_{\mathrm{M}}(v)}$ of $\mathrm{V}_v(\mathrm{M})$ and $\mathrm{T}^*_{\tau_{\mathrm{M}}(v)}\mathrm{M}$, respectively, we can write (14) in coordinates:

$$Q_i(q, \dot{q}, \ddot{q}) = Q_i(q, \dot{q}, \ddot{\bar{q}}) + a_{ij}(q, \dot{q})\left[\ddot{q}^j - \ddot{\bar{q}}^j\right]$$

or

$$Q_i(q, \dot{q}, \ddot{q}) = b_i(q, \dot{q}) + a_{ij}(q, \dot{q})\ddot{q}^j$$

where

$$b_i(q, \dot{q}) = Q_i(q, \dot{q}, \ddot{\bar{q}}) - a_{ij}(q, \dot{q})\ddot{\bar{q}}^j$$

Example 1. In its generality, Birkhoff systems define second order implicit differential equations, which find applications in, for example, flight dynamics, extremal problem from singularity theory and dynamical inequalities occurring in control theory (see LeMasurier (2001) and Davydov (1995) for the latter two).

As a concrete example, consider the model equations of a rigid airplane moving through a quiescent medium. Assuming a flat, stationary Earth, the configuration space is the Euclidean group $\mathrm{SE}(3) = \mathrm{SO}(3) \circledS \mathbb{R}^3$. In local coordinates given by a normal Cartesian coordinate system attached to the surface of the Earth $(x_E, y_E, z_E) \in \mathbb{R}^3$, the standard aeronautical Euler angles $(\phi, \theta, \psi) \in]0, 2\pi[\times]-\pi/2, \pi/2[\times]0, 2\pi[$ called bank, elevation and azimuth, respectively, $(u, v, w) \in \mathbb{R}^3$ the components of the velocity vector along the body axis, $(p, q, r) \in \mathbb{R}^3$ the roll, pitch and yaw rates, respectively, and $(\dot{u}, \dot{v}, \dot{w}, \dot{p}, \dot{q}, \dot{r}) \in \mathbb{R}^6$ with an obvious meaning we obtain: $\omega_U = Q_i \epsilon^i$, $i = 1, \ldots, 6$, with $U \subset J^2(\mathrm{SE}(3))$ the open set corresponding to the image of the domain of the above local coordinates, the Q_i given in the following way:

- Translational dynamics

$$\begin{aligned} Q_1 &= X - mg\sin\theta - m(\dot{u} + qw - rv) \\ Q_2 &= Y - mg\cos\theta\sin\phi - m(\dot{v} + ru - pw) \\ Q_3 &= Z - mg\cos\theta\cos\phi - m(\dot{w} + pv - qu) \end{aligned}$$

  where X, Y and Z are the components of the resultant of the aerodynamical and thrust forces along the body axis, m is the mass of the airplane assumed to be constant, g is the acceleration of the gravity;

- Attitude dynamics

$$\begin{aligned} Q_4 &= L - I_x\dot{p} + I_{zx}(\dot{r} + pq) + (I_y - I_z)qr \\ Q_5 &= M - I_y\dot{q} + I_{zx}(r^2 - p^2) + (I_z - I_x)rp \\ Q_6 &= N - I_z\dot{r} + I_{zx}(\dot{p} + qr) + (I_x - I_y)pq \end{aligned}$$

  where L, M and N are the roll, pitch and yaw aerodynamical and thrust moments, respectively, and I_{ij}, $i, j = x, y, z$ denote the components of the inertia tensor with respect to the body axis. Here ϵ^i, $i = 1, \ldots, 6$, are

$$\begin{aligned} \epsilon^1 &= \cos\theta\cos\psi \mathrm{d}x_E + \cos\theta\sin\psi \mathrm{d}y_E - \sin\theta \mathrm{d}z_E \\ \epsilon^2 &= (\sin\phi\sin\theta\cos\psi - \cos\phi\sin\psi)\mathrm{d}x_E + \\ &\quad (\sin\phi\sin\theta\sin\psi + \cos\phi\cos\psi)\mathrm{d}y_E + \sin\phi\cos\theta \mathrm{d}z_E \\ \epsilon^3 &= (\cos\phi\sin\theta\cos\psi + \sin\phi\sin\psi)\mathrm{d}x_E + \\ &\quad (\cos\phi\sin\theta\sin\psi - \sin\phi\cos\psi)\mathrm{d}y_E + \cos\phi\cos\theta \mathrm{d}z_E \\ \epsilon^4 &= \mathrm{d}\phi - \mathrm{d}\psi\sin\theta \\ \epsilon^5 &= \mathrm{d}\theta\cos\phi + \mathrm{d}\psi\cos\theta\sin\phi \\ \epsilon^6 &= \mathrm{d}\psi\cos\theta\cos\phi - \mathrm{d}\theta\sin\phi. \end{aligned}$$

Note that the coordinate system $(x_E, y_E, z_E, \phi, \theta, \psi, u, v, w, p, q, r, \dot{u}, \dot{v}, \dot{w}, \dot{p}, \dot{q}, \dot{r})$ for the manifold $J^2(\mathrm{SE}(3))$ is not a natural local coordinate system. In classical literature, "non natural" coordinates are often referred to as quasi-coordinates (see, for example, Pars (1965)). Yet, this prefix is a little misleading, since in general, any local coordinate for $J^2(\mathrm{M})$ is valid whether natural or quasi and in some cases, as in present example, the latter may facilitate the writing of and the computations with the equations of motion.

The aerodynamic forces and moments are very hard to estimate in general. However, in many applications, these forces and moments are assumed to be functions of position, velocity and also acceleration. For example, it is common to assume that the lift force is a function of the airplane speed, the angle of attack α, the pitch rate q, the altitude z_E, some control variables and also of $\dot{\alpha}$. Because $\tan \alpha_x = u/w$, with $\alpha_x = \alpha + \text{constant}$, we conclude that even for this elementary model, the system is an implicit second order system of differential equations.

Since these forces and moments depend on the acceleration and are non conservative, problems of this type do not fit into the Lagrangian formalism (it is not conservative in general). Also, since they have an implicit character, which may not be quasi-linear, they can not be included into the general approach put forward by Godbillon (see *Remark* 5). Thus, in a sense, the Birkhoff approach extends the Lagrangian approach to many mathematical-physical models (conservative or not) involving implicit second order differential equations.

We saw that, for each $z \in J^2(\mathrm{M})$, $\bar{\omega}(z)$ is a bilinear form of $\mathrm{T}_2^0(T_p\mathrm{M})$ where $p = \beta(z) = (\tau_\mathrm{M} \circ \tau_J)(z)$.

Definition 9. *A Birkhoffian ω of the configuration space* M *is said to be* normal *if $\bar{\omega}$ is constant along the β-fibres, that is, if $\bar{\omega}|_{\beta^{-1}(p)}$ is constant, for each $p \in \mathrm{M}$.*

Note that, when ω is normal, $\bar{\omega}$ defines a unique section $g_\omega \in \mathcal{T}_2^0(\mathrm{M})$ of $(\tau_2^0)_\mathrm{M} : \mathrm{T}_2^0(\mathrm{M}) \to \mathrm{M}$; in fact, if $p \in \mathrm{M}$: $g_\omega(p) := \bar{\omega}(z)$, for any $z \in \beta^{-1}(p)$.

If, in natural coordinates, ω is locally written as $\omega = Q_i(q, \dot{q}, \ddot{q})\mathrm{d}q^i$, the fact that ω is normal means that $\frac{\partial Q_i}{\partial \ddot{q}^j}(q, \dot{q}, \ddot{q}) = a_{ij}(q)$, that is, $\frac{\partial Q_i}{\partial \ddot{q}^j}$ depends only on $q = (q^1, \ldots, q^n)$.

Example 2. If the configuration space M has the structure of a pseudo-Riemannian manifold

$$(\mathrm{M}, (\,,\,)_g),$$

the function $T = \frac{1}{2}(\cdot,\cdot)_g : \mathrm{TM} \to \mathbb{R}$ is called the *kinetic energy*; given a smooth potential function $U : \mathrm{M} \to \mathbb{R}$, the Lagrangian function $\mathrm{L} = T - U \circ \tau_\mathrm{M}$ is called a *classical Lagrangian* and the corresponding Birkhoffian ω_L is normal with respect to the β-fibres.

Definition 10 (The Gibbs function). *Given a Birkhoffian ω of the configuration space* M *and a fixed smooth second order vector field S on* TM, *the* Gibbs function *for the data (ω, S):*

$$\mathfrak{G} : J^2(\mathrm{M}) \to \mathbb{R}$$

is the smooth function given by

$$\mathfrak{G}(z) = \frac{1}{2}\bar{\omega}(z)(x, x)$$

where $s = S(v)$, $v = \tau_J(z)$ *and* $x \in \mathrm{T}_p\mathrm{M}$, $p = \tau_{\mathrm{M}}(v)$ *being the unique vector such that* $\lambda_v x = z - s$.

Definition 11 (The Gibbs-Appell function). *Let* ω *and* S *be as above, fix a vertical section* $\xi : \mathrm{TM} \to \mathrm{V(M)}$ *and consider the data* (ω, S, ξ). *Let* $z \in J^2(\mathrm{M})$ *and define the bilinear form* $\mathbb{G}(z) : \mathrm{V}_v(\mathrm{M}) \times \mathrm{V}_v(\mathrm{M}) \to \mathbb{R}$ *by*

$$\mathbb{G}(z)(\eta_v, \zeta_v) = \bar{\omega}(z)(x, y)$$

where $v = \tau_J(z)$, $x, y \in \mathrm{T}_p\mathrm{M}$, $p = \tau_{\mathrm{M}}(v)$ *and* $\lambda_v x = \eta$, $\lambda_v y = \zeta$. *The* Gibbs-Appell function *associated to* (ω, S, ξ) *has the following definition, for all* $z \in J^2(\mathrm{M})$:

$$f(z) := \mathfrak{G}(z) + \mathbb{G}(z)(\xi_v, z - S(v)), \ v = \tau_J(z).$$

Definition 12 (The Gibbs-Appell principle). *A Birkhoff vector field* $Y = i \circ X$, $X : \mathrm{TM} \to J^2(\mathrm{M})$, *associated to a Birkhoffian* ω $(\mathrm{Im}\ X \subset D(\omega))$ *satisfies the* Gibbs-Appell principle *corresponding to the data* (ω, S, ξ) *if, for all* $v \in \mathrm{TM}$, $Y(v)$ *is a stationary point of the Gibbs-Appell function of* (ω, S, ξ) *restricted to* $\tau_J^{-1}(v)$.

3.1 The principle of the conservation of energy

Lemma 2. *Given two vectors* $u, v \in \mathrm{TM}$ *there exists a piecewise smooth path*[2] γ : $[0, 1] \to \mathrm{M}$ *such that* $\tau_{\mathrm{M}}(u) = \gamma(0)$, $\tau_{\mathrm{M}}(v) = \gamma(1)$ *and* $\frac{T\gamma}{dt}(0) = u$ *and* $\frac{T\gamma}{dt}(1) = v$.

Proof. Since M is connected, TM is also connected and it suffices to show the result for $\mathrm{T}\mathbb{R}^n$, with normal Cartesian coordinates $(q^i, \dot{q}^i)$. Denote by $(q_0 = q(u), \dot{q}_0 = \dot{q}(u))$ and $(q_1 = q(v), \dot{q}_1 = \dot{q}(v))$. Then, it is enough to consider for the path $\gamma_q^i(t) = q^i(\gamma(t))$ the Hermite interpolating polynomial:

$$\gamma_q^i(t) = (1 + 2t)(1 - t)^2 q_0^i + [1 + 2(1 - t)]t^2 q_1^i + t(1 - t)^2 \dot{q}_0^i - t^2(1 - t)\dot{q}_1^i$$

for all $t \in [0, 1]$. Now, for a general interval $[t_0, t_1]$ one considers the re-parameterization $t \mapsto \frac{t - t_0}{t_1 - t_0}$. □

Lemma 3. *Let* E *be a* C^1*-function on* TM. *Then,* E *is constant on* TM, *if and only if* $\mathrm{d}E(Y) = 0$, *for all second order vector fields* Y *on* TM.

[2] We say a continuous path $\gamma : [0, 1] \to \mathrm{M}$ is piecewise smooth if it is C^∞ for all but a finite number of points in the unit interval.

Proof. If E is constant, then $\mathsf{d}E = 0$. Conversely, assume $\mathsf{d}E(Y) = 0$, for all second order vector fields Y on TM. Given two vectors $u, v \in \mathrm{TM}$, let γ be a piecewise smooth path connecting u to v, that is, $\gamma(0) = \tau_{\mathrm{M}}(u)$, $\gamma(1) = \tau_{\mathrm{M}}(v)$, $\left.\frac{T\gamma}{dt}\right|_{t=0} = u$ and $\left.\frac{T\gamma}{dt}\right|_{t=1} = v$, which by Lemma 2 always exists. Let $t_0 \in [0,1]$ be a point of regularity of γ. Then,

$$\frac{d}{dt}\left[E\left(\frac{T\gamma}{dt}(t_0)\right)\right] = \mathsf{d}E\left(\frac{T}{dt}\frac{T\gamma}{dt}(t_0)\right) = \mathsf{d}E[Y\left(\frac{T\gamma}{dt}(t_0)\right)] = 0$$

where Y is some germ of second order vector field on TM such that $Y(\frac{T\gamma}{dt}(t_0)) = \frac{T}{dt}\frac{T\gamma}{dt}(t_0)$. So, by continuity E is constant along γ and so $E(u) = E(v)$. □

Let ω be a Birkhoffian of a smooth manifold M. Given a pair of vectors $u, v \in \mathrm{TM}$ and any piecewise smooth path joining these two vectors, $\gamma : [0,1] \to \mathrm{M}$, let $0 = t_1 < \cdots < t_{N_\gamma} = 1$ be a partition of the unitary interval where γ is smooth on its cells, the sum

$$W_\gamma(u,v) = \sum_{i=1}^{N_\gamma - 1} \int_{t_i}^{t_{i+1}} \frac{T}{dt}\frac{T\gamma}{dt}^{*}\omega$$

is called the *total work* from u to v over γ.

Definition 13. *A Birkhoffian ω of* M *is said to be* conservative or to satisfy the Principle of the conservation of energy, *if for any pair of vectors $u, v \in \mathrm{TM}$ the total work from u to v does not depend on the piecewise smooth path connecting them.*

Remark 2. Consider a conservative Birkhoffian ω. For a fixed $u \in \mathrm{TM}$ one defines a work W_u such that $W_u(v)$, for all $v \in \mathrm{TM}$, is the total work done between u and v. Given $u, v, w \in \mathrm{TM}$ one has $W_u(v) + W_v(w) = W_u(w)$, so $(W_u - W_v)$ is a constant.

Proposition 5. *A Birkhoffian ω of* M *is conservative if and only if there exists a smooth function $E_\omega : \mathrm{TM} \to \mathbb{R}$ (called energy of ω) such that $(X^*\omega)Y = \mathsf{d}E_\omega(Y)$ for all second order vector fields $Y = i \circ X$ on* TM.

Proof. We start by observing that the condition $(X^*\omega)Y = \mathsf{d}E_\omega(Y)$ for all second order vector fields $Y = i \circ X$ on TM implies that on each local natural coordinate system, the fundamental identity (see Birkhoff (1927), p.16, equation (4)):

$$Q_i(q,\dot q,\ddot q)\dot q^i = \frac{\partial E_\omega}{\partial q^i}\dot q^i + \frac{\partial E_\omega}{\partial \dot q^i}\ddot q^i,$$

holds. Then ω is obviously conservative.

Conversely, suppose ω is conservative and fix a vector $u \in \mathrm{TM}$. Given $v \in \mathrm{TM}$, define $E^u_\omega(v) = W_u(v)$. Clearly, this yields a well defined a function $E^u_\omega : \mathrm{TM} \to \mathbb{R}$.

1. $E^u_\omega : \mathrm{TM} \to \mathbb{R}$ is smooth.

   It is enough to prove it locally. So, let (U, q) be a local chart and $v \in \mathrm{T}U$. By definition we have

$$E^u_\omega(w) = E^u_\omega(v) + \int_0^1 Q_i(\gamma_q(t), \dot{\gamma}_q(t), \ddot{\gamma}_q(t))\dot{\gamma}^i_q(t)\mathrm{d}t,$$

   where γ_q is the path in the proof of Lemma 2 connecting v to any other point w in the local chart $(\mathrm{T}U, (q, \dot{q}))$. The smoothness follows from the smoothness of γ_q with respect to the coordinates of v.

2. $(X^*\omega)Y = \mathrm{d}E^u_\omega(Y)$ for all second order vector fields $Y = i \circ X$ on TM.

   Again we prove the result locally. Thus, let $Y = i \circ X$ be a second order vector field. Take a base integral curve $\gamma : (-\epsilon, \epsilon) \to \mathrm{M}$ of Y at a point $v \in \mathrm{TM}$, that is, $\gamma(0) = \tau_{\mathrm{M}}(v)$, $\frac{T\gamma}{dt}\big|_{t=0} = v$ and $\frac{T}{dt}\frac{T\gamma}{dt}\big|_{t=0} = Y(v)$ (note that $\tau_{\mathrm{TM}}Y(v) = v$). Denote by $q(v) = q_0$, $\dot{q}(v) = \dot{q}_0$ and $\ddot{q}(X(v)) = \ddot{q}_0$. Then,

$$E^u_{\omega,q}(\gamma_q(t), \dot{\gamma}_q(t)) = E^u_\omega(v) + \int_0^t Q_i(\gamma_q(\bar{t}), \dot{\gamma}_q(\bar{t}), \ddot{\gamma}_q(\bar{t}))\dot{\gamma}^i_q(\bar{t})\mathrm{d}\bar{t},$$

   for all $t \in [-\epsilon, \epsilon]$, where $\gamma_q = q(\gamma)$ and $E^u_{\omega,q}(\gamma_q(t), \dot{\gamma}_q(t)) = E^u_\omega\left(\frac{T\gamma}{dt}(t)\right)$. Taking the derivative of the previous equation with respect to t yields

$$\frac{\partial E^u_{\omega,q}}{\partial q^i}\dot{\gamma}^i_q(t) + \frac{\partial E^u_{\omega,q}}{\partial \dot{q}^i}\ddot{\gamma}^i_q(t) = Q_i(\gamma_q(t), \dot{\gamma}_q(t), \ddot{\gamma}_q(t))\dot{\gamma}^i_q(t),$$

   $t \in [-\epsilon, \epsilon]$. Finally, at $t = 0$ we get

$$\frac{\partial E^u_{\omega,q}}{\partial q^i}\dot{q}^i_0 + \frac{\partial E^u_{\omega,q}}{\partial \dot{q}^i}\ddot{q}^i_0 = Q_i(q_0, \dot{q}_0, \ddot{q}_0)\dot{q}^i_0,$$

   and the proof is complete.

□

Remark 3. If ω satisfies the principle of the conservation of energy and $Y = i \circ X$ is a smooth second order vector field, such that $X^*\omega = 0$, then

$$\mathrm{d}E_\omega(Y) = (X^*\omega)Y = 0.$$

This means that E_ω is constant along trajectories of Y on TM (see Definition 4).

Let us recall the notion of fibre derivative (see, for example, Abraham & Marsden (1978)). Let f be a smooth function on TM; the fibre derivative of f, $\mathbb{F}f : \mathrm{TM} \to \mathrm{T}^*\mathrm{M}$, is the smooth morphism defined as follows

$$\mathbb{F}f(v)w = \frac{d}{dt}f(v + tw)\bigg|_{t=0}$$

for all $v, w \in \mathrm{T}_p\mathrm{M}$, with $p = \tau_{\mathrm{M}}(v) = \tau_{\mathrm{M}}(w)$. With the fibre derivative we define the vertical differentiation $\mathsf{d}_v f$ (see Godbillon (1969)) of a function f on TM, as the 1-form on TM given by

$$\mathsf{d}_v f(u)\tilde{w} = \mathbb{F}f(u)\mathrm{T}\tau_{\mathrm{M}}\tilde{w}$$

for all $u \in \mathrm{TM}$ and for all $\tilde{w} \in \tau_{\mathrm{TM}}^{-1}(u)$. In local natural coordinates,

$$\mathsf{d}_v f = \frac{\partial f}{\partial \dot{q}^i}\,\mathrm{d}q^i.$$

Proposition 6. *Given a smooth (Lagrangian) function* $\mathrm{L} : \mathrm{TM} \to \mathbb{R}$, *there exists a unique Birkhoffian* ω_L *which satisfies*

$$X^*\omega_L = \mathsf{i}_Y \mathsf{d}(\mathsf{d}_v \mathrm{L}) + \mathsf{d}\left(Z(\mathrm{L}) - \mathrm{L}\right) \tag{15}$$

for all second order vector fields $Y = i \circ X$ *on* TM, Z *being the Liouville vector field.*

Moreover, in local natural coordinates we have

$$\omega_L = \left(\frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{q}^i} - \frac{\partial \mathrm{L}}{\partial q^i}\right)\mathrm{d}q^i.$$

Proof. In a local natural coordinate system and up to some obvious identifications we can write

$$\begin{aligned}
Y &= i \circ X = \dot{q}^i \frac{\partial}{\partial q^i} + \ddot{q}^i \frac{\partial}{\partial \dot{q}^i}\,, \\
\mathsf{d}\mathsf{d}_v L &= \frac{\partial^2 L}{\partial q^i \partial \dot{q}^j}\mathrm{d}q^i \wedge \mathrm{d}q^j + \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\mathrm{d}\dot{q}^i \wedge \mathrm{d}q^j, \\
\mathsf{i}_X \mathsf{d}\mathsf{d}_v \mathrm{L} &= \frac{\partial^2 L}{\partial q^i \partial \dot{q}^j}\dot{q}^i \mathrm{d}q^j - \frac{\partial^2 L}{\partial q^i \partial \dot{q}^j}\dot{q}^j \mathrm{d}q^i + \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\ddot{q}^i \mathrm{d}q^j - \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\dot{q}^j \mathrm{d}\dot{q}^i, \\
\mathsf{d}L &= \frac{\partial L}{\partial q^i}\,\mathrm{d}q^i + \frac{\partial L}{\partial \dot{q}^i}\,\mathrm{d}\dot{q}^i, \\
Z(L) &= \frac{\partial L}{\partial \dot{q}^i}\,\dot{q}^i, \\
\mathsf{d}(Z(L)) &= \frac{\partial^2 L}{\partial q^i \partial \dot{q}^j}\dot{q}^j \mathrm{d}q^i + \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\dot{q}^j \mathrm{d}\dot{q}^i + \frac{\partial L}{\partial \dot{q}^i}\,\mathrm{d}\dot{q}^i
\end{aligned}$$

As usual, we first prove uniqueness. Hence, suppose a Birkhoffian exists, which satisfies (15) for any second order vector field $Y = i \circ X$. Then, using the above local expressions we get,

$$Q_i(q, \dot{q}, \ddot{q})\mathrm{d}q^i = \left(\frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{q}^i} - \frac{\partial \mathrm{L}}{\partial q^i}\right)\mathrm{d}q^i$$

for all $(q, \dot{q}, \ddot{q}) \in J^2(U)$, which prove uniqueness.

As for the existence, it is enough to define locally the Birkhoffian by the previous expression and show that this definition extends to a (global) 1-form ω_L; but this is a classical result. □

Remark 4. The previous proposition gives an intrinsic characterization of the so called Euler-Lagrange operator. Actually, the latter is $\hat{\omega}_L$. Therefore, in a sense, $\hat{\omega}$ of a Birkhoffian ω is a generalization of the Euler-Lagrange operator for general implicit second-order mechanical systems.

Proposition 7. *The Birkhoffian $\omega = \omega_L$ associated to a smooth Lagrangian function $\mathrm{L} : \mathrm{TM} \to \mathbb{R}$ is affine and satisfies both the principle of reciprocity and conservation of energy. The energy of ω_L is $E_L = Z(\mathrm{L}) - \mathrm{L}$. Moreover, ω_L is regular if and only if L is a regular Lagrangian, that is, if $\mathbb{F}\mathrm{L}$ is a submersion on $\mathrm{T}^*\mathrm{M}$.*

Proof. To check that the Birkhoffian ω_L associated to a smooth Lagrangian function $\mathrm{L} : \mathrm{TM} \to \mathbb{R}$ is affine with respect to the acceleration, it is enough to write it in local natural coordinates,

$$\omega_L = \left(\frac{\partial^2 \mathrm{L}}{\partial \dot{q}^i \partial \dot{q}^j} \ddot{q}^j + \frac{\partial^2 \mathrm{L}}{\partial \dot{q}^i \partial q^j} \dot{q}^j - \frac{\partial \mathrm{L}}{\partial q^i} \right) \mathrm{d}q^i.$$

It is also easy to see that the Birkhoffian ω_L satisfies the principle of reciprocity since by smoothness

$$\frac{\partial Q_i}{\partial \ddot{q}^j} = \frac{\partial^2 \mathrm{L}}{\partial \dot{q}^i \partial \dot{q}^j} = \frac{\partial Q_j}{\partial \ddot{q}^i}$$

From Proposition 6, ω_L is the unique Birkhoffian such that

$$X^* \omega_L = \mathrm{i}_Y \mathsf{dd}_v \mathrm{L} + \mathsf{d} E_L$$

for all second order vector fields $Y = i \circ X$ on TM. Then, since $(\mathrm{i}_Y \mathsf{dd}_v \mathrm{L})(Y) = \mathsf{dd}_v \mathrm{L}(Y, Y) = 0$, one concludes that $(X^* \omega_L) Y = \mathsf{d} E_L(Y)$ for all second order vector fields $Y = i \circ X$ on TM.

Finally, ω_L is regular if, and only if, L is a regular Lagrangian, that is, if we have

$$\det \left[\frac{\partial^2 \mathrm{L}}{\partial \dot{q}^i \partial \dot{q}^j} \right] \neq 0,$$

everywhere. □

We recall that a Pfaffian form π on TM is said to be *semi-basic* if, for each $v \in \mathrm{TM}$, $\pi_v \in [\mathrm{V}_v(\mathrm{M})]^\circ$, that is $\pi_v(\xi) = 0$ for all $\xi \in \mathrm{V}_v(\mathrm{M})$. Locally, $\pi = \pi_i(q, \dot{q}) \mathrm{d}q^i$. A *force field* $\mathscr{F}$ is a smooth fiber preserving mapping $\mathscr{F} : \mathrm{TM} \to \mathrm{T}^*\mathrm{M}$.

Proposition 8. *There is a one-to-one correspondence between semi-basic forms and force fields.*

Proof. Let π be a semi-basic form and define the smooth fibre preserving map $\mathscr{F}_\pi : \mathrm{TM} \to \mathrm{T}^*\mathrm{M}$ by $\mathscr{F}_\pi(v) = \pi_v(\xi_v)$ where $v \in \mathrm{T}_{\tau_\mathrm{M}(v)}\mathrm{M}$ and $\xi_v \in \mathrm{T}_v(\mathrm{TM})$ is any vector that projects onto v under $\mathrm{T}\tau_\mathrm{M}(v)$; $\mathscr{F}_\pi$ is well defined because is $\bar{\xi}_v \in \mathrm{T}_v(\mathrm{TM})$ is another vector field that projects onto v under $\mathrm{T}\tau_\mathrm{M}(v)$ we have $(\mathrm{T}\tau_\mathrm{M}(v))\,(\xi_v - \bar{\xi}_v) = 0$, so $\xi_v - \bar{\xi}_v \in \mathrm{V}_v(\mathrm{M})$ and then $\pi_v(\xi_v - \bar{\xi}_v) = 0$, that is $\pi_v(\xi_v) = \pi_v(\bar{\xi}_v)$. Conversely, given a smooth fibre preserving map $\mathscr{F} : \mathrm{TM} \to \mathrm{T}^*\mathrm{M}$, that is a field of forces, then there exists a unique semi-basic Pfaffian form $\pi_\mathscr{F}$ such that $\mathscr{F}_{\pi_\mathscr{F}} = \pi_\mathscr{F}$. Indeed $\pi_\mathscr{F} = \mathscr{F}^*\theta_0$ where θ_0 is the canonical *Liouville 1-form* of $\mathrm{T}^*\mathrm{M}$, and $\mathscr{F}^*\theta_0(\xi_v) = \theta_0(\mathscr{F}_*\xi_p) = 0$ for all $\xi_v \in \mathrm{V}_v(\mathrm{M})$ since $\mathscr{F}_*\xi_v$ is tangent to the fibre $\mathrm{T}^*_{\tau_\mathrm{M}(v)}\mathrm{M}$. □

Given a semi-basic Pfaffian form π we denote by $\bar{\pi}$ the associated force field.

Remark 5. By Proposition 6 we conclude that Birkhoff systems generalize Lagrangian systems. They also extend the notion of some mechanical systems (non-conservative, in general) as defined by Godbillon in Godbillon (1969). The latter defines a mechanical system by a triple (M, T, π) where M is a smooth manifold, called the configuration space, T is a smooth function on TM, called the kinetic energy and π is a semi-basic Pfaffian form, called the force field. With these data Godbillion associates the notion of a dynamical system to a second order vector field $Y = i \circ X$ which satisfies

$$\mathsf{i}_Y \mathsf{d}(\mathsf{d}_v T) = \mathsf{d}(T - Z \cdot T) + \pi.$$

Indeed, the Birkhoffian ω associated to the triple (M, T, π) is defined by

$$X^*\omega = \mathsf{i}_Y \mathsf{d}(\mathsf{d}_v T) - \mathsf{d}(T - Z \cdot T) - \pi.$$

for all second order vector fields $Y = i \circ X$.

3.2 The inverse problem of the Lagrangian Mechanics

Let M be a smooth manifold and ω be a Birkhoffian of the configuration space M. In the inverse problem we want to know under which conditions $\omega = \omega_L$ for some Lagrangian function L. Let us start with a characterization of classical Lagrangian systems.

3.2.1 External characterization of classical Lagrangian systems

Theorem 2. *Let ω be a Birkhoffian of the configuration space* M. *The following statements are equivalent:*

1. $\omega = \omega_L$ *where* $\mathrm{L} = \frac{1}{2}g - U \circ \tau_\mathrm{M}$ *is a classical Lagrangian.*
2. ω *is such that:*
   - $\bar{\omega}(z)$ *is non degenerate,* $\forall z \in J^2(\mathrm{M})$*;*

- *ω satisfies the principle of reciprocity;*
- *ω is affine;*
- *ω is conservative;*
- *ω is normal (so it defines a pseudo-Riemannian metric g_ω on M);*
- *the differential system $D(\omega)$ is the image of a Birkhoff vector field Y that satisfies the Gibbs-Appell principle for (ω, S, ξ) where S is the geodesic spray of g_ω and $\xi(v) = \lambda_v \operatorname{grad} U$, with $\tau_{\mathrm{M}}^* U = E_\omega - \frac{1}{2} g_\omega$.*

Remark 6. Note that, as an easy computation shows, when ω is normal and conservative with energy E, there exists a unique function $U : \mathrm{M} \to \mathbb{R}$ which satisfies $\tau_{\mathrm{M}}^* U = E - \frac{1}{2} g_\omega$.

Proof. 1) $\Rightarrow$ 2) is clear. So, let us assume that 2) holds. To prove that 2) $\Rightarrow$ 1) it is enough to verify this assertion locally. Now, if ω is non-degenerate, satisfies the principle of reciprocity, is affine, conserves energy and is normal, we get the following equation in local natural coordinates:

$$(g_{ij}(q)\ddot{q}^j + b_i(q,\dot{q}))\dot{q}^i = \frac{\partial E}{\partial q^i}\,\dot{q}^i + \frac{\partial E}{\partial \dot{q}^i}\,\ddot{q}^i,$$

where $g_{ij}(q)$ is the metric matrix of g_ω. Then

$$\frac{\partial E}{\partial \dot{q}^i} = g_{ij}(q)\dot{q}^j \quad \therefore \quad E(q,\dot{q}) = \frac{1}{2} g_{ij}(q)\dot{q}^i\dot{q}^j + U(q)$$

for some $U \in \mathfrak{F}(\mathrm{M})$. Also, from

$$b_i(q,\dot{q})\dot{q}^i = \frac{\partial E}{\partial q^i}\,\dot{q}^i$$

and $\frac{\partial g_{ij}}{\partial q^k} = \Gamma^l_{ik} g_{jl} + \Gamma^l_{jk} g_{il}$ where Γ^i_{jk} are the Christoffel symbols associated with the Levi-Civita connection of g_ω, for the coordinate system $(q^1, \ldots, q^n)$ of M, we obtain, after a simple computation,

$$b_i(q,\dot{q}) = g_{ij}(q)\Gamma^j_{kl}(q)\dot{q}^k\dot{q}^l + \frac{\partial U}{\partial q^i}(q) + R_i(q,\dot{q})$$

with $R_i(q,\dot{q})\dot{q}^i = 0$. So,

$$Q_i(q,\dot{q},\ddot{q}) = g_{ij}(q)(\ddot{q}^j + \Gamma^j_{kl}(q))\dot{q}^k\dot{q}^l) + \frac{\partial U}{\partial q^i} + R_i(q,\dot{q}).$$

To complete the proof let us write the Gibbs-Appell function for (ω, S, ξ) in the local natural coordinates $(q, \dot{q}, \ddot{q})$:

$$f(q,\dot{q},\ddot{q}) = \frac{1}{2} g_{ij}(q)(\ddot{q}^j + \Gamma^j_{kl}(q)\dot{q}^k\dot{q}^l)(\ddot{q}^i + \Gamma^i_{kl}(q)\dot{q}^k\dot{q}^l) + \frac{\partial U}{\partial q^i}(\ddot{q}^i + \Gamma^i_{kl}(q)\dot{q}^k\dot{q}^l)$$

Then, differentiating with respect to $\ddot{q}$ and equating to zero yields

$$\ddot{q}^j + \Gamma^j_{kl}(q)\dot{q}^j\dot{q}^k + g^{ij}(q)\frac{\partial U}{\partial q^i} = 0,$$

for the Birkhoff vector field $Y = i \circ X$, that is, $X(q, \dot{q}) = (q, \dot{q}, \ddot{q})$ with $\ddot{q}$ given by the previous equation. Finally, from $X^*\omega = 0$ we obtain $R_i(q, \dot{q}) = 0$ for an arbitrary $(q, \dot{q})$, as required. □

Remark 7. In the last Theorem, the statement 1. implies that the Birkhoffian $\omega = \omega_L$ is regular. The statement 2. shows also that the stationary point $Y(v)$ is a minimum if and only if the pseudo-Riemannian metric g_ω is a Riemannian metric.

In Birkhoff (1927) a different external characterization of classical Lagrangian systems has been presented. Specifically, apart from regularity, conservation of energy, reciprocity, the notions of affine and normal Birkhoffian it is used the condition that the local functions $b_i(q, \dot{q})$, $i = 1, \ldots, n$ are quadratic with respect to the velocities, the principle of reversibility and also the condition: "If by a particular choice of admissible coordinates, the kinetic energy T is made stationary in the q^j at a certain point q^j_0, then this implies that the Q_i are independent of the velocities." These conditions correspond to: (i) the Birkhoff vector field Y is a spray, that is, for all $s \in \mathbb{R}$ and $v \in \mathrm{TM}$ we have $Y(sv) = (s_{\mathrm{TM}})_* sY(v)$, where $s_{\mathrm{TM}} : \mathrm{TM} \to \mathrm{TM}$, given by: $v \mapsto sv$ is the vector bundle morphism given by scalar multiplication by s and (ii) $Y - S$, where S is the geodesic spray associated to the metric g_ω, satisfies $Y_v - S_v = \lambda_v(\mathrm{grad}\, U(p))$, with $p = \tau_{\mathrm{M}}(v)$ for all $v \in \mathrm{TM}$ and for some function U on M. By a proof analogous to the previous theorem, one can state:

Theorem 3 (Birkhoff). *Let ω be a Birkhoffian of the configuration space* M. *The following statements are equivalent:*

1. $\omega = \omega_L$ *where* $\mathrm{L} = \frac{1}{2}g - U \circ \tau_{\mathrm{M}}$ *is a classical Lagrangian.*
2. ω *is such that:*
    - $\bar{\omega}(z)$ *is non degenerate,* $\forall z \in J^2(\mathrm{M})$;
    - ω *satisfies the principle of reciprocity;*
    - ω *is affine;*
    - ω *conserves an energy* E;
    - ω *is normal (so it defines a pseudo-Riemannian metric* g_ω *on* M*);*
    - *the Birkhoff vector-field* Y *is a spray such that* $Y - S$, *where* S *is the geodesic spray associated to the metric* g_ω, *satisfies* $Y_v - S_v = \lambda_v(\mathrm{grad}\, U(p))$, *with* $p = \tau_{\mathrm{M}}(v)$ *for all* $v \in \mathrm{TM}$ *and* $\tau^*_{\mathrm{M}}U = E - \frac{1}{2}g_\omega$ *(see Remark 6).*

3.2.2 External characterization of general Lagrangian systems

Denote by $C^k(S^1, M)$, $k \in \mathbb{N}$ the Banach manifold of all C^k loops of M, that is, the set of all maps $\gamma : S^1 \to M$ of class C^k, where $S^1 = \mathbb{R}/\mathbb{Z}$. In the sequel, unless otherwise stated we consider $k \geq 2$. Define the Pfaffian form $\Omega : C^k(S^1, M) \to T^*C^k(S^1, M)$ as follows: given $\gamma \in C^k(S^1, M)$ and a Birkhoffian ω:

$$\Omega(\gamma)\eta = -\int_{S^1} \hat{\omega}(\frac{T}{dt}\frac{T\gamma}{dt})\eta \tag{16}$$

for all $\eta \in T_\gamma C^k(S^1, M) \cong \gamma^* TM$. Given a Lagrangian L let us denote by Ω_L the Pfaffian form on $C^k(S^1, M)$ corresponding to ω_L and by $\mathscr{L}$ the function on $C^k(S^1, M)$ defined by

$$\mathscr{L}(\gamma) = \int_{S^1} L(\frac{T\gamma}{dt}).$$

A classical result give us:

Proposition 9. *Given a Lagrangian* L *we have*

$$\Omega_L = \mathsf{d}_{C^k(S^1,M)}\mathscr{L}.$$

From the previous proposition, we conclude that a necessary condition for a given Birkhoffian ω to be equal to ω_L, corresponding to some Lagrangian function L, is that Ω is closed, that is, $\mathsf{d}_{C^k(S^1,M)}\Omega = 0$.

Definition 14. *Let ω be a Birkhoffian on* M. *We say that ω* satisfies the Helmholtz conditions, *if for any natural coordinate system $(q, \dot{q}, \ddot{q})$ of $J^2(M)$ the following equalities are verified:*

$$\begin{aligned} \frac{\partial Q_i}{\partial \ddot{q}^j} &= \frac{\partial Q_j}{\partial \ddot{q}^i}, & (17)\\ \frac{\partial Q_i}{\partial \dot{q}^j} + \frac{\partial Q_j}{\partial \dot{q}^i} &= \frac{d}{dt}\left(\frac{\partial Q_i}{\partial \ddot{q}^j} + \frac{\partial Q_j}{\partial \ddot{q}^i}\right), & (18)\\ \frac{\partial Q_i}{\partial q^j} - \frac{\partial Q_j}{\partial q^i} &= \frac{1}{2}\frac{d}{dt}\left(\frac{\partial Q_i}{\partial \dot{q}^j} - \frac{\partial Q_j}{\partial \dot{q}^i}\right). & (19) \end{aligned}$$

Lemma 4. *Let ω be a Birkhoffian on* M *which satisfies the Helmholtz conditions. Let $(q, \dot{q}, \ddot{q})$ be a local natural coordinate system for $J^2(M)$ associated to a local chart (U, q) of* M. *Then, for all $z \in J^2(U)$ and for all pairs $u, v \in TU$, the number*

$$\begin{aligned} \delta_1(q,\dot{q},\ddot{q})u^j v^i &= \left(\frac{\partial Q_i}{\partial q^j}(q_0,\dot{q}_0,\ddot{q}_0) - \frac{\partial Q_j}{\partial q^i}(q_0,\dot{q}_0,\ddot{q}_0)\right)u^j v^i - \\ &\quad \frac{1}{2}\frac{d}{dt}\left(\frac{\partial Q_i}{\partial \dot{q}^j}(q_0,\dot{q}_0,\ddot{q}_0) - \frac{\partial Q_j}{\partial \dot{q}^i}(q_0,\dot{q}_0,\ddot{q}_0)\right)u^j v^i \in \mathbb{R} \quad (20) \end{aligned}$$

where $q(z) = q_0$, $\dot{q}(z) = \dot{q}_0$, $\ddot{q}(z) = \ddot{q}_0$, $u = u^i \frac{\partial}{\partial q^i}(q_0)$ *and* $v = v^i \frac{\partial}{\partial q^i}(q_0)$, *does not depend on the coordinate system* (q^i).

Also for all $z \in J^2(U)$ *and any pair* $u, v \in \mathrm{T}U$ *the number*

$$\begin{aligned}\delta_2(q,\dot{q},\ddot{q})u^j v^i &= \left(\frac{\partial Q_i}{\partial \dot{q}^j}(q_0,\dot{q}_0,\ddot{q}_0) + \frac{\partial Q_j}{\partial \dot{q}^i}(q_0,\dot{q}_0,\ddot{q}_0)\right) u^j v^i - \\ &\quad \frac{d}{dt}\left(\frac{\partial Q_i}{\partial \ddot{q}^j}(q_0,\dot{q}_0,\ddot{q}_0) + \frac{\partial Q_j}{\partial \ddot{q}^i}(q_0,\dot{q}_0,\ddot{q}_0)\right) u^j v^i \in \mathbb{R} \quad (21)\end{aligned}$$

where $q(z) = q_0$, $\dot{q}(z) = \dot{q}_0$, $\ddot{q}(z) = \ddot{q}_0$, $u = u^i \frac{\partial}{\partial q^i}(q_0)$ *and* $v = v^i \frac{\partial}{\partial q^i}(q_0)$, *does not depend on the coordinate system* (q^i).

Proof. See Appendix A for a proof. □

Remark 8. The previous Lemma shows that equations (20) and (21) define morphisms $\delta_i : J^2(\mathrm{M}) \to \mathrm{T}^0_2(\mathrm{M})$, $i = 1, 2$. In other words, the Helmholtz conditions are globally defined and consequently if the Helmholtz conditions hold for a particular open covering (U_i) of coordinate neighborhoods of M, then it holds for any other coordinate covering.

Theorem 4. *Let* ω *be a Birkhoffian on* M*. Then* Ω *defined in equation* (16) *is closed if and only if* ω *satisfies the Helmholtz conditions.*

Remark 9. Note that, in view of the previous Theorem, we conclude that the geometric content of the Helmholtz conditions corresponds to the closedness of Ω.

Proof. Let us start with an explicit determination of the manifold structure of $\mathrm{C^k}(S^1, \mathrm{M})$. For this, it is convenient to embed the manifold M into $\mathbb{R}^N$ for some $N \in \mathbb{N}$, by the Whitney embedding Theorem. For simplicity of notation we may assume that $\mathrm{M} \subset \mathbb{R}^N$ and so $\mathrm{TM} \subset \mathrm{M} \times \mathbb{R}^N$. Let $\tilde{M}$ denote the normal bundle over M, that is, the union $\tilde{M} = \cup_{p \in \mathrm{M}} \mathrm{T}^\perp_p \mathrm{M} \subset \mathrm{M} \times \mathbb{R}^N$ where $\mathrm{T}^\perp_p \mathrm{M}$ is the subset of $\mathbb{R}^N$ orthogonal to $\mathrm{T}_p\mathrm{M}$ with respect to the usual inner product of $\mathbb{R}^N$. So, we have a direct sum $\mathrm{T}_p\mathrm{M} \oplus \mathrm{T}^\perp_p\mathrm{M} = \mathbb{R}^N$ for each $p \in \mathrm{M}$ and rank $\tilde{M} = N - n$.

Now, take a tubular neighborhood (Π, f) of M in $\mathbb{R}^N$ (see, for example, Cannas da Silva (2001), Oliva (2002)), that means, a smooth diffeomorphism $f : \Pi \to \Delta$ from an open neighborhood Π of the zero section in $\tilde{\mathrm{M}}$ onto an open set $\Delta \subset \mathbb{R}^N$, $\mathrm{M} \subset \Delta$, such that $f(\mathbb{O}_p) = p$ for any zero vector $\mathbb{O}_p \in \tilde{\mathrm{M}}$, $p \in \mathrm{M}$. If $\mathrm{pr}_1 : \mathrm{M} \times \mathbb{R}^N \to \mathrm{M}$ is the first projection, the map $\mathsf{p} = (\mathrm{pr}_1|_{\tilde{M}}) \circ f^{-1} : \Delta \to \mathrm{M}$ is a projection ($\mathsf{p}^2 = \mathsf{p}$); the pair (Δ, p) also represents the tubular neighborhood of M in $\mathbb{R}^N$. The set Δ is called the tube in $\mathbb{R}^N$ and Π is said to be a tube in $\tilde{\mathrm{M}}$; they play the same role and can be identified by the diffeomorphism f.

Given $\gamma_0 \in \mathrm{C^k}(S^1, \Delta)$, consider $U_0 \subset \mathrm{C}^0(S^1, \Delta)$ as the subset of all loops γ such that $||\gamma(t) - \gamma_0(t)||_{C^0(S^1, \mathbb{R}^N)} < r$, with $r > 0$ small enough such that all loops in this neighborhood are contained in the tube Δ. By the continuity of the injection of $\mathrm{C^k}(S^1, \mathbb{R}^N)$ into $\mathrm{C}^0(S^1, \mathbb{R}^N)$ it follows that there exists an open set

$U \subset \mathrm{C}^{\mathrm{k}}(S^1, \Delta)$ such that $\gamma_0 \in U$ and any loop in this neighborhood is contained in the tube Δ. Define,

$$\begin{aligned} \Phi_U &: U \to \mathrm{C}^{\mathrm{k}}(S^1, \mathbb{R}^N) \\ &: \gamma \mapsto \gamma - \gamma_0 \end{aligned}$$

Clearly, the image of Φ_U into the Banach space $\mathrm{C}^{\mathrm{k}}(S^1, \mathbb{R}^N)$ is an open subset and (U, Φ_U) provides an atlas of charts of the manifold structure on $\mathrm{C}^{\mathrm{k}}(S^1, \Delta)$. Let us compute, $\mathsf{d}_{\mathrm{C}^{\mathrm{k}}(S^1,\mathrm{M})}\mathscr{L}(\gamma_0)\eta$ for $\eta \in \mathrm{T}_{\gamma_0}\mathrm{C}^{\mathrm{k}}(S^1, \mathrm{M})$.

Let us extend the Birkhoffian ω on M to a Birkhoffian ω^Δ on Δ: $\omega^\Delta = J^2\mathsf{p}^*\omega$. Given $\gamma \in \mathrm{C}^{\mathrm{k}}(S^1, \Delta)$ we have

$$\Omega^\Delta(\gamma)\eta = -\int_{S^1} Q_i^\Delta(x, \dot{x}, \ddot{x})\eta^i$$

for all $\eta \in \mathrm{T}_\gamma\mathrm{C}^{\mathrm{k}}(S^1, \Delta)$, where (x^i) are the normal Cartesian coordinates of $\mathbb{R}^N$. We recall (see Lang (1995)) that from

$$\Omega^\Delta : U \to \mathrm{C}^{\mathrm{k}}(S^1, \mathbb{R}^N)^*,$$

we have, for each $\gamma \in U$,

$$(\Omega^\Delta)'(\gamma) : \mathrm{C}^{\mathrm{k}}(S^1, \mathbb{R}^N) \to \mathrm{C}^{\mathrm{k}}(S^1, \mathbb{R}^N)^*,$$

which is a continuous linear map. Indeed, for all $\gamma \in U$ we have

$$\langle (\Omega^\Delta)'(\gamma)\xi, \eta\rangle = -\int_{S^1} \left(\frac{\partial Q_i^\Delta}{\partial x^j}\,\xi^j + \frac{\partial Q_i^\Delta}{\partial \dot{x}^j}\,\dot{\xi}^j + \frac{\partial Q_i^\Delta}{\partial \ddot{x}^j}\,\ddot{\xi}^j \right)\eta^i.$$

$\eta \in \mathrm{T}_\gamma\mathrm{C}^{\mathrm{k}}(S^1, \Delta)$.

Taking into account that Ω^Δ extends Ω (recall that, by its definition, ω^Δ extends ω), we conclude that the condition for the closeness of Ω is that for all $\gamma \in \mathrm{C}^{\mathrm{k}}(S^1, \mathrm{M})$:

$$\langle \mathsf{d}_{\mathrm{C}^{\mathrm{k}}(S^1,\Delta)}\Omega(\gamma), \xi \times \eta\rangle = \langle (\Omega^\Delta)'(\gamma)\xi, \eta\rangle - \langle (\Omega^\Delta)'(\gamma)\eta, \xi\rangle = 0,$$

for all $\xi, \eta \in \mathrm{T}_\gamma\mathrm{C}^{\mathrm{k}}(S^1, \mathrm{M})$. The manifold M has been identified with a submanifold of $\mathbb{R}^N$. So, we can use a chart of $\mathbb{R}^N$ adapted to the submanifold M. Then, since $\gamma(S^1)$ is compact, we can cover it by a finite number of adapted coordinate systems (V_l), $l = 1, \ldots, N_\gamma$ such that the previous condition are

$$\begin{aligned} \langle \mathsf{d}_{\mathrm{C}^{\mathrm{k}}(S^1,\Delta)}\Omega(\gamma), \xi \times \eta\rangle &= -\sum_{l=1}^{N_\gamma - 1} \left(\int_{t_l}^{t_{l+1}} \left(\frac{\partial Q_i^l}{\partial q_l^j}\,\xi_l^j + \frac{\partial Q_i^l}{\partial \dot{q}_l^j}\,\dot{\xi}_l^j + \frac{\partial Q_i^l}{\partial \ddot{q}_l^j}\,\ddot{\xi}_l^j \right)\eta_l^i - \right. \\ &\left. \left(\frac{\partial Q_i^l}{\partial q_l^j}\,\eta_l^j + \frac{\partial Q_i^l}{\partial \dot{q}_l^j}\,\dot{\eta}_l^j + \frac{\partial Q_i^l}{\partial \ddot{q}_l^j}\,\ddot{\eta}_l^j \right)\xi_l^i \right), \end{aligned} \tag{22}$$

where $0 = t_1 < \cdots < t_{N_\gamma} = 1$ is a partition of the unit interval such that $\gamma([t_l, t_{l+1}]) \subset V_l$ for each $l = 1, \ldots, N_\gamma$, (q_l^i) is the local system of coordinates on V_l and Q_i^l, ξ_l^i and η_l^i are the local representative on V_l of ω, ξ and η, respectively.

Let us assume that Ω is closed. Then, given a local coordinate system (V, q^i) we consider loops γ lying entirely on this coordinate neighborhood, that is, $N_\gamma = 1$. Then, integrating by parts and taking into account the periodicity of η and ξ, we obtain

$$\langle (\Omega^\Delta)'(\gamma)\xi, \eta\rangle = -\int_{S^1} \left(\frac{\partial Q_i}{\partial q^j}\, \eta^i - \frac{d}{dt}\left(\frac{\partial Q_i}{\partial \dot q^j}\, \eta^i \right) + \frac{d^2}{dt^2}\left(\frac{\partial Q_i}{\partial \ddot q^j}\, \ddot\eta^i \right) \right) \xi^j .$$

Hence,

$$\int_{S^1} \left(\frac{\partial Q_i}{\partial q^j}\, \eta^i - \frac{d}{dt}\left(\frac{\partial Q_i}{\partial \dot q^j}\, \eta^i \right) + \frac{d^2}{dt^2}\left(\frac{\partial Q_i}{\partial \ddot q^j}\, \eta^i \right) \right) \xi^j =$$

$$\int_{S^1} \left(\frac{\partial Q_i}{\partial q^j}\, \eta^j + \frac{\partial Q_i}{\partial \dot q^j}\, \dot\eta^j + \frac{\partial Q_i}{\partial \ddot q^j}\, \ddot\eta^j \right) \xi^i ,$$

and since ξ is arbitrary, we obtain

$$\frac{\partial Q_i}{\partial q^j}\, \eta^i - \frac{d}{dt}\left(\frac{\partial Q_i}{\partial \dot q^j}\, \eta^i \right) + \frac{d^2}{dt^2}\left(\frac{\partial Q_i}{\partial \ddot q^j}\, \eta^i \right) = \frac{\partial Q_i}{\partial q^j}\, \eta^j + \frac{\partial Q_i}{\partial \dot q^j}\, \dot\eta^j + \frac{\partial Q_i}{\partial \ddot q^j}\, \ddot\eta^j .$$

But, η is also arbitrary, whence

$$\begin{aligned} \frac{\partial Q_i}{\partial \ddot q^j} &= \frac{\partial Q_j}{\partial \ddot q^i} \\ \frac{\partial Q_i}{\partial \dot q^j} + \frac{\partial Q_j}{\partial \dot q^i} &= \frac{d}{dt}\left(\frac{\partial Q_i}{\partial \ddot q^j} + \frac{\partial Q_j}{\partial \ddot q^i} \right) \\ \frac{\partial Q_i}{\partial q^j} - \frac{\partial Q_j}{\partial q^i} &= \frac{1}{2}\frac{d}{dt}\left(\frac{\partial Q_i}{\partial \dot q^j} - \frac{\partial Q_j}{\partial \dot q^i} \right), \end{aligned}$$

and ω satisfies the Helmholtz conditions.

Conversely, let us assume that ω satisfies the Helmholtz conditions. So, reverting the computations above on each local coordinate system we obtain

$$\int_{t_l}^{t_{l+1}} \left(\frac{\partial Q_i^l}{\partial q_l^j}\, \eta_l^i - \frac{d}{dt}\left(\frac{\partial Q_i^l}{\partial \dot q_l^j}\, \eta_l^i \right) + \frac{d^2}{dt^2}\left(\frac{\partial Q_i^l}{\partial \ddot q_l^j}\, \eta^i \right) \right) \xi_l^j =$$

$$\int_{t_l}^{t_{l+1}} \left(\frac{\partial Q_i^l}{\partial q_l^j}\, \eta_l^j + \frac{\partial Q_i^l}{\partial \dot q_l^j}\, \dot\eta_l^j + \frac{\partial Q_i^l}{\partial \ddot q_l^j}\, \ddot\eta_l^j \right) \xi_l^i ,$$

for any local section ξ_l and η_l of $\gamma^*\mathrm{TM}$ over each interval $[t_l, t_{l+1}]$. Now, a simple computation shows that

$$\left(\frac{\partial Q_i^l}{\partial q_l^j}\eta^i \ \frac{d}{dt}\left(\frac{\partial Q_i^l}{\partial \dot{q}_l^j}\eta_l^i\right) + \frac{d^2}{dt^2}\left(\frac{\partial Q_i^l}{\partial \ddot{q}_l^j}\eta_l^i\right)\right)\xi_l^j =$$

$$\left(\frac{\partial Q_i^l}{\partial q_l^j}\xi_l^j + \frac{\partial Q_i^l}{\partial \dot{q}_l^j}\dot{\xi}_l^j + \frac{\partial Q_i^l}{\partial \ddot{q}_l^j}\ddot{\xi}_l^j\right)\eta_l^i + \frac{d}{dt}\left[\xi_l^j\frac{d}{dt}(\eta_l^i\frac{\partial Q_i^l}{\partial \ddot{q}_l^j}) - \frac{\partial Q_i^l}{\partial \ddot{q}_l^j}\eta_l^i\dot{\xi}_l^j - \frac{\partial Q_i^l}{\partial \dot{q}_l^j}\eta_l^i\xi_l^j\right]$$

But, using (17) and (18) it is a simple matter to show that

$$\xi_l^j\frac{d}{dt}(\eta_l^i\frac{\partial Q_i^l}{\partial \ddot{q}_l^j}) - \frac{\partial Q_i^l}{\partial \ddot{q}_l^j}\eta_l^i\dot{\xi}_l^j - \frac{\partial Q_i^l}{\partial \dot{q}_l^j}\eta_l^i\xi_l^j = \frac{\partial Q_i^l}{\partial \ddot{q}_l^j}\left(\dot{\eta}_l^i\xi_l^j - \eta_l^i\dot{\xi}_l^j\right) - \frac{1}{2}\left(\frac{\partial Q_i^l}{\partial \dot{q}_l^j} - \frac{\partial Q_j^l}{\partial \dot{q}_l^i}\right)\eta_l^i\xi_l^j.$$

Finally, the right hand side of the previous equation does not depend on the local coordinate system and this concludes the proof. □

The previous Theorem shows that the Helmholtz conditions are necessary for the solution of the inverse problem. But, as the following example shows, they are not sufficient in general.

Example 3. Let M be a manifold with non trivial first de Rham cohomology group and $\tilde{\alpha}$ be any closed Pfaffian form on M which is not exact. Then, given a function T on TM let ω be the Birkhoffian defined as

$$X^*\omega = X^*\omega_T + \tau_{\mathrm{M}}^*\tilde{\alpha}$$

for all cross sections $X : \mathrm{TM} \to J^2(\mathrm{M})$. It is clear that on any local coordinate neighborhood we can define a local Lagrangian function and so ω satisfies the Helmholtz conditions. However, because ω is not conservative, there can be no global Lagrangian function L such that $\omega = \omega_{\mathrm{L}}$.

Definition 15. *Let* M *be a smooth manifold and* ω *a Birkhoffian on* M. *A smooth function* ψ *on* TM *is called a* gauge *function if* $\omega_\psi = 0$. *In other words, a gauge function generates trivial dynamics. A local gauge function is defined in an obvious way.*

Let $\mathfrak{U} = \{\Psi_\alpha\}$ be an open covering of M. A family $\psi = \{\psi_{\alpha\beta}\}$ of smooth functions

$$\psi_{\alpha\beta} : \mathrm{T}(\Psi_\alpha \cap \Psi_\beta) \to \mathbb{R}$$

defined for all α and β such that $\Psi_\alpha \cap \Psi_\beta \neq \varnothing$ is called a *cocycle* of the covering $\mathfrak{U}$ if it satisfies

$$\begin{aligned} \psi_{\alpha\beta} &= -\psi_{\beta\alpha}, \text{ on } \mathrm{T}\Psi_\alpha \cap \mathrm{T}\Psi_\beta && (23)\\ \psi_{\alpha\gamma} &= \psi_{\alpha\beta} + \psi_{\beta\gamma}, \text{ on } \mathrm{T}\Psi_\alpha \cap \mathrm{T}\Psi_\beta \cap \mathrm{T}\Psi_\gamma && (24) \end{aligned}$$

for all α,β and γ such that $\Psi_\alpha \cap \Psi_\beta \cap \Psi_\gamma \neq \varnothing$.

We say that two cocycles $\psi = \{\psi_{\alpha\beta}\}$ and $\bar{\psi} = \{\bar{\psi}_{\alpha\beta}\}$ are *cohomologous* if for each α there exists a local gauge function

$$\phi_\alpha : \mathrm{T}\Psi_\alpha \to \mathbb{R}$$

which for any pair of indexes α and β such that $\Psi_\alpha \cap \Psi_\beta \neq \varnothing$ and any point point $u \in \mathrm{T}\Psi_\alpha \cap \mathrm{T}\Psi_\beta$ the following equality holds

$$\psi_{\alpha\beta}(u) - \bar{\psi}_{\alpha\beta}(u) = \phi_\alpha(u) - \phi_\beta(u).$$

Clearly, the cohomology relation between two cocycles is an equivalence relation. The corresponding classes are called cohomology classes of the covering $\mathfrak{U}$. We denote by $[\psi]$ the cohomology class of the cocycle ψ and by $\mathrm{H}^1(\mathfrak{U})$ the set of all cohomology classes of $\mathfrak{U}$.

Definition 16. *A smooth* Helmholtz covering $\mathfrak{H} = (\mathfrak{U}, \{\psi_\alpha\})$ *on a manifold* M *consists of an open covering* $\mathfrak{U} = (\Psi_\alpha)$ *of* M *and a family of functions*

$$\psi_\alpha : \mathrm{T}\Psi_\alpha \to \mathbb{R}$$

such that if $\Psi_\alpha \cap \Psi_\beta \neq \varnothing$, *then* $\omega_{\psi_\alpha}|_{\Psi_\alpha \cap \Psi_\beta} = \omega_{\psi_\beta}|_{\Psi_\alpha \cap \Psi_\beta}$, *where* ω_{ψ_α} *denotes the local Birkhoffian on* Ψ_α *associated to* ψ_α.

The cocycle $\psi_{\alpha\beta} = \psi_\alpha - \psi_\beta$ *for all* $\Psi_\alpha \cap \Psi_\beta$ *is called the* Helmholtz cocycle *associated with the Helmholtz covering.*

Proposition 10. *Let* M *be a smooth manifold and* ω *a Birkhoffian on* M. *Suppose* ω *satisfies the principle of reciprocity, is affine and satisfies the Helmholtz conditions, then for any covering* $\mathfrak{U} = \{\Psi_\alpha\}$ *of* M *we can define a smooth Helmholtz covering* $\mathfrak{H} = (\mathfrak{U}, \{\mathrm{L}_\alpha\})$ *on* M.

Proof. Under the conditions of the Proposition a classical result shows that on any open set $\Psi_\alpha \in \mathfrak{U}$ there exits a local Lagrangian function L_α, such that, in local natural coordinates

$$Q_i(q,\dot{q},\ddot{q}) = \frac{d}{dt}\frac{\partial \mathrm{L}_\alpha}{\partial \dot{q}^i} - \frac{\partial \mathrm{L}_\alpha}{\partial q^i}\,.$$

Now, if $\Psi_\alpha \cap \Psi_\beta \neq \varnothing$ we have

$$\omega_{\mathrm{L}_\alpha}|_{\Psi_\alpha \cap \Psi_\beta} = \omega_{\mathrm{L}_\beta}|_{\Psi_\alpha \cap \Psi_\beta},$$

and the result follows by the $\mathbb{R}$-linearity of the map $\mathfrak{F}(\mathrm{TM}) \ni \mathrm{L} \mapsto \omega_\mathrm{L} \in \mathfrak{U}_1(J^2(\mathrm{M}))$. $\square$

Theorem 5. *Let* ω *be a Birkhoffian of the configuration space* M. *The following statements are equivalent:*

1. $\omega = \omega_L$ *for some function* L *on* TM.

2. ω *is such that:*

- ω *is affine;*
- ω *satisfies the Helmholtz conditions;*
- *if* $\mathfrak{U} = \{\Psi_\alpha\}$ *is any covering of* M *the associated Helmholtz cocycle* $\{\mathrm{L}_\alpha\}$ *is cohomologous to* $[0]$.

Proof. (1) $\Rightarrow$ (2) is clear. Now, suppose ω is affine and satisfies the Helmholtz conditions, then given any open covering $\mathfrak{U} = \{\Psi_\alpha\}$ of M by Proposition 10 it is defined a Helmholtz covering $\mathfrak{H} = (\mathfrak{U}, \{\mathrm{L}_\alpha\})$.

Fix a point $v \in \mathrm{TM}$ and let $u \in \mathrm{TM}$ be another point. Now, since M is connected, TM is connected and also pathwise connected. So, let γ be a (compact) path connecting v to u. Let $\{\Psi_1, \dots, \Psi_{N_\gamma}\}$ be a finite open covering of γ such that $v \in \mathrm{T}\Psi_{i_v}$ and $u \in \mathrm{T}\Psi_{i_u}$, for some $i_u, i_v \in \{1, \dots, N_\gamma\}$. Then, by the condition on the cohomology of the Helmholtz covering there exists $\{\phi_1, \dots, \phi_n\}$, ϕ_i defined on $\mathrm{T}\Psi_i$, $i = 1, \dots, N_\gamma$ such that $\mathrm{L}_i - \phi_i = \mathrm{L}_j - \phi_j$ on $\mathrm{T}\Psi_i \cap \mathrm{T}\Psi_j \neq \varnothing$. We define the value $\mathrm{L}(u) = \mathrm{L}_{i_u}(u) - \phi_{i_u}(u)$, that does not depend on γ. Indeed, let γ' be another path connecting v to u. Then $\mathrm{L}_{i'_v} - \phi_{i'_v} = \mathrm{L}_{i_v} - \phi_{i_v}$ and so $\mathrm{L}'(u) = \mathrm{L}_{i'_u} - \phi_{i'_u} = \mathrm{L}_{i_u} - \phi_{i_u} = \mathrm{L}(u)$, as required.

Hence, it is well defined a function $\mathrm{L} : \mathrm{TM} \to \mathbb{R}$. Since its restriction to a set $\Psi_\alpha \in \mathfrak{U}$ coincides with the difference of two smooth functions, it is smooth. Also, because ϕ_α are gauge functions we have $\omega|_{\Psi_\alpha} = \omega_{\mathrm{L}_\alpha - \phi_\alpha} = \omega_{\mathrm{L}}|_{\Psi_\alpha}$ on any $\Psi_\alpha \in \mathfrak{U}$. □

Example 4. Another important class of Birkhoffians which conserve energy, but, in general, are not Lagrangians, are the so-called *generalized magnetic fields* (see, for example, Wojtkowski (2000)). A closed 2-form $\mathfrak{b}$ on TM is called a generalized magnetic field (gmf), if there is a vector bundle morphism $B : \mathrm{TM} \to \mathrm{T}^*\mathrm{M}$ such that

$$\mathfrak{b}(z, \bar{z}) = B(\mathrm{T}\tau_{\mathrm{M}} z)(\mathrm{T}\tau_{\mathrm{M}} \bar{z})$$

for all $z, \bar{z} \in \mathrm{T}(\mathrm{TM})$.

In this case given a gmf $\mathfrak{b}$ and a Lagrangian L the Birkhoffian ω for the data $(\mathrm{L}, \mathfrak{b})$ is the 1-form which satisfies

$$X^*(\omega - \omega_L) = \mathrm{i}_Y \mathfrak{b},$$

for all second order vector fields $Y = i \circ X$. Conservation of the energy $E = Z(\mathrm{L}) - \mathrm{L}$ follows at once from $X^*\omega(Y) = X^*\omega_L(Y) + \mathfrak{b}(Y, Y) = X^*\omega_L(Y) = \mathsf{d}E_L(Y)$, for any second order vector field $Y = i \circ X$.

In local natural coordinates one obtains

$$\omega = \left(\frac{d}{dt} \frac{\partial \mathrm{L}}{\partial \dot{q}^i} - \frac{\partial \mathrm{L}}{\partial q^i} + \mathfrak{b}_{ji}(q, \dot{q}) \dot{q}^j \right) \mathrm{d}q^i$$

where $\mathfrak{b} = \mathfrak{b}_{ij}(q,\dot{q})\mathrm{d}q^i \wedge \mathrm{d}q^j$. So, in order to satisfy the Helmholtz conditions we must have

$$\begin{aligned}\left(\frac{\partial \mathfrak{b}_{ji}}{\partial \dot{q}^k} + \frac{\partial \mathfrak{b}_{jk}}{\partial \dot{q}^i}\right)\dot{q}^j &= 0\\ \left(\frac{\partial \mathfrak{b}_{ji}}{\partial q^k} - \frac{\partial \mathfrak{b}_{jk}}{\partial q^i}\right)\dot{q}^j &= \frac{1}{2}\dot{q}^l\frac{\partial}{\partial q^l}\left(\frac{\partial \mathfrak{b}_{ji}\dot{q}^j}{\partial \dot{q}^k} - \frac{\partial \mathfrak{b}_{jk}\dot{q}^j}{\partial \dot{q}^i}\right)\end{aligned}$$

which are always true when B can be identified with a two-covariant tensor over M, but clearly do not hold in general.

Remark 10. Liouville's Theorem states that the flow of a Hamiltonian vector field preserves the phase volume (see Proposition 3.3.4 in Abraham & Marsden (1978)). This result can be easily extended to the non-conservative case. Let ζ be a symplectic form on TM, π a force field, H a smooth function on TM. Suppose that the triple (ζ, H, π) defines a regular Birkhoffian ω on M (for instance, if $\zeta = -\mathsf{d}\mathsf{d}_v\mathrm{L}$ for some regular Lagrangian) such that

$$\bar{Y}^*\omega = -\mathrm{i}_Y\zeta + \mathsf{d}H - \pi,$$

for all second order vector fields $Y = i \circ \bar{Y}$. Denote the phase volume by

$$\Omega_\zeta := \underbrace{\zeta \wedge \ldots \wedge \zeta}_{n \text{ times}},$$

Then, the flow of the Birkhoff vector field Y given by Theorem 1 preserves the phase volume Ω_ζ if and only if $\mathsf{d}\pi \wedge \underbrace{\zeta \wedge \ldots \wedge \zeta}_{n-1 \text{ times}} = 0$. Indeed, denoting by Θ the Lie derivative, we have

$$\Theta(Y)\Omega_\zeta = (\Theta(Y)\zeta) \wedge \ldots \wedge \zeta + \cdots + \zeta \wedge \ldots \wedge (\Theta(Y)\zeta) = n(\Theta(Y)\zeta) \wedge \ldots \wedge \zeta,$$

but, by Cartan's magic formula, we have

$$\Theta(Y)\zeta = \mathrm{i}_Y\mathsf{d}\zeta + \mathsf{d}\mathrm{i}_Y\zeta = \mathsf{d}(\mathsf{d}H - \pi) = -\mathsf{d}\pi.$$

4 Constrained Birkhoff systems

Start with $(\mathrm{M}, \omega, \mathscr{C})$ where ω is a Birkhoffian of M and $\mathscr{C}$ is a smooth constant rank affine sub-bundle of the affine bundle $\tau_J : J^2(\mathrm{M}) \to \mathrm{TM}$. By definition there exists a vector sub-bundle C ($\tau_{\mathrm{TM}}|_{\mathrm{C}} : \mathrm{C} \to \mathrm{TM}$) of $\tau_{\mathrm{TM}}|_{\mathrm{V(M)}} : \mathrm{V(M)} \to \mathrm{TM}$ and a second order vector field $S : \mathrm{TM} \to \mathscr{C}$ such that

$$S(v) + \mathrm{C}_v = \mathscr{C}_v, \quad \forall v \in \mathrm{TM}$$

Example 5. Given a smooth distribution $\mathscr{D}$ on TM one uses $\mathrm{C}_v := \mathscr{D}_v \cap \mathrm{V}_v(\mathrm{M})$ to construct a vector bundle $\mathrm{C} = \bigcup_{v \in \mathrm{TM}} \mathrm{C}_v$ that one assumes to have constant rank; so, if there exits a cross section $S : \mathrm{TM} \to J^2(\mathrm{M}) \cap \mathscr{D}$ one obtains a smooth constant rank affine sub-bundle $\mathscr{C} = S + \mathrm{C}$ of $J^2(\mathrm{M})$. This affine sub-bundle is independent of the choice of the cross section S, that is, given another cross section $S' : \mathrm{TM} \to J^2(\mathrm{M}) \cap \mathscr{D}$ we have $S' + \mathrm{C} = S + \mathrm{C}$. Indeed, it is an immediate consequence of $S - S' \in \mathrm{C}$ for any two cross sections $S : \mathrm{TM} \to J^2(\mathrm{M}) \cap \mathscr{D}$ and $S' : \mathrm{TM} \to J^2(\mathrm{M}) \cap \mathscr{D}$. This example is motivated by Weber (1986), who defined a system of constraints to be a set of m linearly independent 1-forms $\psi^k : P \to \mathrm{T}^*P$, $k = 1, \ldots, m$, where P is a $2n$-dimensional smooth manifold which admits a symplectic structure.

Example 6. The usual notion of linear, both holonomic or non-holonomic, affine constraints or more generally non-linear constraints can be dealt with by defining the constraint $\mathscr{C}$ on a submanifold of TM. For the sake of simplicity in the exposition, in what follows we present the results for the constraint defined on TM only.

The annihilator of the vector space C_v is the sub-space of $\mathrm{T}^*_p\mathrm{M}$, $p = \tau_\mathrm{M}(v)$:

$$\mathrm{C}^o_v = \left\{ \alpha \in \mathrm{T}^*_p\mathrm{M} \,\middle|\, \alpha(\lambda_v^{-1} u) = 0, \forall u \in \mathrm{C}_v \right\}.$$

Definition 17. *A motion compatible with $\mathscr{C}$ is a smooth curve $\phi : I \to \mathrm{M}$ such that its lifting $\frac{T}{dt}\left(\frac{T\phi}{dt}\right) : I \to J^2(\mathrm{M})$ has its values on $\mathscr{C}$.*

4.1 The d'Alembert-Birkhoff principle for constrained mechanical systems

A triple $(\mathrm{M}, \omega, \mathscr{C})$ as above is called a *constrained Birkhoff system.*

Definition 18. *A constrained Birkhoff system $(\mathrm{M}, \omega, \mathscr{C})$ is* regular *if:*

1. *$\bar{\omega}(z)$ is non-degenerate for all $z \in J^2(\mathrm{M})$ and there exists $z_0 \in \mathscr{C}_v$ such that $\hat{\omega}(z_0) \in \mathrm{C}^o_v$ with $v = \tau_J(z)$;*

2. *the following bilinear form is non-degenerate:*

$$(\alpha, \bar{\alpha}) \in \mathrm{C}^o_v \times \mathrm{C}^o_v \mapsto \bar{\omega}(z)(\alpha^\sharp, \bar{\alpha}^\sharp) \in \mathbb{R}$$

for all $z \in \mathscr{C}$, with $v = \tau_J(z)$ and $\alpha^\sharp$ defined as $\alpha(\cdot) = \bar{\omega}(z)(\alpha^\sharp, \cdot)$ for each $\alpha \in \mathrm{C}^o_v$.

Remark 11. We note that a sufficient condition for regularity is that, for each $z \in J^2(\mathrm{M})$, $\bar{\omega}(z)$ be either positive or negative defined.

Definition 19. *A motion $\phi : I \to \mathrm{M}$ compatible with $\mathscr{C}$ is a* d'Alembert-Birkhoff trajectory *for $(\mathrm{M}, \omega, \mathscr{C})$ if its lifting $\frac{T}{dt}\left(\frac{T\phi}{dt}\right)$ satisfies*

$$\hat{\omega}\left[\frac{T}{dt}\left(\frac{T\phi}{dt}\right)\right] \in \mathrm{C}^{o}_{\frac{T\phi}{dt}}.$$

The constrained differential system corresponding to the constrained Birkhoff system *$(\mathrm{M}, \omega, \mathscr{C})$ is the set*

$$D(\omega, \mathscr{C}) = \left\{ z \in \mathscr{C} \mid \hat{\omega}(z) \in \mathrm{C}^{o}_{\tau_J(z)} \right\}.$$

The Birkhoffian ω is said to satisfy the d'Alembert-Birkhoff principle *if for any point $z = \frac{T}{dt}\big|_{t=0} \frac{T\gamma}{dt}$ of the constrained differential system $D(\omega, \mathscr{C})$, the curve γ is a d'Alembert-Birkhoff trajectory.*

Theorem 6. *Let $(\mathrm{M}, \omega, \mathscr{C})$ be a regular constrained Birkhoff system. Then, there exists well defined in a neighborhood of $v \in \mathrm{TM}$ a local d'Alembert-Birkhoff vector field, that is, a local smooth second order vector-field whose base curves are d'Alembert-Birkhoff trajectories of $(\mathrm{M}, \omega, \mathscr{C})$.*

Proof. We prove the result in local coordinates. Let V be a coordinate neighborhood where the vector bundle C is the annihilator of some local differential forms, that is,

$$\mathrm{C}_v = \mathrm{Ann}(\theta^1_v, \ldots, \theta^m_v), \quad m \leq n$$

for all $v \in V$. In local natural coordinates these differential forms can be written as:

$$\theta^\nu = b^\nu_i \mathrm{d}\dot{q}^i, \quad \nu = 1, \ldots, m, i = 1, \ldots, n$$

We make the convention that $\mu, \nu = 1, \ldots, m$ and $i, j = 1, \ldots, n$.

Now, fix a second order vector field $S : \mathrm{TM} \to \mathscr{C}$, which in local natural coordinates is given by $S(q, \dot{q}) = (q, \dot{q}, \ddot{q}_s(q, \dot{q}))$. Then, in these coordinates the constraint equations are:

$$b^\nu_j(q, \dot{q})(\ddot{q}^j - \ddot{q}^j_s(q, \dot{q})) = 0$$

or

$$b^\nu_j(q, \dot{q})\ddot{q}^j + a^\nu(q, \dot{q}) = 0$$

with $a^\nu(q, \dot{q}) = -b^\nu_j(q, \dot{q})\ddot{q}^j_s(q, \dot{q})$.

The Birkhoffian is locally represented as $\omega = Q_i(q, \dot{q}, \ddot{q})\mathrm{d}q^i$. So, the d'Alembert-Birkhoff principle is locally equivalent to the following equations:

$$Q_i(q, \dot{q}, \ddot{q}) = b^\nu_i(q, \dot{q})\lambda_\nu \tag{25}$$
$$b^\nu_j(q, \dot{q})\ddot{q}^j + a^\nu(q, \dot{q}) = 0 \tag{26}$$

for some functions λ_ν.

Now, let $(q, \dot{q}, \ddot{q}_0)$ be a point which satisfy both (25) and (26) for some λ_0. From regularity of the Birkhoffian and the Implicit Function Theorem there is a neighborhood of $(q, \dot{q}, \ddot{q}_0)$ where we can write

$$\ddot{q}^i = \ddot{Q}^i(q, \dot{q}, \lambda).$$

Substituting this equality into (26) we get the auxiliary functions

$$h^\nu(q, \dot{q}, \lambda) = b^\nu_j(q, \dot{q})\ddot{Q}^j(q, \dot{q}, \lambda) + a^\nu(q, \dot{q})$$

Differentiating the latter with respect to λ we obtain:

$$\frac{\partial h^\nu}{\partial \lambda_\mu}(q, \dot{q}, \lambda) = b^\nu_j(q, \dot{q}) a^{ji}(q, \dot{q}, \ddot{Q}(q, \dot{q}, \lambda)) b^\mu_i(q, \dot{q})$$

where $a_{ij}(q, \dot{q}, \ddot{q}) = \frac{\partial Q_i}{\partial \ddot{q}^j}(q, \dot{q}, \ddot{q})$ and $a^{ij}(q, \dot{q}, \ddot{q})$ is the inverse of $a_{ij}(q, \dot{q}, \ddot{q})$. Condition 2 means that $\frac{\partial h^\nu}{\partial \lambda_\mu}(q, \dot{q}, \lambda)$ is an isomorphism in an open neighborhood of $(q, \dot{q}, \lambda_0)$. Finally, applying the Implicit Function Theorem we can locally write $\lambda_\nu = \Lambda_\nu(q, \dot{q})$ and the local vector field $\ddot{q}^i(q, \dot{q}) = \ddot{Q}^i(q, \dot{q}, \Lambda(q, \dot{q}))$ satisfies the principle of d'Alembert-Birkhoff as required. $\square$

Corollary 1 (Principle of determinism). *If in condition 1. of the definition of a regular* $(\mathrm{M}, \omega, \mathscr{C})$ *the element* $z_0 \in \mathscr{C}_v$ *is unique, one can define a global smooth d'Alembert-Birkhoff vector-field* $Y_\mathscr{C} = i \circ X_\mathscr{C}$ *and we have* $\mathrm{Im}\ X_\mathscr{C} = D(\omega, \mathscr{C})$*, that is, the principle of determinism holds.*

Next we consider time reversibility of constrained Birkhoff systems (for a constrained Lagrangian system, see Gorni & Zampieri (2000)). Let us define the mapping $\rho_2 : J^2(\mathrm{M}) \to J^2(\mathrm{M})$ as

$$\rho_2(z) = \left.\frac{T}{dt}\right|_{t=0} \frac{T\gamma(-t)}{dt}$$

where $z = \left.\frac{T}{dt}\right|_{t=0} \frac{T\gamma}{dt}$, for some curve $\gamma : I \to \mathrm{M}$. In local natural coordinates we have $\rho_2(q, \dot{q}, \ddot{q}) = (q, -\dot{q}, \ddot{q})$.

Proposition 11. *Let* $(\mathrm{M}, \omega, \mathscr{C})$ *be a constrained Birkhoff system with a regular and reversible Birkhoffian* ω*. Suppose* $\rho_2(\mathscr{C}) \subset \mathscr{C}$ *and let* $c(t)$ *be an integral curve of a d'Alembert-Birkhoff vector field* $Y_\mathscr{C}$*, then* $\rho \circ c(-t)$ *is also an integral curve of* $Y_\mathscr{C}$*. Thus,* $F_{-t}(x) = \rho(F_t(\rho(x)))$*, where* F_t *is the flow of* $Y_\mathscr{C}$*.*

Proof. We prove the result in local natural coordinates $(q, \dot{q}, \ddot{q})$. Let $q : I \to U$ denote a local base curve of $Y_\mathscr{C}$, then using the same notation as in Theorem 6 we have

$$\begin{aligned} Q(q(t), \dot{q}(t), \ddot{q}(t)) &\in \mathrm{C}^o_{(q(t), \dot{q}(t))} \\ b^\mu_i(q(t), \dot{q}(t))\ddot{q}^i(t) + a^\nu(q(t), \dot{q}(t)) &= 0, \end{aligned}$$

for all $t \in I$, where $Q(q(t), \dot{q}(t), \ddot{q}(t)) := (Q_1(q(t), \dot{q}(t), \ddot{q}(t)), \ldots, Q_n(q(t), \dot{q}(t), \ddot{q}(t)))$. Define $u(t) := q(-t)$, for all $t \in \mathbb{R}$ such that $-t \in I$; we have to show that

$$\begin{array}{rcl} Q(u(-t), \dot{u}(-t), \ddot{u}(-t)) & \in & \mathrm{C}^o_{(u(-t),\dot{u}(-t))} \\ b^\mu_i(u(-t), \dot{u}(-t))\ddot{u}^i(-t) + a^\nu(u(-t), \dot{u}(-t)) & = & 0, \end{array}$$

for all $t \in I$. But, $u(-t) = q(t)$, $\dot{u}(-t) = -\dot{q}$ and $\ddot{u}(-t) = \ddot{q}(t)$. The result then follows from:

$$Q_i(u(-t), \dot{u}(-t), \ddot{u}(-t)) = Q_i(q(t), -\dot{q}(t), \ddot{q}(t)) = Q_i(q(t), \dot{q}(t), \ddot{q}(t)),$$

by the hypothesis of reversibility of ω; and from

$$\begin{aligned} b^\mu_i(u(-t), \dot{u}(-t))\ddot{u}^i(-t) + a^\nu(u(-t), \dot{u}(-t)) &= b^\mu_i(q(t), -\dot{q}(t))\ddot{q}^i(t) + a^\nu(q(t), -\dot{q}(t)) \\ &= b^\mu_i(q(t), \dot{q}(t))\ddot{q}^i(t) + a^\nu(q(t), \dot{q}(t)) \\ &= 0, \end{aligned}$$

and $\mathrm{C}^o_{(u(-t),\dot{u}(-t))} = \mathrm{C}^o_{(q(t),-\dot{q}(t))} = \mathrm{C}^o_{(q(t),\dot{q}(t))}$, since $\rho_2(\mathscr{C}) \subset \mathscr{C}$. □

Definition 20. *Let $(\mathrm{M}, \omega, \mathscr{C})$ be a constrained Birkhoff system which satisfies the principle of determinism. Then, it defines a vector bundle morphism $R : \mathrm{TM} \to \mathrm{T^*M}$, called the* reaction field, *such that $\hat{\omega}(Y_{\mathscr{C}}(v)) - R(v) = 0$ for all $v \in \mathrm{TM}$, where Y denotes the d'Alembert-Birkhoff vector field.*

Definition 21 (Principle of least constraint of Gauss). *Let $(\mathrm{M}, \omega, \mathscr{C})$ be a constrained Birkhoff system. The Birkhoffian ω is said to satisfy the* principle of least constraint of Gauss *if any d'Alembert-Birkhoff vector field $Y_{\mathscr{C}}$ is such that, for all $v \in \mathrm{TM}$, $Y_{\mathscr{C}}(v)$ is a stationary point of the Gibbs function corresponding to the data (ω, Y_0) restricted to $\mathscr{C}_v$, where $Y_0 : \mathrm{TM} \to J^2(\mathrm{M})$ is the Birkhoff vector field associated to the (unconstrained) Birkhoff system (M, ω).*

The case where $\omega = \omega_L$, L being a regular Lagrangian function has been considered in Lewis (1996).

Proposition 12. *Let $(\mathrm{M}, \omega, \mathscr{C})$ be a constrained Birkhoff system. Assume ω is affine and satisfies the principle of reciprocity. Then, ω satisfies the d'Alembert-Birkhoff principle if and only if it satisfies the principle of least constraint of Gauss.*

Proof. In local coordinates we have

$$\mathfrak{G}(q, \dot{q}, \ddot{q}) = \frac{1}{2} a_{ij}(q, \dot{q})(\ddot{q}^i - \ddot{q}^i_0)(\ddot{q}^j - \ddot{q}^j_0)$$

with $b^\mu_j(q\dot{q})\ddot{q}^j + a^\mu(q, \dot{q}) = 0$ and where $(q, \dot{q}, \ddot{q}_0)$ is the local representation of Y_0. Then the requirement that $Y(v)$ be stationary is equivalent to the equations:

$$\begin{array}{rcl} \dfrac{\partial \mathfrak{G}}{\partial \ddot{q}^i}(q, \dot{q}, \ddot{q}) & = & \dfrac{1}{2}(a_{ij}(q, \dot{q}) + a_{ji}(q, \dot{q}))(\ddot{q}^j - \ddot{q}^j_0) = \lambda_\mu b^\mu_i \\ b^\mu_j(q, \dot{q})\ddot{q}^j + a^\mu(q, \dot{q}) & = & 0 \end{array}$$

And the result follows from the symmetry of $a_{ij}(q,\dot{q}) = a_{ji}(q,\dot{q})$. □

Example 7. There is a class of non-linear constraints which provides useful models in nonequilibrium statistical mechanics: the so called Gauss thermostats, or isokinetic dynamics (see, for example, Hoover (1986), Ruelle (1999), Gallavotti & Ruelle (1997) and Wojtkowski (2000)). Given a Riemannian manifold, the equations of motions are obtained by imposing the conservation of the kinetic energy and then applying the Gauss least constraint principle for the Birkhoffian associated with a classical Lagrangian. By the previous Proposition, this procedure fits in the proposed formulation and is clearly non-linear. More generally, given a conservative Birkhoffian ω, we may require the constancy of the energy and derive the equations for the isoenergetic dynamics using the d'Alembert-Birkhoff principle (in the case of classical Lagrangian see, for example, Wojtkowski (2000)).

Remark 12. It is easy to find counter-examples for the Gauss least constraint principle, if ω is neither affine nor symmetric. For example, let $\mathrm{M} = \mathbb{R}^2$ and consider the local representation in normal Cartesian coordinates (x,y) of the constrained Birkhoff system

$$\begin{bmatrix} 0 & 2 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \& \quad \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} + 1 = 0$$

Solution of the previous equations is $x(t) = -t^2 + \dot{x}_0 t + x_0$, $y(t) = \frac{1}{4}t^2 + \dot{y}_0 t + y_0$ and $\lambda(t) = 1$. The equations for the principle of least constraint are

$$\begin{bmatrix} 0 & 1/2 \\ 1/2 & 0 \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \& \quad \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} + 1 = 0$$

A simple computation shows that the solution which satisfies the principle of least constraint is $x(t) = -\frac{1}{4}t^2 + \dot{x}_0 t + x_0$, $y(t) = -\frac{1}{8}t^2 + \dot{y}_0 t + y_0$ and $\lambda(t) = -\frac{1}{8}$.

Corollary 2. *Let $(\mathrm{M}, \omega, \mathscr{C})$ be a constrained Birkhoff system. The following statements are equivalent:*

1. $\omega = \omega_L$ *for some regular Lagrangian function* L *on* TM *and* ω *satisfies the d'Alembert-Birkhoff principle for the constrained Birkhoff system* $(\mathrm{M}, \omega, \mathscr{C})$.
2. ω *is such that:*
   - ω *is regular;*
   - ω *satisfies the principle of reciprocity;*
   - ω *is affine;*
   - ω *satisfies the Helmholtz conditions;*
   - *if* $\mathfrak{U} = \{\Psi_\alpha\}$ *is any covering of* M *the associated Helmholtz co-cycle* $\{\mathrm{L}_\alpha\}$ *is co-homologous to* $[0]$*;*

- *ω satisfies the principle of least constraint of Gauss for $(\mathrm{M}, \omega, \mathscr{C})$.*

Remark 13. Assume that to the conditions of Remark 10 we add a regular constraint; then, an analogous computation shows that the flow of the d'Alembert-Bikhoff vector field $Y_{\mathscr{C}}$ preserves the phase volume Ω_ζ if and only if $\mathrm{d}(R+\pi) \wedge \underbrace{\zeta \wedge \ldots \wedge \zeta}_{n-1 \text{ times}} = 0$.

4.2 A generalization of Liouville's Theorem for classical affine constraints

In this subsection we prove a version of the Liouville's Theorem for classical Lagrangians and classical affine constraints. So, (M, g) is a Riemannian manifold and the function V on M is the potential energy. We denote by ∇ the associated Levi-Civita connection. The affine constraint $\mathscr{A}$ is defined by a pair $(\mathscr{D}, X_a)$ where $\mathscr{D}$ is a smooth, constant rank distribution on M (rank $\mathscr{D} = m$) and X_a is a (global) cross section of $\mathscr{D}^\perp$. With $(\mathscr{D}, X_a)$ we define

$$\mathscr{A} = \mathscr{A}(\mathscr{D}, X_a) := \{v \in \mathrm{TM} \mid v - X_a(\tau_{\mathrm{M}}(v)) \in \mathscr{D}\}.$$

In order to construct a volume form on $\mathscr{A}$ we start, locally, choosing an adapted system of coordinates for TM. Let $(U, q^1, \ldots, q^n)$ be a local system of coordinates for M, $(\xi_1, \ldots, \xi_n)$ a local orthonormal basis of vector fields defined on U such that $(\xi_1, \ldots, \xi_m)$ is a local orthonormal basis of sections of $\mathscr{D}$ and $(\epsilon^1, \ldots, \epsilon^n)$ its dual basis; they induce the functions $\bar{\epsilon}^i : \mathrm{T}U \to \mathbb{R}$

$$\bar{\epsilon}^i \stackrel{def}{=} \sum_{j=1}^{n} (a^i_j \circ \tau_M) \dot{q}^j, \quad i = 1, 2, \ldots, n \tag{27}$$

and one can also consider new local coordinates for TM: $(\mathrm{T}U; q^1, \ldots, q^n, \bar{\epsilon}^1, \ldots, \bar{\epsilon}^n)$.

The affine bundle $\mathscr{A}$ is an embedded affine subbundle of TM (considered as an affine bundle modelled by itself) and

$$\mathscr{A} \cap \mathrm{T}U = \{v \in \mathrm{T}U \mid \bar{\epsilon}^{m+1}(v - X_a(\tau_{\mathrm{M}}(v))) = \ldots = \bar{\epsilon}^n(v - X_a(\tau_{\mathrm{M}}(v))) = 0\}.$$

Then the restrictions of $q^1, \ldots, q^n, \bar{\epsilon}^1, \ldots, \bar{\epsilon}^m$ to the open set $\mathscr{A} \cap \mathrm{T}U$ define a local system of coordinates for $\mathscr{A}$ on $\mathscr{A} \cap \mathrm{T}U$; by consequence a volume form on $\mathscr{A} \cap \mathrm{T}U$ is defined by the restriction $\tilde{\Omega}$ to $\mathscr{A} \cap \mathrm{T}U$ of the $(m+n)$- form

$$\Omega = \epsilon^1 \wedge \ldots \wedge \epsilon^n \wedge \mathrm{d}_{\mathrm{TM}}\bar{\epsilon}^1 \wedge \ldots \wedge \mathrm{d}_{\mathrm{TM}}\bar{\epsilon}^m, \tag{28}$$

where each ϵ^i, in (28), means $\tau^*_{\mathrm{M}}\epsilon^i$, the pull-back of ϵ^i to $\mathrm{T}U$; so,

$$\epsilon^i \cong \tau^*_{\mathrm{M}}\epsilon^i = \sum_{j=1}^{n} (a^i_j \circ \tau_{\mathrm{M}}) \mathrm{d}_{\mathrm{TM}}(q^j \circ \tau_{\mathrm{M}}), \quad i = 1, \ldots, n. \tag{29}$$

Analogously, we also define the local $(n-m)$ form $\Omega^{\perp}$ on U as

$$\Omega^{\perp} := \epsilon^{m+1} \wedge \ldots \wedge \epsilon^{n}.$$

To obtain a (global) volume form on $\mathscr{A}$ one needs to assume that $\mathscr{A}$ is orientable as manifold. One way to obtain it is the following:

Definition 22. *The affine bundle $\mathscr{A} = \mathscr{A}(\mathscr{D}, X_a)$ on (M, g) is orientable if $\mathscr{D}$ is orientable, that is, there exists a differentiable exterior $(n-m)$-form Ψ on* M *such that, for any $p \in \mathrm{M}$ and $z_1, \ldots, z_{n-m} \in \mathscr{D}_q^{\perp}$ then $\Psi_p(z_1, \ldots, z_{n-m}) \neq 0$ if, and only if, $(z_1, \ldots, z_{n-m})$ is a basis of $\mathscr{D}_p^{\perp}$.*

Remark that in the codimension one case $(m = n-1)$, $\mathscr{A}$ orientable is equivalent to the existence of a globally defined unitary vector field N on M, orthogonal to $\mathscr{D}_p$, $\forall p \in \mathrm{M}$.

Let us recall the definition of the *total second fundamental form of a distribution* $\mathscr{D}$ (see Kupka & Oliva (2001)). One defines a bilinear vector bundle morphism $B_{\mathscr{D}} : \mathrm{TM} \times_{\mathrm{M}} \mathscr{D} \to \mathscr{D}^{\perp}$ where $\mathrm{TM} \times_{\mathrm{M}} \mathscr{D}$ is the fiber product of the bundles TM and $\mathscr{D}$ and $\mathscr{D}^{\perp}$ is the distribution orthogonal complement of $\mathscr{D}$ with respect to g: let $(\bar{X}, \bar{Y}) \in \mathrm{T}_p\mathrm{M} \times \mathscr{D}_p$ and choose two germs of vector fields at p, X, Y, so that $X(p) = \bar{X}$, $Y(p) = \bar{Y}$ and $Y \in \mathscr{D}$ (that means Y is a germ of section of $\mathscr{D}$). Then

$$B_{\mathscr{D}} := P_{\mathscr{D}^{\perp}}\left[(\nabla_X Y)(p)\right]. \tag{30}$$

This does not depend on the choice of the germs X and Y.

Theorem 7. *The (local) volume form $\tilde{\Omega}$ defined on $\mathscr{A} \cap \mathrm{T}U$ by formula (28) is invariant under the flow of the d'Alembert vector field X if, and only if,*

$$\Theta(X_a)\Omega^{\perp} = 0$$

and the trace of $B_{\mathscr{D}^{\perp}}|_{\mathscr{D}^{\perp} \times_M \mathscr{D}^{\perp}}$ vanishes ($B_{\mathscr{D}^{\perp}}$ is the total second fundamental form of $\mathscr{D}^{\perp}$). If $\mathscr{A}$ is orientable, both $\tilde{\Omega}$ and $\Omega^{\perp}$ can be extended to a global volume on $\mathscr{A}$ and a global section of $\underbrace{(\mathscr{D}^{\perp})^* \wedge \cdots \wedge (\mathscr{D}^{\perp})^*}_{n-m-times}$.

Proof. We use the technique proposed in Kupka & Oliva (2001) for the case with $X_a = 0$. Let $q_0 : [a_0, a_1] \to U \subset \mathrm{M}$ be a solution of

$$P_{\mathscr{D}} \nabla_t \frac{Tq_0}{dt} = 0,$$

where $P_{\mathscr{D}} : \mathrm{TM} \to \mathscr{D}$ is the orthogonal projection. Define $v_0 : [a_0, a_1] \to \mathrm{T}U \cap \mathscr{D}$ as

$$v_0 = \frac{Tq_0}{dt} - X_a \circ q_0,$$

then

$$\nabla_t v_0 = B_{\mathscr{D}}(v_0, v_0) + B_{\mathscr{D}}(X_a, v_0) - B_{\mathscr{D}^{\perp}}(v_0, X_a) - B_{\mathscr{D}^{\perp}}(X_a, X_a), \tag{31}$$

so that,

$$
\begin{aligned}
X \;=\; & \sum_{i=1}^{n} \dot{q}^i \frac{\partial}{\partial q^i} - \sum_{i=1}^{m} \sum_{j,k=1}^{m} A^i_{kj} \bar{\epsilon}^k \bar{\epsilon}^j \frac{\partial}{\partial \bar{\epsilon}^i} \\
- & \sum_{i=1}^{m} \left(\sum_{\alpha=m+1}^{n} \sum_{k=1}^{m} A^i_{\alpha k} \bar{\epsilon}^k X^\alpha_a + \sum_{j=1}^{m} \sum_{\alpha=m+1}^{n} A^i_{j\alpha} \bar{\epsilon}^j X^\alpha_a + \right. \\
& \left. \sum_{\alpha=m+1}^{n} \sum_{\beta=m+1}^{n} A^i_{\alpha\beta} X^\alpha_a X^\beta_a \right) \frac{\partial}{\partial \bar{\epsilon}^i}.
\end{aligned} \tag{32}
$$

We want to know under which conditions the volume form Ω, given by

$$\Omega = \epsilon^1 \wedge \ldots \wedge \epsilon^n \wedge \mathsf{d}_{\mathrm{TM}} \bar{\epsilon}^1 \wedge \ldots \wedge \mathsf{d}_{\mathrm{TM}} \bar{\epsilon}^m,$$

is invariant under X. Let us assume, for a moment, that $V = 0$; we start by computing the Lie derivative $\Theta(X)\Omega$:

$$\Theta(X)\Omega = \textstyle\sum_{i=1}^{n} \epsilon^1 \wedge \ldots \wedge \Theta(X)\, \epsilon^i \wedge \ldots \wedge \epsilon^n \wedge \mathsf{d}_{\mathrm{TM}} \bar{\epsilon}^1 \wedge \ldots \wedge \mathsf{d}_{\mathrm{TM}} \bar{\epsilon}^m +$$

$$\textstyle\sum_{j=1}^{m} \epsilon^1 \wedge \ldots \wedge \epsilon^n \wedge \mathsf{d}_{\mathrm{TM}} \bar{\epsilon}^1 \wedge \ldots \wedge \Theta(X)\, \mathsf{d}_{\mathrm{TM}} \bar{\epsilon}^j \wedge \ldots \wedge \mathsf{d}_{\mathrm{TM}} \bar{\epsilon}^m.$$

Thus, in order to obtain $\Theta(X)\Omega$ it is enough to compute $\Theta(X)\epsilon^i$ modulo

$$\epsilon^1, \ldots, \epsilon^{i-1}, \epsilon^{i+1}, \ldots, \epsilon^n, \mathsf{d}_{\mathrm{TM}} \bar{\epsilon}^1, \ldots, \mathsf{d}_{\mathrm{TM}} \bar{\epsilon}^m,$$

and, analogously, to compute $\Theta(X)\mathsf{d}_{\mathrm{TM}} \bar{\epsilon}^j$ modulo

$$\mathsf{d}_{\mathrm{TM}} q^1, \ldots, \mathsf{d}_{\mathrm{TM}} q^n, \mathsf{d}_{\mathrm{TM}} \bar{\epsilon}^1, \ldots, \mathsf{d}_{\mathrm{TM}} \bar{\epsilon}^{j-1}, \mathsf{d}_{\mathrm{TM}} \bar{\epsilon}^{j+1}, \ldots, \mathsf{d}_{\mathrm{TM}} \bar{\epsilon}^m.$$

But $\Theta(X)\epsilon^i = \mathrm{i}_X \mathsf{d}_{\mathrm{TM}} \epsilon^i + \mathsf{d}_{\mathrm{TM}}(\epsilon^i(X))$, so

$$\Theta(X)\epsilon^i \equiv \begin{cases} \mathrm{i}_X \mathsf{d}_{\mathrm{TM}} \epsilon^i & \text{for } i = 1, \ldots, m \\ \mathrm{i}_X \mathsf{d}_{\mathrm{TM}} \epsilon^i + \left(\mathsf{d}_{\mathrm{TM}} X^i_a\right)_i \epsilon^i & \text{for } i = m+1, \ldots, n. \end{cases}$$

(observe that $\epsilon^\alpha(X) = X^\alpha_a, \alpha = m+1, \ldots, n$). We also have $\mathsf{d}_{\mathrm{TM}} \epsilon^i = -\sum_{k=1}^{n} \omega^i_k \wedge \epsilon^k$ where, with the simplified notation, ω^i_k and ϵ^k mean $\tau^*_{\mathrm{M}} \omega^i_k$ and $\tau^*_{\mathrm{M}} \epsilon^k$, respectively. Thus

$$\Theta(X)\epsilon^i \equiv \begin{cases} -\sum_{k=1}^{n} \omega^i_k(X)\epsilon^k + \sum_{k=1}^{n} \epsilon^k(X)\omega^i_k & \text{for } i = 1, \ldots, m \\ -\sum_{k=1}^{n} \omega^i_k(X)\epsilon^k + \sum_{k=1}^{n} \epsilon^k(X)\omega^i_k + \left(\mathsf{d}_{\mathrm{TM}} X^i_a\right)_i \epsilon^i & \text{for } i = m+1, \ldots, n \end{cases}$$

Since $\omega^i_k = \sum_{r=1}^{n} A^i_{rk} \epsilon^r$ and $A^i_{ri} = 0$ $(\omega^i_i = 0)$ for all $r = 1, \ldots, n$, we write

$$\Theta(X)\epsilon^i \equiv \sum_{j=1}^{n} A^i_{ij} \epsilon^j(X) \epsilon^i \equiv \sum_{j=1}^{m} A^i_{ij} \bar{\epsilon}^j \epsilon^i + \sum_{\beta=m+1}^{n} A^i_{i\beta} X^\beta_a \epsilon^i. \tag{33}$$

for $i = 1, \ldots, m$. Similarly, we have,

$$\Theta(X)\epsilon^{\alpha} \equiv \sum_{j=1}^{m} A^{\alpha}_{\alpha j}\bar{\epsilon}^{j}\epsilon^{\alpha} + \sum_{\beta=m+1}^{n} A^{\alpha}_{\alpha\beta}X^{\beta}_{a}\epsilon^{\alpha} + (\mathsf{d}_{\mathrm{TM}}X^{\alpha}_{a})_{\alpha}\,\epsilon^{\alpha}. \tag{34}$$

for $\alpha = m+1, \ldots, n$.

On the other hand $\Theta(X)\mathsf{d}_{\mathrm{TM}}\bar{\epsilon}^{j} = \mathsf{d}_{\mathrm{TM}}(\Theta(X)\bar{\epsilon}^{j})$ and $\Theta(X)\bar{\epsilon}^{j}$ is the component of X along $\bar{\epsilon}^{j}$. Then, from (32) we obtain

$$\Theta(X)\mathsf{d}_{\mathrm{TM}}\bar{\epsilon}^{i} \equiv -\sum_{l=1}^{m} A^{i}_{il}\bar{\epsilon}^{l}\mathsf{d}_{\mathrm{TM}}\bar{\epsilon}^{i} - \sum_{\alpha=m+1}^{n} A^{i}_{i\alpha}X^{\alpha}_{a}\mathsf{d}_{\mathrm{TM}}\bar{\epsilon}^{i}, \tag{35}$$

Finally, from (33), (34) and (35) we have

$$\Theta(X)\Omega = \left[\sum_{\alpha=m+1}^{n}\sum_{j=1}^{m} A^{\alpha}_{\alpha_j}\bar{\epsilon}^{j} + \sum_{\alpha=m+1}^{n}(\mathsf{d}_{\mathrm{TM}}X^{\alpha}_{a})_{\alpha} + \sum_{\alpha=m+1}^{n}\sum_{\beta=1}^{n} A^{\alpha}_{\alpha\beta}X^{\beta}_{a}\right]\Omega. \tag{36}$$

This last equation (36) shows that $\Theta(X)\Omega = 0$ if, and only if,

$$\sum_{\alpha=m+1}^{n} A^{\alpha}_{\alpha j} = 0, \quad j = 1, \ldots, m \tag{37}$$

and

$$\sum_{\alpha=m+1}^{n}(\mathsf{d}_{\mathrm{TM}}X^{\alpha}_{a})_{\alpha} + \sum_{\alpha=m+1}^{n}\sum_{\beta=1}^{n} A^{\alpha}_{\alpha\beta}X^{\beta}_{a} = 0. \tag{38}$$

The intrinsic interpretation of conditions (37) comes from the consideration of the total second fundamental form $B_{\mathscr{D}^{\perp}}$ of the distribution $\mathscr{D}^{\perp}$. In fact

$$\mathrm{tr}\ B^{j}|_{\mathscr{D}^{\perp}\times_M\mathscr{D}^{\perp}} = \sum_{\alpha=m+1}^{n} B^{j}(\xi_{\alpha},\xi_{\alpha}) = \sum_{\alpha=m+1}^{n} A^{j}_{\alpha\alpha} = -\sum_{\alpha=m+1}^{n} A^{\alpha}_{\alpha j}. \tag{39}$$

where $B_{\mathscr{D}^{\perp}}(x,y) = \sum_{j=1}^{m} B^{j}(x,y)\xi_{j}$, for all $(x,y) \in \mathrm{TM} \oplus \mathscr{D}^{\perp}$. On the other hand, an analogous computation shows that condition (38) is equivalent to

$$\Theta(X_a)(\epsilon^{m+1} \wedge \cdots \wedge \epsilon^{n}) = 0 \tag{40}$$

Thus from (37), (38), (39) and (40) the proof of Theorem 7 follows for $V = 0$. When $V \neq 0$ we have a similar proof because in that more general case, denoting the vector field by $X(V)$, one can write

$$X(V) = X - \sum_{j=1}^{m} V^{j}\frac{\partial}{\partial\bar{\epsilon}^{j}}$$

where $V^j = \epsilon^j(\text{grad} \quad V)$, $j = 1, \ldots, m$. Then, clearly, $\Theta(X(V))\epsilon^i \equiv \Theta(X)\epsilon^i$, $i = 1, \ldots, n$, and $\Theta(X(V))\bar{\epsilon}^j = \Theta(X)\bar{\epsilon}^j + V^j$, $j = 1, \ldots, m$. So, $\Theta(X(V))\mathsf{d}_{\mathrm{TM}}\bar{\epsilon}^j = \Theta(X)\mathsf{d}_{\mathrm{TM}}\bar{\epsilon}^j + \mathsf{d}_{\mathrm{TM}}V^j$; but, we have $\mathsf{d}_{\mathrm{TM}}V^j \equiv 0$ modulo $\epsilon^1, \ldots, \epsilon^n$, because V depends only on q. Then $\Theta(X(V))\Omega = \Theta(X)\Omega$ and the proof is now complete. □

5 Symmetry and reduction of Birkhoff systems

Definition 23. *Let* (M, ω) *be a Birkhoff system and* $\Phi : \mathrm{G} \times \mathrm{M} \to \mathrm{M}$ *be a smooth action of a Lie group* G *on* M. *Denote by* $\mathfrak{g}$ *the Lie algebra of the Lie group* G *and by* $\mathfrak{g}^*$ *its dual. We say that a mapping*

$$\mu : \mathrm{TM} \to \mathfrak{g}^*$$

is a momentum mapping *relative to* ω *for the action* Φ, *provided that for each* $X \in \mathfrak{g}$,

$$\mathsf{d}\mu^X(Y) = (\bar{Y}^*\omega)X^{\#}_{\mathrm{TM}}$$

for all second order vector fields $Y = i \circ \bar{Y}$. *Here* $\mu^X : \mathrm{TM} \to \mathbb{R}$, $\mu^X(v) := \langle \mu(v), X \rangle$, $v \in \mathrm{TM}$, *is the* component of μ along X *and* $X^{\#}_{\mathrm{TM}}$ *is infinitesimal generator corresponding to* X *with respect to the tangent action* $\Phi^T : \mathrm{G} \times \mathrm{TM} \to \mathrm{TM}$, $\Phi^T_g = \mathrm{T}\Phi_g : \mathrm{TM} \to \mathrm{TM}$. *The quadruple* $(\mathrm{M}, \omega, \Phi, \mu)$ *will be called a* Noether G-space.

In the present context, Noether's Theorem is an immediate consequence of the definitions of Birkhoff vector field and momentum mapping.

Proposition 13 (Noether). *Let* $(\mathrm{M}, \omega, \Phi, \mu)$ *be a Noether G-space. Then,* μ *is an integral of any Birkhoff vector field* Y *associated with the Birkhoff system* (M, ω), *that is, if* F_t *is the flow of* Y,

$$\mu \circ F_t = \mu.$$

Example 8. Let G acting smoothly on M by $\Phi : \mathrm{G} \times \mathrm{M} \to \mathrm{M}$, ζ a symplectic form on TM, $H : \mathrm{TM} \to \mathbb{R}$ a smooth function, π a smooth semi-basic Pfaffian form satisfying, $\hat{\pi}(X^{\#}_{\mathrm{TM}}) = 0$, for all $X \in \mathfrak{g}$, where $X^{\#}_{\mathrm{TM}}$ is the infinitesimal generator corresponding to X with respect to the action Φ^T and suppose that the triple (ζ, H, π) defines a Birkhoffian ω on M such that

$$\bar{Y}^*\omega = -\mathrm{i}_Y\zeta + \mathsf{d}H - \pi,$$

for all second order vector fields $Y = i \circ \bar{Y}$ on TM. If $(\mathrm{TM}, \zeta, \Phi^T, \mu)$ is a Hamiltonian G-space, that is, Φ^T is symplectic and

$$\mathsf{d}\mu^X = \mathrm{i}_{X^{\#}_{\mathrm{TM}}}\zeta$$

for all $X \in \mathfrak{g}$; and H is Φ^T-invariant, then μ is a momentum mapping for the Birkhoff system (M, ω). Indeed, for each $X \in \mathfrak{g}$ and for each $v \in \mathrm{TM}$, we have $H(\Phi^T(\exp(tX), v)) = H(v)$ since H is Φ^T-invariant. Differentiating at $t = 0$, $\mathsf{d}_v H \cdot X^{\#}_{\mathrm{TM}}(v) = 0$; so $\mathsf{d}H(X^{\#}) = 0$ for all $X \in \mathfrak{g}$. Thus, $\mathsf{d}\mu^X(Y) = \mathrm{i}_{X^{\#}_{\mathrm{TM}}}\zeta(Y) = -\mathrm{i}_Y\zeta(X^{\#}_{\mathrm{TM}}) + \mathsf{d}H(X^{\#}_{\mathrm{TM}}) - \pi(X^{\#}_{\mathrm{TM}}) = \bar{Y}^*\omega(X^{\#}_{\mathrm{TM}})$ for all second order vector field $Y = i \circ \bar{Y}$ and for all $X \in \mathfrak{g}$.

In particular, if L is a smooth, regular and Φ^T-invariant Lagrangian on TM, then, $\mu^X(v) = \mathbb{F}\mathrm{L}(v)X^{\#}_{\mathrm{M}}(\tau_{\mathrm{M}}(v))$ for all $X \in \mathfrak{g}$ and for all $v \in \mathrm{TM}$, is a momentum mapping for the action. Indeed, from Corollary 4.2.12 in Abraham & Marsden (1978) we have $(\mathsf{d}_{\mathrm{TM}}\mu^X)Y = -(\mathrm{i}_{X^{\#}}\mathsf{d}_{\mathrm{TM}}\theta_L)Y = -(\mathrm{i}_{X^{\#}_{\mathrm{TM}}}\mathsf{d}_{\mathrm{TM}}\mathsf{d}_v\mathrm{L})Y = (\mathrm{i}_Y\mathsf{d}_{\mathrm{TM}}\mathsf{d}_v\mathrm{L})X^{\#}_{\mathrm{TM}} + \mathsf{d}E(X^{\#}_{\mathrm{TM}}) = (\bar{Y}^*\omega)X^{\#}_{\mathrm{TM}}$, for all second order vector fields $Y = i \circ \bar{Y}$ and for all $X \in \mathfrak{g}$. Here $\theta_L = (\mathbb{F}\mathrm{L})^*\theta_0$, with $\theta_0 : \mathrm{T}^*\mathrm{M} \to \mathrm{T}^*(\mathrm{T}^*\mathrm{M})$ being the canonical Liouville 1-form of $\mathrm{T}^*\mathrm{M}$.

Remark 14. Let L be a smooth, regular and Φ^T-invariant Lagrangian on TM, where $\Phi : \mathrm{G} \times \mathrm{M} \to \mathrm{M}$ is some smooth action. Define $\mu_L : \mathrm{TM} \to \mathfrak{g}^*$ as $\mu^X_L(v) = \mathbb{F}\mathrm{L}(v)X^{\#}_{\mathrm{M}}(\tau_{\mathrm{M}}(v))$. Then, under the Godbillon formalism, the equation of motion of μ_L is

$$\mathsf{d}\mu^X_L(Y) = \hat{\pi}(X^{\#}_{\mathrm{M}}),$$

in other words, it generalizes Appell's Theorem on the projection of the linear momentum (see Appell (1941), p.335) which states that

La derivée par rapport au temps de la projection de la quantité de mouvement sur un axe est égale à la somme des projections sur le même axe des forces appliquées au mobile.

This shows in passing that μ_L is a constant of the motion if and only if $\hat{\pi}(X^{\#}_{\mathrm{M}}) = 0$ for all $X \in \mathfrak{g}$. If $\mathsf{p} : \mathrm{M} \to \mathrm{M}/\mathrm{G}$ is a principal bundle with structural group G, then this condition is equivalent to the force field π being horizontal, that is, $\hat{\pi}(\ker \mathsf{p}) = 0$.

Next we prove that, as in the case of the momentum mapping for symplectic actions on symplectic manifolds, the momentum mapping for an equivariant Birkhoffian is not unique, but defines a unique cohomology class of coadjoint cocycles.

Proposition 14. *Let $(\mathrm{M}, \omega, \Phi, \mu)$ be a Noether G-space and assume that the Birkhoffian ω is equivariant with respect to Φ^{J^2} and Φ^{T^*}, where the latter is $\Phi^{\mathrm{T}^*}_g := \mathrm{T}^*\Phi_{g^{-1}}$, for all $g \in \mathrm{G}$. Define, for $g \in \mathrm{G}$ and $X \in \mathfrak{g}$,*

$$\begin{aligned} \psi_{g,X} &: \mathrm{TM} \to \mathbb{R} \\ &: v \mapsto \mu^X(\Phi^T_g(v)) - \mu^{\mathrm{Ad}_{g^{-1}}X}(v). \end{aligned}$$

Then $\psi_{g,X}$ is constant on TM. *We let $\sigma : G \to \mathfrak{g}^*$ be defined by $\sigma(g)X = \psi_{g,X}(v)$, for any $v \in \mathrm{TM}$, and call it the* coadjoint cocycle *associated to μ. It satisfies the* cocycle identity*: $\sigma(gh) = \sigma(g) + \mathrm{Ad}^*_{g^{-1}}\sigma(h)$.*

Proof. We compute the derivative of $\psi_{g,X}$ at a second order vector field $Y = i \circ \bar{Y}$ on TM

$$\begin{aligned}
\mathsf{d}\psi_{g,X}(Y) &= \mathsf{d}\mu^X(T\Phi_g^T Y) - \mathsf{d}\mu^{\mathrm{Ad}_{g^{-1}} X}(Y) \\
&= ((\Phi_g^{J^2}\bar{Y})^*\omega)X_{\mathrm{TM}}^{\#} - (\bar{Y}^*\omega)(\mathrm{Ad}_{g^{-1}} X)_{\mathrm{TM}}^{\#} \\
&= \hat{\omega}(\Phi_g^{J^2}\bar{Y})X_{\mathrm{M}}^{\#} - \hat{\omega}(\bar{Y})(\mathrm{Ad}_{g^{-1}} X)_{\mathrm{M}}^{\#} \\
&= \hat{\omega}(\Phi_g^{J^2}\bar{Y})X_{\mathrm{M}}^{\#} - \hat{\omega}(\bar{Y})\Phi_g^* X_{\mathrm{M}}^{\#} \\
&= \hat{\omega}(\Phi_g^{J^2}\bar{Y})X_{\mathrm{M}}^{\#} - \left[\Phi_g^{T^*}\hat{\omega}(\bar{Y})\right] X_{\mathrm{M}}^{\#} \\
&= 0
\end{aligned}$$

Thus, by Lemma 3, $\psi_{g,X}$ is constant on any connected component of TM and since TM is connected $\psi_{g,X}$ is a constant on TM.

As for the cocycle identity, we have

$$\begin{aligned}
\sigma(gh) &= \mu^X(\Phi_{gh}^T(v)) - \mu^{\mathrm{Ad}_{gh^{-1}} X}(v) \\
&= \mu^X(\Phi_g^T\Phi_h^T(v)) - \mu^{\mathrm{Ad}_{g^{-1}} X}(\Phi_h^T v) + \\
&\quad\ \mu^{\mathrm{Ad}_{g^{-1}} X}(\Phi_h^T v) - \mu^{\mathrm{Ad}_{h^{-1}}\mathrm{Ad}_{g^{-1}} X}(v) \\
&= \psi_{g,X}(\Phi_h^T(v)) + \psi_{g,\mathrm{Ad}_{g^{-1}} X}(v) \\
&= \sigma(g)X + \sigma(g)\,\mathrm{Ad}_{g^{-1}} X
\end{aligned}$$

□

Proposition 15. *Let* (M,ω) *be a Birkhoff system and* Φ *a smooth action of* G *on* M *such that* ω *is equivariant with respect to* Φ^{J^2} *and* Φ^{T^*}. *If* μ_1 *and* μ_2 *are two momentum mappings with cocycles* σ_1 *and* σ_2, *respectively, then* $[\sigma_1] = [\sigma_2]$. *Thus to any triple* (M,ω,Φ) *satisfying the previous conditions there is a well defined cohomology class* $[\sigma]$.

Proof. Under the hypothesis of the Proposition, we have for each $g \in \mathrm{G}$,

$$\sigma_1(g) - \sigma_2(g) = \mu_1(\Phi_g^T(v)) - \mu_2(\Phi_g^T(v)) - \mathrm{Ad}_{g^{-1}}^*(\mu_1(v) - \mu_2(v))$$

for any $v \in \mathrm{TM}$. But $\mathsf{d}\mu_1^X(Y) = (\bar{Y}^*\omega)X_{\mathrm{TM}}^{\#} = \mathsf{d}\mu_2^X(Y)$, for all second order vector fields Y and all $X \in \mathfrak{g}$, so by Lemma 3, $\mu_1 - \mu_2 = l$, for some constant element $l \in \mathfrak{g}^*$. Then,

$$\sigma_1(g) - \sigma_2(g) = l - \mathrm{Ad}_{g^{-1}}^* l,$$

and $\sigma_1 - \sigma_2$ is a coboundary. □

The reduction of the dynamics of a Birkhoff system (M,ω) with a momentum mapping $\mu : \mathrm{TM} \to \mathfrak{g}^*$ for a smooth action $\Phi : \mathrm{G} \times \mathrm{M} \to \mathrm{M}$ can be effected both in the conservative and non-conservative cases.

Proposition 16. *Let* $(\mathrm{M}, \omega, \Phi, \mu)$ *be a Noether G-space. Assume* ω *regular,* $\hat{\omega}$ *is equivariant with respect to* Φ^{J^2} *and* Φ^{T^*}, $l \in \mathfrak{g}^*$ *is a regular value of* μ *(which by Sard's Theorem takes place for almost all* l*),* μ *is an* Ad^**-equivariant momentum map and the isotropy subgroup of the coadjoint action* $G_l = \{g \in \mathrm{G} \mid \mathrm{Ad}^*_{g^{-1}} l = l\}$ *acts freely and properly on* $\mu^{-1}(l)$*, then, there exists a unique vector field* Y_l *on* $\mu^{-1}(l)/\mathrm{G}_l$ *which is* π_l*-related to the Birkhoff vector field* Y*, that is,* $\mathrm{T}\pi_l Y = Y_l \circ \pi_l$*, where* $\pi_l : \mu^{-1}(l) \to \mu^{-1}(l)/\mathrm{G}_l$ *denotes the canonical projection.*

Proof. Under the assumptions of the Proposition, assume there exists a vector field Y_l on $\mu^{-1}(l)/\mathrm{G}_l$ which is π_l-related to the Birkhoff vector field Y, that is,

$$\mathrm{T}\pi_l Y = Y_l \circ \pi_l, \tag{41}$$

then, uniqueness of Y_l is a consequence of π_l being a submersion.

Now, equivariance of $\hat{\omega}$ with respect to Φ^{J^2} and Φ^{T^*} means

$$\hat{\omega}(\Phi_g^{J^2} z)\eta = \hat{\omega}(z)\mathrm{T}\Phi_{g^{-1}}\eta$$

for all $z \in J^2(\mathrm{M})$, all $\eta \in \mathrm{T}_{\beta(z)}\mathrm{M}$ and all $g \in \mathrm{G}$. So, if $z \in D(\omega)$ then $\hat{\omega}(\Phi_g^{J^2} z)\eta = 0$ for all $\eta \in \mathrm{T}_{\beta(z)}\mathrm{M}$, so that $\Phi_g^{J^2} z \in D(\omega)$. But $z \in D(\omega) \Leftrightarrow z = Y(v)$ for some $v \in \mathrm{TM}$, where Y denotes the Birkhoff vector field; thus, by uniqueness of Y we obtain

$$\Phi_g^{J^2}(Y(v)) = Y(\Phi_g^T v)$$

for all $g \in \mathrm{G}$. In particular, if $l \in \mathfrak{g}$ is a regular value of the momentum mapping μ, then, for $v \in \mu^{-1}(l)$ we have $Y(v) \in \mathrm{T}(\mu^{-1}(l))$ (by Proposition 13) and

$$(\mathrm{T}\Phi_g^T)Y(v) = Y(\Phi_g^T v)$$

for all $g \in G_l$. The previous equation implies that it is well defined a flow H_t on $\mu^{-1}(l)/\mathrm{G}_l$ (note that $Y|_{\mu^{-1}(l)}$ is Φ_g^T-related with itself for all $g \in G_l$) satisfying

$$\pi_l \circ F_t = H_t \circ \pi_l. \tag{42}$$

Define Y_l as the generator of H_t, then, differentiating (42) we obtain (41) and this completes the proof. □

Proposition 17. *Let* μ *be a momentum mapping for a smooth action* Φ*, with cocycle* σ*. Then:*

1. *the map* $\Psi : (g, l) \mapsto \mathrm{Ad}^*_{g^{-1}} l + \sigma(l)$ *is an action of* G *on* $\mathfrak{g}^*$*;*

2. μ *is equivariant with respect to the action in 1.*

Proof. (1) follows from the cocycle identity and (2) from the definition of σ and Ψ. □

From the previous proposition, we can perform a reduction in the case μ is not required to be Ad^*-equivariant. The proof is analogous to that of Proposition 16

Proposition 18. *Let* $(\mathrm{M}, \omega, \Phi, \mu)$ *be a Noether G-space. Assume* ω *regular,* $\hat{\omega}$ *is equivariant with respect to* Φ^{J^2} *and* Φ^{T^*}, $l \in \mathfrak{g}^*$ *is a regular value of* μ *(which by Sard's Theorem takes place for almost all* l*), and the isotropy subgroup of the coadjoint action* $G_l = \{g \in \mathrm{G} \mid \Psi_g l = l\}$ *acts freely and properly on* $\mu^{-1}(l)$*, then, there exists a unique vector field* Y_l *on* $\mu^{-1}(l)/\mathrm{G}_l$ *which is* π_l*-related to the Birkhoff vector field* Y*, that is,* $\mathrm{T}\pi_l Y = Y_l \circ \pi_l$*, where* $\pi_l : \mu^{-1}(l) \to \mu^{-1}(l)/\mathrm{G}_l$ *denotes the canonical projection.*

Without a momentum mapping, a similar computation shows that,

Proposition 19. *Let* (M, ω) *be a regular Birkhoff system and* $\Phi : \mathrm{G} \times \mathrm{M} \to \mathrm{M}$ *a smooth action of a Lie group* G *on the configuration space* M*. Assume* $\hat{\omega}$ *is equivariant with respect to* Φ^{J^2} *and* Φ^{T^*}*, then, there exists a unique vector field* Y_G *on* TM/G *which is* π*-related to the Birkhoff vector field* Y*, that is,* $\mathrm{T}\pi Y = Y_G \circ \pi$*, where* $\pi : \mathrm{TM} \to \mathrm{TM}/\mathrm{G}$ *denotes the canonical projection.*

Example 9. Consider the Birkhoff system associated to the Godbillon formulation (see Remark 5). If T is regular and Φ^T-invariant, $\pi(X^{\#}_{\mathrm{TM}}) = 0$, for all $X \in \mathfrak{g}$ and the hypotheses on the action hold, then we can reduce the dynamics to $\mu^{-1}(l)/\mathrm{G}_l$, for any regular value $l \in \mathfrak{g}$ of the momentum map $\mu^X(v) = \mathbb{F}T(v) \cdot X^{\#}_{\mathrm{M}}(\tau_{\mathrm{M}}(v))$, for all $v \in \mathrm{TM}$ and for all $X \in \mathfrak{g}$.

As a concrete example, consider the rigid body model of an artificial satellite orbiting around the Earth under the influence of the gravity and the drag only. The equations of motion can be cast in the Godbillon formalism, with the configuration space SE(3), kinetic energy $T = \frac{1}{2}\left((\dot{R}, \dot{x}), (\dot{R}, \dot{x})\right)_{\mathrm{SE}(3)}$ where $x \in \mathbb{R}^3$ is the vector from the center of the Earth to the center of mass of the satellite, $R \in \mathrm{SO}(3)$ is the proper orthogonal matrix corresponding to the attitude of the satellite and $(\cdot, \cdot)_{\mathrm{SE}(3)}$ denotes the usual metric on SE(3) (see Oliva (2002)), that is,

$$((s, u), (\bar{s}, \bar{u}))_{\mathrm{SE}(3)} = m(u, u)_g + \int_{\mathcal{B}} (s\xi, \bar{s}\xi)_g \mathrm{d}m(\xi)$$

for all $(s, u), (\bar{s}, \bar{u}) \in \mathrm{TSE}(3)$. Here $(\cdot, \cdot)_g$ is the usual metric on $\mathbb{R}^3$, $m \in \mathbb{R}_+$ is the mass of the satellite, $\mathcal{B} \subset \mathbb{R}^3$ is the body (satellite) and $\xi \in \mathcal{B}$ denotes the position in the body frame of a material particle of the body with respect to its center of mass. It is convenient to define the body angular velocity $\hat{\Omega} = R^{-1}\dot{R}$ where $\hat{}: \mathbb{R}^3 \to \mathfrak{so}(3)$ is the usual identification of $\mathbb{R}^3$ with $\mathfrak{so}(3)$, that is,

$$\hat{\xi} = \begin{bmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{bmatrix}$$

for all $\xi = (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3$. So,

$$T = \frac{1}{2}m(\dot{x}, \dot{x})_g + \frac{1}{2}(I\Omega, \Omega)_g$$

where I is the moment of inertia tensor of the body in the body frame.

As for the force field, again we use the identification of semi-basic forms and force fields to write

$$\hat{\pi} = \mathrm{d}_{\mathrm{TSE}(3)}\widehat{(V \circ \tau_{\mathrm{SE}(3)})} + D$$

where V is the potential energy of the body

$$V = -\int_{\mathcal{B}} \frac{GM}{|x + R\xi|} \mathrm{d}m(\xi)$$

$|\cdot|$ is the Euclidean norm, G is the universal gravitational constant, M is the mass of the Earth and D is the drag force

$$D = -\frac{1}{2}B^* \rho |\dot{x}| \operatorname{pr}_2^* \dot{x}^\flat$$

where $\flat : \mathbb{R}^3 \to \mathbb{R}^3$ is the index-lowering operator with respect to the usual metric on $\mathbb{R}^3$, pr_2 is the projection $\operatorname{pr}_2 : \mathrm{SE}(3) = \mathrm{SO}(3) \circledS \mathbb{R}^3 \to \mathbb{R}^3$ and B^* is the ballistic coefficient of the satellite.

Since, in general $|x| \gg |\xi|$ it is usual to consider an approximation (see Wang, Krishnaprasad & Maddocks (1991))

$$\begin{aligned} V &= -\int_{\mathcal{B}} \frac{GM}{|x + R\xi|} \mathrm{d}m(\xi) \\ &= -\frac{GM}{|x|} \int_{\mathcal{B}} \left\{ 1 - \frac{(x, R\xi)_g}{|x|^2} - \frac{1}{2}\frac{|\xi|^2}{|x|^2} + \frac{3}{2}\frac{(x, R\xi)_g^2}{|x|^4} + \mathrm{o}\left(|x|^{-2}\right) \right\} \mathrm{d}m(\xi) \\ &\approx -\frac{GMm}{|x|} - \frac{1}{2}\frac{GM \operatorname{tr}(I)}{|x|^3} + \frac{3}{2}\frac{GM}{|x|^5}(IR^{-1}x, R^{-1}x)_g. \end{aligned}$$

Note that the system is dissipative and the translational and attitude dynamics are coupled, so that the reduction of Poisson manifolds is not directly applicable.

We consider the body $\mathcal{B}$ to be a symmetry top, that is, two of its principal momentum of inertia are equal. In this case, we can define a right S^1-action $\Phi : S^1 \times \mathrm{SE}(3) \to \mathrm{SE}(3)$ on $\mathrm{SE}(3)$, as the rotation around the axis of symmetry of the ellipsoid of inertia:

$$\Phi : (S, (R, x)) \mapsto (RS, x)$$

where we identify S^1 with the abelian, connected Lie subgroup of $\mathrm{SO}(3)$ given by $S^1 \cong \{\exp(t\hat{v}) \mid t \in \mathbb{R}\}$, where $v \in \mathbb{R}^3$ is a non-null vector in the direction of the symmetry axis. Note that the symmetry of the ellipsoid of inertia implies that $SIS^{-1} = I$ for all $S \in S^1$.

An easy computation shows that T is Φ^T-invariant and that $\pi(X^{\#}_{\mathrm{SE}(3)}) = 0$. So, the angular momentum μ_v around the symmetry axis is an Ad*-equivariant momentum map for the action, whence a constant of the motion. Also, the action is free and proper, $\hat{\omega}$ is equivariant and because the group is abelian, the adjoint action is trivial. So, given $l \in \mathbb{R}$, the isotropy group for the coadjoint action G_l is the whole group S^1 and the dynamics can be reduced to the submanifold $\mu_v^{-1}(l)/S^1$ for any real number $l \in \mathbb{R}^n$. Note that $\dim(\mu_v^{-1}(l)/S^1) = \dim(\mathrm{SE}(3)) - 2$.

From this example, we conclude that the proposed reduction is an extension of the Routh reduction (see, for example, Marsden & Ratiu (1999)) to non-conservative mechanical systems and to possibly non-abelian groups. This may be relevant to engineering in general, and to control theory, in particular, where non-conservative systems are the norm.

Remark 15. If we know the flow H_t of the reduced Birkhoff vector field Y_l, then we can reconstruct the flow F_t of the Birkhoff vector field Y restricted to $\mu^{-1}(l)$. Essentially, it follows the lines proposed in Abraham & Marsden (1978) for the Hamiltonian case (see pp. 304 and 305, *ibidem*). For the sake of completeness, next we present the latter with the adaptations needed. Let $v_0 \in \mu^{-1}(l)$ and let $\dot{\gamma} : I \to \mu^{-1}(l)$ and $[\dot{\gamma}] : I \to \mu^{-1}(l)/\mathrm{G}_\mu$, for some neighborhood of the origin $0 \in I \subset \mathbb{R}$; be integral curves of Y and Y_l, respectively, with $\dot{\gamma}(0) = v_0$. Now, let $d : I \to \mu^{-1}(l)$ be a smooth curve with $d(0) = v_0$ and $[d] = [\dot{\gamma}]$. Define for each $t \in I$, $g(t) \in G_l$ such that $\dot{\gamma}(t) = \Phi^T_{g(t)} d(t)$. Thus,

$$\begin{aligned} Y(\dot{\gamma}(t)) &= \frac{T\dot{\gamma}}{dt} \\ &= \mathrm{T}\Phi^T_{g(t)}(d(t)) \cdot \frac{Td}{dt}(t) + \mathrm{T}\Phi^T_{g(t)}(d(t)) \cdot \left(\mathrm{T}L_{g(t)^{-1}} \frac{Tg}{dt}(t)\right)^{\#}_{\mathrm{TM}} (d(t)) \end{aligned}$$

for all $t \in I$. Taking into account the Φ^T-invariance of Y yields,

$$Y(d(t)) = \frac{Td}{dt}(t) + \left(\mathrm{T}L_{g(t)^{-1}} \frac{Tg}{dt}(t)\right)^{\#}_{\mathrm{TM}} (d(t))$$

for all $t \in I$.

This is an equation for $g : I \to G_l$ in terms of d. We solve the latter by first solving the next algebraic equation for $\xi(t) \in \mathfrak{g}$

$$\xi^{\#}_{\mathrm{TM}}(d(t)) = Y(d(t)) - \frac{Td}{dt}(t),$$

for all $t \in I$ and then

$$\frac{Tg}{dt} = \mathrm{T}L_g \xi$$

for g.

Finally, the solution $\dot{\gamma}$ sought is $\dot{\gamma} = \Phi_g d$, whereas the curve in the configuration space is $\frac{T\gamma}{dt} = \dot{\gamma}$ (recall that Y is a second order vector field).

Let M be a smooth manifold and $a_0, a_1 \in \mathbb{R}$ be real numbers with $a_0 < a_1$; we denote by $\mathrm{C^k}([a_0, a_1], \mathrm{M})$ the set of all curves $\gamma : [a_0, a_1] \to \mathrm{M}$ of class $\mathrm{C^k}$. Let P and Q be smooth submanifolds of M; one defines the set

$$\mathrm{C^k}([a_0, a_1], \mathrm{M}; P, Q) := \left\{\gamma \in \mathrm{C^k}([a_0, a_1], \mathrm{M}) \mid \gamma(a_0) \in P \ \ \& \ \ \gamma(a_1) \in Q\right\}.$$

$\mathrm{C^k}([a_0, a_1], \mathrm{M}; P, Q)$ can be endowed with a differentiable structure of a smooth Banach manifold (see, for instance, Piccione & Tausk (2001) and Piccione & Tausk (2002)) and its tangent space at $\gamma \in \mathrm{C^k}([a_0, a_1], \mathrm{M}; P, Q)$ can be identified with

$$\mathrm{T}_\gamma \mathrm{C^k}([a_0, a_1], \mathrm{M}; P, Q) \cong \left\{\eta \in \gamma^* \mathrm{TM} \mid \eta(a_0) \in \mathrm{T}_{\gamma(a_0)} P \ \ \& \ \ \eta(a_1) \in \mathrm{T}_{\gamma(a_1)} Q\right\}.$$

Let (M, ω) be a Birkhoff system and $\Phi : \mathrm{G} \times \mathrm{M} \to \mathrm{M}$ a free and proper smooth action of a Lie group G on M; let $G \cdot p_0$ and $G \cdot p_1$ be two orbits, which by the hypothesis of Φ being free and proper are smooth closed submanifolds of M. The action Φ induces a smooth action $\Phi\circ : \mathrm{G} \times \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1) \to \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$ on $\mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$ defined by $(\Phi\circ)_g(\gamma) := (\Phi_g \circ)(\gamma) = \Phi_g \circ \gamma$. Indeed, this follows at once from $\mathrm{C^k}([a_0, a_1], \cdot)$ being a functor, from the category of smooth manifolds to the category of smooth Banach manifolds with smooth maps as morphisms (see, for example, Palais (1968) and Piccione & Tausk (2001)),

$$\begin{aligned} \mathrm{C^k}([a_0, a_1], \cdot) &: \mathrm{M} \mapsto \mathrm{C^k}([a_0, a_1], \mathrm{M}) \\ &: f \mapsto f\circ \end{aligned}$$

and the fact that $G \cdot p_i$, $i = 0, 1$ are orbits by Φ.

Analogously, Φ induces a smooth action $\Phi^{J^2} : \mathrm{G} \times J^2(\mathrm{M}) \to J^2(\mathrm{M})$ defined as $\Phi_g^{J^2} := J^2(\Phi_g)$. That Φ^{J^2} is an action is also a consequence of J^2 being a functor.

Lemma 5. *Let (M, ω_L) be a Birkhoff system for some Lagrangian function* L *and $\Phi : \mathrm{G} \times \mathrm{M} \to \mathrm{M}$ be a free and proper smooth action on* M*. Then, the Lagrangian function* L *is Φ^T-invariant if and only if for all $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$ and for all orbits $G \cdot p_0 \subset \mathrm{M}$ and $G \cdot p_1 \subset \mathrm{M}$, the Lagrangian functional $\mathscr{L}$: $\mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1) \to \mathbb{R}$ is $\Phi\circ$-invariant.*

Proof. Given $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$ and two orbits $G \cdot p_0$, $G \cdot p_1$, the Lagrangian functional being $\Phi\circ$-invariant means

$$\mathscr{L}\left((\Phi\circ)_g \gamma\right) = \mathscr{L}\left(\gamma\right)$$

for all $g \in \mathrm{G}$. That is,

$$\begin{aligned}\int_{a_0}^{a_1} \mathrm{L}\,(\gamma(t))\,\mathrm{d}t &= \int_{a_0}^{a_1} \mathrm{L}\left(\frac{T}{dt}\left((\Phi\circ)_g\gamma\right)(t)\right)\mathrm{d}t\\ &= \int_{a_0}^{a_1} \mathrm{L}\left(\frac{T}{dt}\left(\Phi_g\circ\gamma\right)(t)\right)\mathrm{d}t\\ &= \int_{a_0}^{a_1} \mathrm{L}\left(\left(\mathrm{T}\Phi_g\circ\frac{T\gamma}{dt}\right)(t)\right)\mathrm{d}t\\ &= \int_{a_0}^{a_1} \mathrm{L}\left(\left(\Phi_g^T\circ\frac{T\gamma}{dt}\right)(t)\right)\mathrm{d}t\end{aligned}$$

for all $\gamma \in \mathrm{C}^{\mathrm{k}}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1)$ and all $g \in \mathrm{G}$.

But, this requires that $\mathrm{L} = \mathrm{L}\circ\Phi_g^T$. For, if for a given $g \in \mathrm{G}$ there is a vector $v \in \mathrm{TM}$ such that $\mathrm{L}(v) \neq \mathrm{L}(\Phi_g^T v)$, say $\mathrm{L}(v) > \mathrm{L}(\Phi_g^T v)$, then by continuity, there is a neighborhood $V \subset \mathrm{TM}$ of $v \in V$ such that $\mathrm{L}(u) > \mathrm{L}(\Phi_g^T u)$ for all $u \in V$. Finally, choosing a curve $\mathrm{C}^{\mathrm{k}}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1)$, for some $a_0,a_1 \in \mathbb{R}$, $a_0 < a_1$ and orbits $G\cdot p_0$, $G\cdot p_1$ such that $\frac{T\gamma}{dt}([a_0,a_1]) \subset V$, we obtain $\int_{a_0}^{a_1}\mathrm{L}\,(\gamma(t))\,\mathrm{d}t > \int_{a_0}^{a_1}\mathrm{L}\left(\Phi_g^T\circ\frac{T\gamma}{dt}(t)\right)\mathrm{d}t$, which is a contradiction. The converse is trivial and this completes the proof. □

Given $a_0,a_1 \in \mathbb{R}$, $a_0 < a_1$ and two orbits $G\cdot p_0$, $G\cdot p_1$, we can define the Pfaffian form Ω_L on $\mathrm{C}^{\mathrm{k}}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1)$ as

$$\Omega_L := \mathrm{d}_{\mathrm{C}^{\mathrm{k}}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1)}\mathscr{L}.$$

Proposition 20. *Let $\Phi : \mathrm{G}\times\mathrm{M}\to\mathrm{M}$ be a smooth action on a configuration space M and L a regular Lagrangian function. Given $a_0,a_1 \in \mathbb{R}$, $a_0 < a_1$ and orbits $G\cdot p_0$ and $G\cdot p_1$ we have,*

$$\begin{aligned}\Omega_L(\gamma)\eta = -\int_{a_0}^{a_1}\hat{\omega}_L\left(\frac{T}{dt}\frac{T\gamma}{dt}(t)\right)\eta(t)\mathrm{d}t + \\ \langle\mu_L\left(\left.\frac{T\gamma}{dt}\right|_{t=a_1}\right),X_1\rangle - \langle\mu_L\left(\left.\frac{T\gamma}{dt}\right|_{t=a_0}\right),X_0\rangle\end{aligned}$$

for all $\gamma \in \mathrm{C}^{\mathrm{k}}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1)$ and all $\eta \in \mathrm{T}_\gamma\mathrm{C}^{\mathrm{k}}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1)$, where $\mu_L : \mathrm{TM}\to\mathfrak{g}^*$ is defined as

$$\langle\mu_L(v),X\rangle = \mathbb{F}\mathrm{L}(v)\cdot X_{\mathrm{M}}^{\#}(\tau_{\mathrm{M}}(v)),$$

for all $v \in \mathrm{TM}$ and all $X \in \mathfrak{g}$ and $X_0,X_1 \in \mathfrak{g}$ are the unique vectors such that $X_{\mathrm{M},i}^{\#}(\gamma(a_i)) = \eta(a_i)$, $i = 0,1$.

Proof. Fix $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$ and $G \cdot p_0$ and $G \cdot p_1$. Given a regular Lagrangian function L we have,

$$\begin{aligned}
\mathsf{d}_{\mathrm{C^k}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1)}\mathscr{L}(\gamma)(\eta) &= -\int_{a_0}^{a_1} \hat{\omega}_L\left(\left(\frac{T}{dt}\frac{T\gamma}{dt}\right)(t)\right)\eta(t)\mathrm{d}t+ \\
&\mathbb{F}\mathrm{L}\left(\left.\frac{T\gamma}{dt}\right|_{t=a_1}\right)\eta(a_1) - \mathbb{F}\mathrm{L}\left(\left.\frac{T\gamma}{dt}\right|_{t=a_0}\right)\eta(a_0) \\
&= -\int_{a_0}^{a_1} \hat{\omega}_L\left(\left(\frac{T}{dt}\frac{T\gamma}{dt}\right)(t)\right)\eta(t)\mathrm{d}t+ \\
&\langle \mu_L\left(\left.\frac{T\gamma}{dt}\right|_{t=a_1}\right), X_1\rangle - \langle \mu_L\left(\left.\frac{T\gamma}{dt}\right|_{t=a_0}\right), X_0\rangle
\end{aligned}$$

for all $\gamma \in \mathrm{C^k}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1)$ and all $\eta \in \mathrm{T}_\gamma\mathrm{C^k}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1)$. Here $X_0, X_1 \in \mathfrak{g}$ are the unique vectors such that $X^{\#}_{\mathrm{M},i}(\gamma(a_i)) = \eta(a_i)$, $i = 0, 1$ (recall that $\eta(a_i) \in \mathrm{T}_{\gamma(a_i)}(G\cdot p_i)$). □

Remark 16. Let (M,ω) be a Birkhoff system and $\Phi : \mathrm{G}\times\mathrm{M} \to \mathrm{M}$ a smooth action on M. Then, the action Φ is free if and only if for each $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$ and for each pair of orbits $G\cdot p_0$ and $G\cdot p_1$, the induced action $\Phi\circ : \mathrm{G}\times\mathrm{C^k}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1) \to \mathrm{C^k}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1)$ is also free. Indeed, when $\Phi\circ$ is free, the consideration of constant curves proves that Φ is free too. Conversely, assume Φ is free. Given $\gamma \in \mathrm{C^k}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1)$, let $g \in \mathrm{G}$ be such that $(\Phi\circ)_g\gamma = \gamma$. Then, for any $t \in [a_0,a_1]$, we have $(\Phi\circ)_g(\gamma)(t) = \Phi_g(\gamma(t)) = \gamma(t) \Leftrightarrow g = e$.

Similarly, the action Φ is proper if and only if so is the action $\Phi\circ$.

Proposition 21. *Let (M,ω_L) be a Birkhoff system for some regular Lagrangian function* L *and $\Phi : \mathrm{G}\times\mathrm{M} \to \mathrm{M}$ be a smooth, free and proper action on* M. *If* L *is Φ^T-invariant, then for each $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$ and for each pair of orbits $G\cdot p_0$ and $G\cdot p_1$, there exists a unique reduced Lagrangian functional $\mathscr{L}_G$: $\mathrm{C^k}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1)/\mathrm{G} \to \mathbb{R}$ such that $\mathscr{L}_G \circ \Pi = \mathscr{L}$ and $\Pi^*\Omega_{L_G} = \Omega_L$, where $\Pi : \mathrm{C^k}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1) \to \mathrm{C^k}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1)/\mathrm{G}$ is the canonical projection.*

Proof. By hypothesis, the Lagrangian function L is Φ^T-invariant. Now, by Lemma 5 the Lagrangian functional $\mathscr{L}$ is $\Phi\circ$-invariant, that is, the following diagram is commutative

$$\begin{array}{ccc} C^k([a_0,a_1],M;G\cdot p_0,G\cdot p_1) & \xrightarrow{\mathscr{L}} & \mathbb{R} \\ \downarrow (\Phi\circ)_g & & \downarrow 1 \\ C^k([a_0,a_1],M;G\cdot p_0,G\cdot p_1) & \xrightarrow[\mathscr{L}]{} & \mathbb{R} \end{array}$$

So, $\mathscr{L}$ is equivariant with respect to the $\Phi\circ$ and identity actions and this defines a smooth function $\mathscr{L}_G$ on $C^k(S^1,M)/G$, which makes the following diagram commutative

$$\begin{array}{ccc} C^k([a_0,a_1],M;G\cdot p_0,G\cdot p_1) & \xrightarrow{\mathscr{L}} & \mathbb{R} \\ \downarrow \Pi & & \downarrow 1 \\ C^k([a_0,a_1],M;G\cdot p_0,G\cdot p_1)/G & \xrightarrow[\mathscr{L}_G]{} & \mathbb{R}/1=\mathbb{R} \end{array}$$

Finally,

$$\begin{aligned} \Pi^*\Omega_{L_G} &= \Pi^*\mathsf{d}_{C^k([a_0,a_1],M;G\cdot p_0,G\cdot p_1)/G}\mathscr{L}_G \\ &= \mathsf{d}_{C^k([a_0,a_1],M;G\cdot p_0,G\cdot p_1)}(\mathscr{L}_G\circ\Pi) \\ &= \mathsf{d}_{C^k([a_0,a_1],M;G\cdot p_0,G\cdot p_1)}\mathscr{L} = \Omega_L. \end{aligned}$$

□

Proposition 22. *Let* L *be a regular Lagrangian function and* $\Phi : G\times M \to M$ *a free and proper smooth action. Then,* L *is* Φ^T*-invariant if and only if for each* $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$ *and for each pair of orbits* $G\cdot p_0$ *and* $G\cdot p_1$, Ω_L *is* $\Phi\circ$-invariant, *that is,* $(\Phi\circ)_g^*\Omega = \Omega$, *for all* $g\in G$, *and* Ω_L *is* horizontal, *that is,* $\Omega_L(\ker T\Pi) = 0$.

Proof. When L is Φ^T invariant, the result follows from Lemma 5 and the previous Proposition. Conversely, let us assume that for each $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$, and for each pair of orbits $G\cdot p_0$ and $G\cdot p_1$, Ω_L is $\Phi\circ$-invariant and horizontal. By Lemma 5 it is only necessary to show that $\mathscr{L}$ is $\Phi\circ$ invariant.

$\Phi\circ$-invariance means $\mathsf{d}_{C^k([a_0,a_1],M;G\cdot p_0,G\cdot p_1)}\mathscr{L}$
$= \mathsf{d}_{C^k([a_0,a_1],M;G\cdot p_0,G\cdot p_1)}(\mathscr{L}\circ(\Phi\circ)_g)$ for all $g\in G$. So, there is a function $\mathsf{h} : G \to \mathbb{R}$

such that $\mathscr{L}(\gamma) = \mathscr{L} \circ (\Phi\circ)_g(\gamma) + \mathsf{h}(g)$ for all $\gamma \in \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$ and for all $g \in \mathrm{G}$. It is easy to see that h is a smooth homomorphism between G and $(\mathbb{R}, +)$. Indeed, smoothness follows at once from the smoothness of the action $\Phi\circ$. Now, fix a curve $\gamma \in \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$ to obtain

$$\mathsf{h}(e) = \mathscr{L}(\gamma) - \mathscr{L} \circ (\Phi\circ)_e(\gamma) = 0,$$

where $e \in \mathrm{G}$ is the identity, and

$$\begin{aligned} \mathscr{L}(\gamma) &= \mathscr{L} \circ (\Phi\circ)_{gh}(\gamma) + \mathsf{h}(gh) \\ &= \mathscr{L} \circ (\Phi\circ)_g(\gamma) - \mathsf{h}(h) + \mathsf{h}(gh) \\ &= \mathscr{L}(\gamma) - \mathsf{h}(g) - \mathsf{h}(h) + \mathsf{h}(gh) \end{aligned}$$

for all $g, h \in \mathrm{G}$.

Because Φ is free and proper so is $\Phi\circ$. The latter implies that $\Pi : \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1) \to \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)/\mathrm{G}$ is a principal bundle. Thus, for any curve $\gamma \in \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$ the orbit $G \cdot \gamma$ is a closed smooth submanifold of $\mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$ and its tangent space at γ is $\mathrm{T}_\gamma(G \cdot \gamma) = \ker \mathrm{T}_\gamma \Pi$.

The Pfaffian Ω_L being horizontal means

$$\mathsf{d}_{\mathrm{C^k}([a_0,a_1],\mathrm{M};G\cdot p_0,G\cdot p_1)} \mathscr{L}(\eta) = 0$$

for all $\eta \in \mathrm{T}_\gamma(G \cdot \gamma)$. So, given $g_0 \in \mathrm{G}$ and $X \in \mathrm{T}_{g_0}\mathrm{G}$ let $g :]-\epsilon, \epsilon[\to \mathrm{G}$ be a curve in G such that $g(0) = g_0$ and $\left.\frac{Tg}{ds}\right|_{s=0} = X$. Consider $\eta = \left.\frac{T}{ds}(\Phi\circ)_{g(s)}\gamma\right|_{s=0} \in \mathrm{T}_\gamma(G \cdot \gamma)$, we obtain

$$\begin{aligned} 0 &= \left.\frac{T}{ds}\mathscr{L}\left[(\Phi\circ)_{g(s)}\gamma\right]\right|_{s=0} \\ &= \left.\frac{T}{ds}\left[\mathscr{L}(\gamma) - \mathsf{h}(g(s))\right]\right|_{s=0} \\ &= -\mathrm{T}_{g_0}\mathsf{h}X. \end{aligned}$$

Whence $\mathsf{h}(\mathrm{G}) = \{0\}$, since Im h is a subgroup of $\mathbb{R}$. □

Let (M, ω) be a Birkhoff system and $\Phi : \mathrm{G} \times \mathrm{M} \to \mathrm{M}$ a smooth, free and proper action of a smooth Lie group G on the configuration space M. Given a smooth mapping $\mu : \mathrm{TM} \to \mathfrak{g}^*$, for each $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$, and each pair of orbits $G \cdot p_0$ and $G \cdot p_1$, we define the Pfaffian form Ω for the data $(\mathrm{M}, \omega, \Phi, \mu)$ on $\mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$ as

$$\begin{aligned} \Omega_\gamma(\eta) := -\int_{a_0}^{a_1} \hat{\omega}\left(\left(\frac{T}{dt}\frac{T\gamma}{dt}\right)(t)\right)\eta(t)\mathrm{d}t + \\ \langle\mu\left(\left.\frac{T\gamma}{dt}\right|_{t=a_1}\right), X_1\rangle - \langle\mu\left(\left.\frac{T\gamma}{dt}\right|_{t=a_0}\right), X_0\rangle \end{aligned}$$

for all $\gamma \in \mathrm{C}^{\mathrm{k}}([a_0, a_1], \mathrm{M}; G\cdot p_0, G\cdot p_1)$ and for all $\eta \in \mathrm{T}_\gamma \mathrm{C}^{\mathrm{k}}([a_0, a_1], \mathrm{M}; G\cdot p_0, G\cdot p_1)$, where $\mu^X(v) = \langle \mu(v), X\rangle$, for all $v \in \mathrm{TM}$ and for all $X \in \mathfrak{g}$.

Theorem 8. *Let* (M, ω) *be a Birkhoff system and* $\Phi : \mathrm{G} \times \mathrm{M} \to \mathrm{M}$ *a smooth, free and proper action of a smooth Lie group* G *on the configuration space* M. *Also, let* $\mu : \mathrm{TM} \to \mathfrak{g}^*$ *be a given smooth mapping. Then, the following are equivalent:*

1. $(\mathrm{M}, \omega, \Phi, \mu)$ *is a Noether G-space;*
2. *For any* $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$ *and any pair of orbits* $G \cdot p_0$ *and* $G \cdot p_1$, *the Pfaffian form* Ω *is horizontal.*

Proof. Given $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$ and a pair of orbits $G \cdot p_0$ and $G \cdot p_1$, the requirement of Ω being horizontal is that

$$\begin{aligned} 0 &= \Omega_\gamma(\eta) \\ &= -\int_{a_0}^{a_1} \hat{\omega}\left(\left(\frac{T}{dt}\frac{T\gamma}{dt}\right)(t)\right)\eta(t)\mathrm{d}t + \\ &\quad \langle \mu\left(\left.\frac{T\gamma}{dt}\right|_{t=a_1}\right), X_1\rangle - \langle \mu\left(\left.\frac{T\gamma}{dt}\right|_{t=a_0}\right), X_0\rangle \end{aligned}$$

for all $\gamma \in \mathrm{C}^{\mathrm{k}}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$ and for all $\eta \in \mathrm{T}_\gamma(G \cdot \gamma)$, where $X_i \in \mathfrak{g}$ such that $\eta(a_i) = X^{\#}_{\mathrm{M},i}(\gamma(a_i))$, $i = 0, 1$.

Claim: $\eta \in \mathrm{T}_\gamma(G \cdot \gamma)$ if and only if $\eta = X^{\#}_{\mathrm{M}} \circ \gamma$, for some $X \in \mathfrak{g}$.

Indeed, given such an $\eta \in \mathrm{T}_\gamma(G \cdot \gamma)$ we have

$$\begin{aligned} \eta(t) &= \left(\left.\frac{T}{ds}\right|_{s=0} (\Phi\circ)_{\exp(sX)}(\gamma)\right)(t) \\ &= \left(\left.\frac{T}{ds}\right|_{s=0} \Phi_{\exp(sX)}\gamma\right)(t) \\ &= \left.\frac{T}{\partial s}\right|_{s=0} \Phi(\exp(sX), \gamma(t)) \\ &= X^{\#}_{\mathrm{M}}(\gamma(t)). \end{aligned}$$

for all $t \in [a_0, a_1]$.

Thus, the condition for horizontality is equivalent to

$$\int_{a_0}^{a_1} \hat{\omega}\left(\left(\frac{T}{dt}\frac{T\gamma}{dt}\right)(t)\right)\eta(t)\mathrm{d}t = \langle \mu\left(\left.\frac{T\gamma}{dt}\right|_{t=a_1}\right) - \mu\left(\left.\frac{T\gamma}{dt}\right|_{t=a_0}\right), X\rangle \tag{43}$$

for all $\gamma \in \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$ and all $\eta \in \mathrm{T}_\gamma(G \cdot \gamma)$, with $\eta = X^{\#}_{\mathrm{M}} \circ \gamma$.

Now, assume that $(\mathrm{M}, \omega, \Phi, \mu)$ is a Noether G-space. Then,

$$\hat{\omega}(z)(X^{\#}_{\mathrm{M}}(p)) = (\bar{Y}^*\omega)X^{\#}_{\mathrm{TM}}(v) = \mathsf{d}\mu^X(i \circ z)$$

for all $z \in J^2(\mathrm{M})$, where $p = \beta(z)$, $v = \tau_J(z)$ and $Y = i \circ \bar{Y}$ is a germ of second order vector field such that $Y(v) = z$. Here we have used $\mathrm{T}\tau_{\mathrm{M}} \circ X^{\#}_{\mathrm{TM}} = X^{\#}_{\mathrm{M}} \circ \tau_{\mathrm{M}}$ (use the fact that $\tau_{\mathrm{M}} : \mathrm{TM} \to \mathrm{M}$ is equivariant and Proposition 4.1.28 in Abraham & Marsden (1978)) and the fact that ω is a Birkhoffian, so that $(\bar{Y}^*\omega)X^{\#}_{\mathrm{TM}}(v) = \hat{\omega}(z)(\mathrm{T}\tau_{\mathrm{M}} X^{\#}_{\mathrm{TM}}(v))$ (see Remark 1). Thus,

$$\hat{\omega}(j^2\gamma)\eta = \hat{\omega}(j^2\gamma)(X^{\#}_{\mathrm{M}} \circ \gamma) = \mathsf{d}\mu^X\left(\frac{T}{dt}\frac{T\gamma}{dt}\right) = \frac{d}{dt}\langle \mu\left(\frac{T\gamma}{dt}\right), X\rangle$$

for all $\gamma \in \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$ and all $\eta \in \mathrm{T}_\gamma(G \cdot \gamma)$, where $\eta = X^{\#}_{\mathrm{M}} \circ \gamma$ and $\frac{T}{dt}\frac{T\gamma}{dt} = i \circ j^2\gamma$. Thus, identity (43) holds for all $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$, for all orbits $G \cdot p_0$ and $G \cdot p_1$, for all $\gamma \in \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$ and all $\eta \in \mathrm{T}_\gamma(G \cdot \gamma)$ and Ω is horizontal.

Conversely, let us assume that Ω is horizontal. Suppose, by contradiction, that there is a second order vector field $Y = i \circ \bar{Y}$, a vector $X \in \mathfrak{g}$ and $v \in \mathrm{TM}$ such that $(\bar{Y}^*\omega)_v X^{\#}_{\mathrm{TM}}(v) \neq \mathsf{d}_v\mu^X(Y_v)$ say $(\bar{Y}^*\omega)_v X^{\#}_{\mathrm{TM}}(v) > \mathsf{d}_v\mu^X(Y_v)$. By continuity, there exists a neighborhood $W \subset \mathrm{TM}$ of $v \in W$ such that $(\bar{Y}^*\omega)_u X^{\#}_{\mathrm{TM}}(u) > \mathsf{d}_u\mu^X(Y_u)$, for all $u \in W$. Consider a base curve γ of Y and the vector field $\eta = X^{\#}_{\mathrm{M}} \circ \gamma$, in an interval $[a_0, a_1]$ such that $\frac{T\gamma}{dt}([a_0, a_1]) \subset W$, then

$$\int_{a_0}^{a_1} \hat{\omega}\left(\left(\frac{T}{dt}\frac{T\gamma}{dt}\right)(t)\right)\eta(t)\mathrm{d}t > \langle \mu\left(\left.\frac{T\gamma}{dt}\right|_{t=a_1}\right) - \mu\left(\left.\frac{T\gamma}{dt}\right|_{t=a_0}\right), X\rangle$$

which is a contradiction with the hypothesis of Ω being horizontal. So, μ is a momentum mapping, as required. □

Theorem 9. *Let* (M, ω) *be a Birkhoff system and* $\Phi : \mathrm{G} \times \mathrm{M} \to \mathrm{M}$ *a smooth, free and proper action of a smooth Lie group* G *on the configuration space* M. *Also, let* $\mu : \mathrm{TM} \to \mathfrak{g}^*$ *be a smooth momentum mapping. Then, the following are equivalent:*

1. $\hat{\omega}$ *is equivariant with respect to the actions* Φ^{J^2} *and* Φ^{T^*} *and* μ *is* Ad^*-*equivariant.*

2. *For any* $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$, *and any pair of orbits* $G \cdot p_0$ *and* $G \cdot p_1$, *the Pfaffian form* Ω *is* $\Phi\circ$-*invariant.*

Proof. Given $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$ and a pair of orbits $G \cdot p_0$ and $G \cdot p_1$, the Pfaffian Ω being $\Phi\circ$-invariant means

$$\Omega_\gamma(\eta) = \Omega_{\Phi_g \circ \gamma}(\mathrm{T}(\Phi\circ)_g \eta)$$

for all $\gamma \in \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$, for all $\eta \in \mathrm{T}_\gamma \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$ and for all $g \in \mathrm{G}$. This is equivalent to

$$\begin{aligned}
&-\int_{a_0}^{a_1} \hat{\omega}\left(\left(\frac{T}{dt}\frac{T\gamma}{dt}\right)(t)\right)\eta(t)\mathrm{d}t+ \\
&\qquad \langle \mu\left(\left.\frac{T\gamma}{dt}\right|_{t=a_1}\right), X_1\rangle - \langle \mu\left(\left.\frac{T\gamma}{dt}\right|_{t=a_0}\right), X_0\rangle = \\
&-\int_{a_0}^{a_1} \hat{\omega}\left(\Phi_g^{J^2} \circ \left(\frac{T}{dt}\frac{T\gamma}{dt}\right)(t)\right)\mathrm{T}_{\gamma(t)}\Phi_g \eta(t)\mathrm{d}t+ \\
&\qquad \langle \mu\left(\Phi_g^T \circ \left.\frac{T\gamma}{dt}\right|_{t=a_1}\right), \mathrm{Ad}_g X_1\rangle - \langle \mu\left(\Phi_g^T \circ \left.\frac{T\gamma}{dt}\right|_{t=a_0}\right), \mathrm{Ad}_g X_0\rangle, \qquad (44)
\end{aligned}$$

where we have used $\mathrm{T}(\Phi_g \circ) = (\mathrm{T}\Phi_g \circ)$ (see Piccione & Tausk (2001)) and $(\mathrm{Ad}_g X)^\#_{\mathrm{M}} = \Phi^*_{g^{-1}}(X^\#_{\mathrm{M}})$ (see Proposition 4.1.26 in Abraham & Marsden (1978)).

So, assume the momentum mapping is Ad^*-equivariant, then we obtain

$$\langle \mu\left(\left.\frac{T\gamma}{dt}\right|_{t=a_i}\right), X_i\rangle = \langle \mu\left(\Phi_g^T \left.\frac{T\gamma}{dt}\right|_{t=a_i}\right), \mathrm{Ad}_g X_i\rangle,$$

for $i = 0, 1$ and for all $g \in \mathrm{G}$; so that, the former condition reduces to

$$\int_{a_0}^{a_1} \hat{\omega}\left(\left(\frac{T}{dt}\frac{T\gamma}{dt}\right)(t)\right)\eta(t)\mathrm{d}t = \int_{a_0}^{a_1} \hat{\omega}\left(\Phi^{J^2} \circ \left(\frac{T}{dt}\frac{T\gamma}{dt}\right)(t)\right)\mathrm{T}_{\gamma(t)}\Phi_g\eta(t)\mathrm{d}t \quad (45)$$

for all $\gamma \in \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$, for all $\eta \in \mathrm{T}_\gamma \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$ and for all $g \in \mathrm{G}$.

Now, if $\hat{\omega}$ is equivariant with respect to the actions Φ^{J^2} and Φ^{T^*} then (45) clearly holds.

Conversely, let us assume that for any $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$ and any pair of orbits $G \cdot p_0$ and $G \cdot p_1$, the Pfaffian form Ω is $\Phi\circ$-invariant.

Then, by taking $\eta \in \mathrm{T}_\gamma \mathrm{C^k}([a_0, a_1], \mathrm{M}; G \cdot p_0, G \cdot p_1)$ such that $\eta(a_i) = 0$, $i = 0, 1$ and by an argument analogous to the last part of the proof of Lemma 5 we prove that $\hat{\omega}$ is equivariant with respect to the actions Φ^{J^2} and Φ^{T^*}. Finally, taking $\eta(a_0) = 0$ and an arbitrary $\eta(a_1)$ we obtain the Ad^*-invariance. □

Theorem 10. *Let* (M, ω_L) *be a Birkhoff system for some Lagrangian function* L *and* $\Phi : \mathrm{G} \times \mathrm{M} \to \mathrm{M}$ *a smooth action of a Lie group* G *on the configuration space* M. *Then, the following are equivalent:*

1. L *is* Φ^T*-invariant;*
2. *the Pfaffian form* θ_L *(see Example 8) and the energy* $E_L = Z(\mathrm{L}) - \mathrm{L}$ *are* Φ^T*-invariant;*

3. $(\mathrm{M}, \omega_L, \Phi, \mu_L)$ *is a Noether G-space,* $\hat{\omega}_L$ *is equivariant with respect to the actions* Φ^{J^2} *and* Φ^{T^*} *and* μ_L *is* Ad^**-equivariant.*

Proof. • 1. $\Leftrightarrow$ 2. That 1. $\Rightarrow$ 2. is proved in Corollary 4.2.14 in Abraham & Marsden (1978). Conversely, assume that θ_L and the energy $E_L = Z(\mathrm{L}) - \mathrm{L}$ are Φ^T-invariant. We have,

$$\mathrm{L}(v) = \hat{\theta}_L(v)v - E_L(v)$$

for all $v \in \mathrm{TM}$. Now, $\theta_L = (\Phi_g^T)^*\theta_L$ means

$$\theta_L(v)\eta = \theta_L(\Phi_g^T(v))(\mathrm{T}\Phi_g^T\eta)$$

for all $v \in \mathrm{TM}$, all $\eta \in \mathrm{T}_v(\mathrm{TM})$ and all $g \in \mathrm{G}$. But,

$$\theta_L(v)\eta = \langle \hat{\theta}_L(v), \mathrm{T}\tau_{\mathrm{M}}\eta \rangle,$$

for all $v \in \mathrm{TM}$ and all $\eta \in \mathrm{T}_v(\mathrm{TM})$ (compare Remark 1), so that,

$$\begin{aligned} \langle \hat{\theta}_L(v), \mathrm{T}\tau_{\mathrm{M}}\eta \rangle &= \theta_L(v)\eta \\ &= \theta_L(\Phi_g^T(v))(\mathrm{T}\Phi_g^T\eta) \\ &= \langle \hat{\theta}_L(\Phi_g^T(v)), \mathrm{T}\tau_{\mathrm{M}}\mathrm{T}\Phi_g^T\eta \rangle \\ &= \langle \hat{\theta}_L(\Phi_g^T(v)), \mathrm{T}(\Phi_g \circ \tau_{\mathrm{M}})\eta \rangle \\ &= \langle \hat{\theta}_L(\Phi_g^T(v)), \Phi_g^T\tau_{\mathrm{M}}\eta \rangle \end{aligned}$$

for all $v \in \mathrm{TM}$, all $\eta \in \mathrm{T}_v(\mathrm{TM})$ and all $g \in \mathrm{G}$. Now, τ_{M} is a submersion and so $\hat{\theta}_L(v)v = \hat{\theta}_L(\Phi^T v)\Phi^T v$ and L is Φ^T-invariant.

• 1. $\Leftrightarrow$ 3. That, 1. $\Rightarrow$ 3. follows from Corollary 4.2.12 *ibidem* and Example 8. The converse, that is, 3. $\Rightarrow$ 1. follows from 22–9.

□

6 Symmetry and reduction of constrained Birkhoff systems

In this sub-section, we extend the symmetry and reduction procedures for constrained Birkhoff systems. Because of the generality of Birkhoff systems, this extension turn out to be quite simple.

Definition 24. *Let* $(\mathrm{M}, \omega, \mathscr{C})$ *be a constrained Birkhoff system and* $\Phi : \mathrm{G} \times \mathrm{M} \to \mathrm{M}$ *a smooth action of a Lie group* G *on the configuration space* M. *Denote by* $\mathfrak{g}$ *the Lie algebra of the Lie group* G *and by* $\mathfrak{g}^*$ *its dual. We say that a mapping*

$\mu_{\mathscr{C}} : \mathrm{TM} \to \mathfrak{g}^*$ *is a* constrained momentum mapping *relative to* $(\mathrm{M}, \omega, \mathscr{C})$ *for the action* Φ, *provided that for each* $X \in \mathfrak{g}$,

$$\mathsf{d}\mu_{\mathscr{C}}^X(Y) = (\bar{Y}^*\omega|_{\mathscr{C}})X_{\mathrm{TM}}^{\#}$$

for all second order vector fields $Y = i \circ \bar{Y}$, *with* $\bar{Y}$ *a cross section of* $\mathscr{C}$. *The quintuple* $(\mathrm{M}, \omega, \Phi, \mu_{\mathscr{C}}, \mathscr{C})$ *will be called a* constrained Noether G-space.

Proposition 23. *Let* $(\mathrm{M}, \omega, \Phi, \mu, \mathscr{C})$ *be a given constrained Noether G-space, with* $(\mathrm{M}, \omega, \mathscr{C})$ *a regular constrained mechanical system which satisfies the principle of the determinism. Assume that the action* Φ is tangent to the constraint $\mathscr{C}$, *that is,* $\lambda_v(X_{\mathrm{M}}^{\#}(p)) \subset \mathrm{C}_v$, *for all* $v \in \mathrm{TM}$, *with* $p = \tau_{\mathrm{M}}(v)$. *Then, the constrained momentum mapping* $\mu_{\mathscr{C}}$ *is a constant of the motion, that is,*

$$\mu_{\mathscr{C}} \circ F_t^{\mathscr{C}} = \mu_{\mathscr{C}}$$

where $F_t^{\mathscr{C}} : \mathrm{TM} \to \mathrm{TM}$ *denotes the flow of the d'Alembert-Birkhoff vector field* $Y_{\mathscr{C}}$.

Proof. Let $Y_{\mathscr{C}} = i \circ \bar{Y}_{\mathscr{C}}$ be the d'Alembert-Birkhoff vector field. Then, taking into account that $\hat{\omega}(\bar{Y}_{\mathscr{C}}) \in \mathrm{C}^o_{\tau_J \circ \bar{Y}_{\mathscr{C}}}$ we obtain

$$\begin{aligned} \mathsf{d}\mu_{\mathscr{C}}^X(Y_{\mathscr{C}}) &= (\bar{Y}_{\mathscr{C}}^*\omega|_{\mathscr{C}})X_{\mathrm{TM}}^{\#} \\ &= \hat{\omega}(\bar{Y}_{\mathscr{C}})X_{\mathrm{M}}^{\#} \\ &= 0, \end{aligned}$$

where we have used the fact that $\mathrm{T}\tau_{\mathrm{M}} X_{\mathrm{TM}}^{\#} = X_{\mathrm{M}}^{\#} \circ \tau_{\mathrm{M}}$ and the tangency of Φ: $X_{\mathrm{M}}^{\#} \circ \tau_{\mathrm{M}}(v)\lambda_v^{-1}(\mathrm{C}_v)$. □

Proposition 24. *Let* $(\mathrm{M}, \omega, \mathscr{C})$ *be a constrained Birkhoff system and* $\Phi : \mathrm{G} \times \mathrm{M} \to \mathrm{M}$ *a smooth action of a Lie group* G *on the configuration space* M, *such that* $\pi : \mathrm{M} \to \mathrm{M}/\mathrm{G}$ *is a principal bundle with structural group* G *and the action* Φ *is tangent to the constraint* $\mathscr{C}$. *Assume that the action* Φ^{J^2} *leaves invariant the constraint, that is, that* $\Phi^{J^2}(\mathscr{C}) \subset \mathscr{C}$, $\hat{\omega}|_{\mathscr{C}}$ *is equivariant with respect to* $\Phi^{J^2}|_{\mathscr{C}}$ *and* Φ^{T^*}, *the constrained Birkhoff system* $(\mathrm{M}, \omega, \mathscr{C})$ *satisfies the principle of determinism,* $l \in \mathfrak{g}^*$ *is a regular value of* $\mu_{\mathscr{C}}$, $\mu_{\mathscr{C}}$ *is an* Ad^**-equivariant constrained momentum map and the isotropy subgroup of the coadjoint action* G_l *acts freely and properly on* $\mu_{\mathscr{C}}^{-1}(l)$, *then, there exists a unique vector field* $Y_{\mathscr{C}}^l$ *on* $\mu_{\mathscr{C}}^{-1}(l)/\mathrm{G}_l$ *which is* π_l*-related to the d'Alembert-Birkhoff vector field* $Y_{\mathscr{C}}$, *that is,* $\mathrm{T}\pi_l Y_{\mathscr{C}} = Y_{\mathscr{C}}^l \circ \pi_l$, *where* $\pi_l : \mu_{\mathscr{C}}^{-1}(l) \to \mu_{\mathscr{C}}^{-1}(l)/\mathrm{G}_l$ *denotes the canonical projection.*

Again, without a momentum mapping,

Proposition 25. *Let* (M, ω) *be a regular Birkhoff system which satisfies the principle of reciprocity and* $\Phi : \mathrm{G} \times \mathrm{M} \to \mathrm{M}$ *a smooth action of a Lie group* G *on* *the configuration space* M. *Assume* Φ^{J^2} *leaves* $\mathscr{C}$ *invariant and* $\hat{\omega}|_{\mathscr{C}}$ *is equivariant with respect to* Φ^{J^2} *and* Φ^{T^*}, *then, there exists a unique vector field* $Y_{\mathscr{C}}^G$ *on* TM/G *which is* π*-related to the d'Alembert-Birkhoff vector field* $Y_{\mathscr{C}}$, *that is,* $\mathrm{T}\pi Y_{\mathscr{C}} = Y_{\mathscr{C}}^G \circ \pi$, *where* $\pi : \mathrm{TM} \to \mathrm{TM/G}$ *denotes the canonical projection.*

The proofs are analogous to that of Proposition 16.

Example 10. Consider a particle moving in $\mathrm{M} = \mathbb{R}^3$ with kinetic energy

$$T = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2),$$

where $(x, y, z) \in \mathbb{R}^3$ are normal Cartesian coordinates for $\mathbb{R}^3$. In Rosenberg (1977) it is proposed the following non-holonomic constraint (see also Bloch, Krishnaprasad, Marsden & Murray (1996) and Tavares (2001))

$$\dot{z} - y\dot{x} = 0.$$

We note in passing that the associated form $\epsilon^3 = \mathrm{d}z - y\mathrm{d}x$ is a contact form.

We define the constraint $\mathscr{C}$ as

$$\mathscr{C} = \left\{(x, y, z, \dot{x}, \dot{y}, \dot{z}, \ddot{x}, \ddot{y}, \ddot{z}) :\in J^2(\mathbb{R}^3) \mid \ddot{z} - \dot{y}\dot{x} - y\ddot{x} = 0\right\}.$$

First, we note that the constraint $\mathscr{C}$ is *integrable*, in the sense that the foliation $\mathcal{L} = \{\mathcal{L}_k\}_{k\in\mathbb{R}}$ of $\mathrm{T}(\mathbb{R}^3)$ defined by $\mathcal{L}_k := \left\{(x, y, z, \dot{x}, \dot{y}, \dot{z}) \in \mathrm{T}(\mathbb{R}^3) \mid \dot{z} - y\dot{x} = k\right\}$ is such that, at each pair $v \in \mathrm{T}\mathbb{R}^3$, there is a unique leaf $\mathcal{L}_k$ containing v and satisfying $\mathrm{T}_v\mathcal{L}_k \cap J_v^2(\mathrm{M}) = \mathscr{C}_v$.

The constrained Birkhoff system $(\mathrm{M}, \omega, \mathscr{C})$ is clearly regular (see Remark 11). We consider the $\mathbb{R}$-action

$$\Phi : (\delta, (x, y, z)) \mapsto (x, y + \delta, z).$$

This action is tangent to the constraint and the momentum mapping $\mu_L = p_y$ given by the linear momentum in the y-direction

$$p_y = \dot{y}$$

is a constant of the motion. However, Φ does not leave the constraint invariant, so that we can not use Proposition 24 to reduce the dynamics. This is called an action of the second type in Marle (2002). We note, however, that the constant μ_L is still useful and can be used to integrate the equations. Indeed, we have $y = y_0 + p_y t$, for all $t \in \mathbb{R}$, where $y(0) = y_0$. The remaining equations of motion are

$$\begin{aligned} \ddot{x} &= -\frac{\dot{x}\dot{y}y}{1+y^2} \\ \ddot{z} &= \frac{\dot{x}\dot{y}}{1+y^2}, \end{aligned}$$

yielding

$$\begin{aligned} x &= x_0 + \dot{x}_0 t \\ z &= z_0 + \dot{z}_0 t \end{aligned}$$

if $p_y = 0$ and

$$\begin{aligned} x &= x_0 + \frac{\dot{x}_0\sqrt{1+y_0^2}}{p_y}\left[\sinh^{-1}(y_0+p_y t) - \sinh^{-1}(y_0)\right] \\ z &= \frac{1}{p_y}\left[\dot{x}_0\sqrt{1+y_0^2}\sqrt{1+(y_0+p_y t)^2} + p_y(z_0+\dot{z}_0 t) - \dot{x}_0(1+y_0^2+t y_0 p_y)\right] \end{aligned}$$

otherwise, where $x(0) = x_0$, $\dot{x}(0) = \dot{x}_0$, $z(0) = z_0$ and $\dot{z}(0) = \dot{z}_0$.

The reaction field is

$$R = -\frac{\dot{x}\dot{y}y}{1+y^2}\mathrm{d}x + \frac{\dot{x}\dot{y}}{1+y^2}\mathrm{d}z.$$

Let us check if the flow of the d'Alembert-Birkhoff vector field leads to the conservation of volume. The symplectic form is $\zeta = \mathrm{d}x \wedge \mathrm{d}\dot{x} + \mathrm{d}y \wedge \mathrm{d}\dot{y} + \mathrm{d}z \wedge \mathrm{d}\dot{z}$, so that,

$$\mathsf{d}R \wedge \zeta \wedge \zeta = -\frac{2y\dot{y}}{1+y^2}\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z \wedge \mathrm{d}\dot{x} \wedge \mathrm{d}\dot{y} \wedge \mathrm{d}\dot{z}.$$

By Remark 13, we conclude that the flow of the d'Alembert-Birkhoff vector field $Y_{\mathscr{C}}$ does not preserve the phase volume $\Omega_\zeta = \zeta \wedge \zeta \wedge \zeta$.

The leaves of the above foliation are affine hyperplanes. Indeed, define $e_3 = \frac{1}{1+y^2}\left(\frac{\partial}{\partial z} - y\frac{\partial}{\partial x}\right)$, then

$$v \in \mathcal{L}_k \Leftrightarrow v - k e_3 \in \mathscr{D}$$

where $\mathscr{D} := \ker \epsilon^3$ is the contact distribution.

Now, for $k = 0$, $N = \frac{1}{\sqrt{1+y^2}}\left(\frac{\partial}{\partial z} - y\frac{\partial}{\partial x}\right)$ is a unitary vector field orthogonal to $\mathscr{D}$. Liouville's Theorem for linear non-holonomic constraints (see Kupka & Oliva (2001) and Castro & Oliva (1999)) states that the volume form $\Omega_0|_{\mathcal{L}_0}$ for the leaf $\mathcal{L}_0$ is conserved by the d'Alembert vector field if and only if $\nabla_N N = 0$, where

$$\Omega_0 = \epsilon^1 \wedge \epsilon^2 \wedge \epsilon^3 \wedge \mathrm{d}\bar{\epsilon}^1 \wedge \mathrm{d}\bar{\epsilon}^2,$$

$\epsilon^1 = \mathrm{d}y$, $\epsilon^2 = \frac{\mathrm{d}x + y\mathrm{d}z}{\sqrt{1+y^2}}$ and $\bar{\epsilon}^i$, $i = 1,2,3$ are the corresponding forms view as a function on TM. A simple computation shows that the latter condition is fulfilled and we conclude that the flow of $Y_{\mathscr{C}}|_{\mathcal{L}_0}$ preserves $\Omega_0|_{\mathcal{L}_0}$.

For $k \neq 0$ the non-holonomic constraint is affine and we may use Liouville's Theorem 7 for affine constraints. For that purpose we compute

$$\Theta(k e_3)\epsilon^3 = k\Theta(e_3)\epsilon^3 = k[\mathrm{i}_{e_3}\mathsf{d}\epsilon^3 + \mathsf{d}(\mathrm{i}_{e_3}\epsilon^3)] = k(\mathrm{i}_{e_3}\mathsf{d}\epsilon^3) = -\frac{ky}{1+y^2}\mathrm{d}y,$$

and we conclude that the flow of $Y_{\mathscr{C}}|_{\mathcal{L}_k}$ does not preserve $\Omega_0|_{\mathcal{L}_k}$ for $k \neq 0$.

Example 11. In Benenti (1996) a non-linear constraint is proposed, in which the velocity vectors of two particles moving in a plane are required to remain parallel to each other at all times. The configuration space is $\mathbb{R}^2 \times \mathbb{R}^2$ and the constraint is $v_1 \times v_2 = 0$, where v_1 and v_2 denote the velocity vector of the particles. It is singular when both v_1 and v_2 are null, so we consider a slightly more general and regular non-linear constraint $\mathscr{C}$ defined on the open submanifold $v_1^2 + v_2^2 > 0$ ($v_i^2 = |v_i|^2$, $i = 1, 2$) of $\mathrm{T}(\mathbb{R}^2 \times \mathbb{R}^2)$ given by

$$\begin{aligned}\mathscr{C} := \{ &(x_1, y_1, x_2, y_2, \dot{x}_1, \dot{y}_1, \dot{x}_2, \dot{y}_2, \ddot{x}_1, \ddot{y}_1, \ddot{x}_2, \ddot{y}_2) \in \mathbb{R}^{12} \mid \\ &\dot{x}_1^2 + \dot{y}_1^2 + \dot{x}_2^2 + \dot{y}_2^2 > 0 \quad \& \\ &\ddot{x}_1\dot{y}_2 + \dot{x}_1\ddot{y}_2 - \ddot{x}_2\dot{y}_1 - \dot{x}_2\ddot{y}_1 = 0 \},\end{aligned}$$

where (x_i, y_i), $i = 1, 2$ are normal Cartesian coordinates.

$\mathscr{C}$ is integrable with foliation $\mathcal{L} = \{\mathcal{L}_k\}_{k \in \mathbb{R}}$ of $\mathrm{T}(\mathbb{R}^2 \times \mathbb{R}^2)$ defined by

$$\mathcal{L}_k := \{ (v_1, v_2) \in \mathrm{T}(\mathbb{R}^2 \times \mathbb{R}^2) \mid v_1 \times v_2 = k \}.$$

We consider both particles with equal mass m and electrical charges q_1 and q_2, subjected to a uniform magnetic field

$$B = \begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix}.$$

The equations of motion are then

$$\begin{aligned} m\ddot{x}_1 &= \lambda\dot{y}_2 + q_1\dot{y}_1 b && (46)\\ m\ddot{y}_1 &= -\lambda\dot{x}_2 - q_1\dot{x}_1 b && (47)\\ m\ddot{x}_2 &= -\lambda\dot{y}_1 + q_2\dot{y}_2 b && (48)\\ m\ddot{y}_1 &= \lambda\dot{x}_1 - q_2\dot{x}_2 b, && (49)\end{aligned}$$

and the constraint equation

$$\ddot{x}_1\dot{y}_2 + \dot{x}_1\ddot{y}_2 - \ddot{x}_2\dot{y}_1 - \dot{x}_2\ddot{y}_1 = 0. \tag{50}$$

Now, substituting (46)—(49) into (50) we obtain

$$\lambda = b(q_2 - q_1)\frac{(v_1, v_2)_{\mathbb{R}^2}}{v_1^2 + v_2^2}.$$

Thus, the d'Alembert-Birkhoff vector field is

$$\begin{aligned} Y_{\mathscr{C}}^{x_1} &= \dot{x}_1 \\ Y_{\mathscr{C}}^{y_1} &= \dot{y}_1 \\ Y_{\mathscr{C}}^{x_1} &= \dot{x}_2 \\ Y_{\mathscr{C}}^{y_2} &= \dot{y}_2 \end{aligned}$$

$$\begin{aligned}
Y_{\mathscr{C}}^{\dot{x}_1} &= \frac{b}{m}\left[(q_2-q_1)\frac{\dot{x}_1\dot{x}_2+\dot{y}_1\dot{y}_2}{\dot{x}_1^2+\dot{x}_2^2+\dot{y}_1^2+\dot{y}_2^2}\dot{y}_2+q_1\dot{y}_1\right]\\
Y_{\mathscr{C}}^{\dot{y}_1} &= -\frac{b}{m}\left[(q_2-q_1)\frac{\dot{x}_1\dot{x}_2+\dot{y}_1\dot{y}_2}{\dot{x}_1^2+\dot{x}_2^2+\dot{y}_1^2+\dot{y}_2^2}\dot{x}_2+q_1\dot{x}_1\right]\\
Y_{\mathscr{C}}^{\dot{x}_2} &= -\frac{b}{m}\left[(q_2-q_1)\frac{\dot{x}_1\dot{x}_2+\dot{y}_1\dot{y}_2}{\dot{x}_1^2+\dot{x}_2^2+\dot{y}_1^2+\dot{y}_2^2}\dot{y}_1-q_2\dot{y}_2\right]\\
Y_{\mathscr{C}}^{\dot{y}_2} &= \frac{b}{m}\left[(q_2-q_1)\frac{\dot{x}_1\dot{x}_2+\dot{y}_1\dot{y}_2}{\dot{x}_1^2+\dot{x}_2^2+\dot{y}_1^2+\dot{y}_2^2}\dot{x}_1-q_2\dot{x}_2\right]
\end{aligned}$$

for all $(x_1, y_1, x_2, y_2, \dot{x}_1, \dot{y}_1, \dot{x}_2, \dot{y}_2)$ such that $\dot{x}_1^2+\dot{y}_1^2+\dot{x}_2^2+\dot{y}_2^2>0$.

We remark, in passing, that $Y_{\mathscr{C}}|_{\mathcal{L}_0}$ coincides with the constrained vector field defined by Marle (see Marle (1995)). And from the homogeneity of $\dot{x}_1\dot{y}_2-\dot{x}_2\dot{y}_1=0$, it follows that the Liouville vector field is tangent to $\mathcal{L}_0$ so that the kinetic energy $T=\frac{m}{2}(\dot{x}_1^2+\dot{y}_1^2+\dot{x}_2^2+\dot{y}_2^2)$ is constant along the flow of $Y_{\mathscr{C}}|_{\mathcal{L}_0}$.

The SO(2)-action

$$(R,(r_1,r_2))\mapsto(Rr_1,Rr_2)$$

with $r_i=(x_i,y_i)$, $i=1,2$ leaves both the constraint and the Lagrangian invariant. Yet it is not tangent to the constraint and though Proposition 24 does not apply (this is analogous to an action of the first type defined in Marle (1995)) we can still use Proposition 25. However, next we integrate the system directly.

1. Case I—$b(q_2-q_1)=0$.

   If $b=0$, the base curves of $Y_{\mathscr{C}}$ consist of rectilinear uniform motion of each particle along parallel straight lines determined by the initial conditions of the particles. Otherwise, if $b\neq0$ and $q_1=q_2=q\neq0$, then the problem corresponds to that of two independent charged particles subject to the same uniform magnetic field. Thus, the solution is a harmonic motion around a pair of circles,

$$\begin{aligned}
x_i &= x_{i,0}+\frac{\dot{y}_{i,0}}{qb/m}+\frac{1}{qb/m}\left[\dot{x}_{i,0}\sin(qbt/m)-\dot{y}_{i,0}\cos(qbt/m)\right]\\
y_i &= y_{i,0}-\frac{\dot{x}_{i,0}}{qb/m}+\frac{1}{qb/m}\left[\dot{x}_{i,0}\cos(qbt/m)+\dot{y}_{i,0}\sin(qbt/m)\right]
\end{aligned}$$

   $i=1,2$ for all $t\in\mathbb{R}$, where $x_{i,0}=x_i(0)$, $y_{i,0}=y_i(0)$, $\dot{x}_{i,0}=\dot{x}_i(0)$ and $\dot{y}_{i,0}=\dot{y}_i(0)$.

2. Case II—$b(q_2-q_1)\neq0$.

   We consider first the regular version of the original constraint in Benenti (1996), that is, $v_1\times v_2=0$ and $v_1^2+v_2^2>0$. Suppose $v_1=0$ and $v_2\neq0$, then inspection of the d'Alembert-Birkhoff vector field indicates that the solution consists of the particle 1 fixed, while particle 2 moves around a circle

$$\begin{aligned} x_2 &= x_{2,0} + \frac{\dot{y}_{2,0}}{q_2 b/m} + \frac{1}{q_2 b/m}\left[\dot{x}_{2,0}\sin(q_2 bt/m) - \dot{y}_{2,0}\cos(q_2 bt/m)\right] \\ y_2 &= y_{2,0} - \frac{\dot{x}_{2,0}}{q_2 b/m} + \frac{1}{q_2 b/m}\left[\dot{x}_{2,0}\cos(q_2 bt/m) + \dot{y}_{2,0}\sin(q_2 bt/m)\right] \end{aligned}$$

where $x_{2_0} = x_2(v_2)$, $y_{2_0} = y_2(v_2)$, $\dot{x}_{2_0} = \dot{x}_2(v_2)$ and $y_{2_0} = y_2(v_2)$. The case of $v_1 \neq 0$ and $v_2 = 0$ is analogous. So, assume $v_i \neq 0$, $i = 1, 2$ so that $v_1 \times v_2 = 0$ yields $v_2 = \kappa v_1$, for some $\kappa \in \mathbb{R} \setminus \{0\}$. We look for a solution with a constant κ. Then, $\lambda = \frac{b(q_2 - q_1)\kappa}{1+\kappa^2}$ and we look for a solution of the following linear system of ODE

$$\begin{aligned} \dot{x}_1 &= u_1 \\ \dot{y}_1 &= w_1 \\ \dot{x}_2 &= u_2 \\ \dot{y}_2 &= w_2 \end{aligned}$$

$$\begin{aligned} \dot{u}_1 &= \frac{b(q_2 - q_1)\kappa^2}{m(1+\kappa^2)} w_1 + \frac{q_1 b}{m} w_1 \\ \dot{w}_1 &= -\frac{b(q_2 - q_1)\kappa^2}{m(1+\kappa^2)} u_1 - \frac{q_1 b}{m} u_1 \\ \dot{u}_2 &= -\frac{b(q_2 - q_1)}{m(1+\kappa^2)} w_2 + \frac{q_2 b}{m} w_2 \\ \dot{w}_1 &= \frac{b(q_2 - q_1)}{m(1+\kappa^2)} u_2 - \frac{q_2 b}{m} u_2 \end{aligned}$$

with initial conditions $x_1(0) = x_{1,0}$, $y_1(0) = y_{1,0}$, $x_2(0) = x_{2,0}$, $y_2(0) = y_{2,0}$, $\dot{x}_1(0) = \dot{x}_{1,0}$, $\dot{y}_1(0) = \dot{y}_{1,0}$, $\dot{x}_2(0) = \kappa\dot{x}_{1,0}$ and $\dot{y}_2(0) = \kappa\dot{y}_{1,0}$.

Now, a unique solution for the previous system of ODE clearly exists and again is given by a pair of circles: particle 1 moving along,

$$x_1 = x_{1,0} + \frac{\dot{y}_{1,0}}{b(q_2\kappa^2 + q_1)/m(1+\kappa^2)} + \frac{1}{b(q_2\kappa^2 + q_1)/m(1+\kappa^2)}$$
$$\left[\dot{x}_{1,0}\sin(b(q_2\kappa^2 + q_1)/m(1+\kappa^2)t) - \dot{y}_{1,0}\cos(b(q_2\kappa^2 + q_1)/m(1+\kappa^2)t)\right]$$
$$y_1 = y_{1,0} - \frac{\dot{x}_{1,0}}{b(q_2\kappa^2 + q_1)/m(1+\kappa^2)} + \frac{1}{b(q_2\kappa^2 + q_1)/m(1+\kappa^2)}$$
$$\left[\dot{x}_{1,0}\cos(b(q_2\kappa^2 + q_1)/m(1+\kappa^2)t) + \dot{y}_{1,0}\sin(b(q_2\kappa^2 + q_1)/m(1+\kappa^2)t)\right]$$

for all $t \in \mathbb{R}$, where $x_{1,0} = x_1(v_1)$, $y_{1,0} = y_1(v_1)$, $\dot{x}_{1,0} = \dot{x}_1(v_1)$ and $\dot{y}_{1,0} = \dot{y}_1(v_1)$; and particle 2 along,

$$x_2 = x_{2,0} + \frac{\dot{y}_{2,0}}{b(q_2\kappa^2 + q_1)/m(1+\kappa^2)} + \frac{1}{b(q_2\kappa^2 + q_1)/m(1+\kappa^2)}$$
$$\left[\dot{x}_{2,0}\sin(b(q_2\kappa^2+q_1)/m(1+\kappa^2)t) - \dot{y}_{2,0}\cos(b(q_2\kappa^2+q_1)/m(1+\kappa^2)t)\right]$$
$$y_2 = y_{2,0} - \frac{\dot{x}_{2,0}}{b(q_2\kappa^2 + q_1)/m(1+\kappa^2)} + \frac{1}{b(q_2\kappa^2 + q_1)/m(1+\kappa^2)}$$
$$\left[\dot{x}_{2,0}\cos(b(q_2\kappa^2+q_1)/m(1+\kappa^2)t) + \dot{y}_{2,0}\sin(b(q_2\kappa^2+q_1)/m(1+\kappa^2)t)\right]$$

for all $t \in \mathbb{R}$, where $x_{2,0} = x_2(v_2)$, $y_{2,0} = y_2(v_2)$, $\dot{x}_{2,0} = \kappa\dot{x}_{1,0}$ and $\dot{y}_{2,0} = \kappa\dot{y}_{1,0}$. But, from uniqueness of the flow we conclude that this is the solution of the d'Alembert-Birkhoff vector field.

Finally, for the general case of $v_1 \times v_2 \neq 0$ we have $v_2 = \kappa R_\theta v_1$, where $R_\theta \in \mathrm{SO}(2)$ and motivated by the previous solution, we obtain, by direct substitution, that the solution is again two circles: particle 1 moving around

$$x_1 = x_{1,0} + \frac{\dot{y}_{1,0}}{b(q_2\kappa^2 + q_1)/m(1+\kappa^2)} + \frac{1}{b(q_2\kappa^2 + q_1)/m(1+\kappa^2)}$$
$$\left[\dot{x}_{1,0}\sin(b(q_2\kappa^2+q_1)/m(1+\kappa^2)t) - \dot{y}_{1,0}\cos(b(q_2\kappa^2+q_1)/m(1+\kappa^2)t)\right]$$
$$y_1 = y_{1,0} - \frac{\dot{x}_{1,0}}{b(q_2\kappa^2 + q_1)/m(1+\kappa^2)} + \frac{1}{b(q_2\kappa^2 + q_1)/m(1+\kappa^2)}$$
$$\left[\dot{x}_{1,0}\cos(b(q_2\kappa^2+q_1)/m(1+\kappa^2)t) + \dot{y}_{1,0}\sin(b(q_2\kappa^2+q_1)/m(1+\kappa^2)t)\right]$$

for all $t \in \mathbb{R}$, where $x_{1,0} = x_1(v_1)$,

$y_{1,0} = y_1(v_1)$,

$\dot{x}_{1,0} = \dot{x}_1(v_1)$ and $\dot{y}_{1,0} = \dot{y}_1(v_1)$; and particle 2 along,

$$x_2 = x_{2,0} + \tfrac{\dot{y}_{2,0}}{b(q_2\kappa^2+q_1)/m(1+\kappa^2)} + \tfrac{1}{b(q_2\kappa^2+q_1)/m(1+\kappa^2)}$$

$$\begin{aligned}&\left[\dot{x}_{2,0}\sin(b(q_2\kappa^2+q_1)/m(1+\kappa^2)t+\theta)-\right.\\&\left.\dot{y}_{2,0}\cos(b(q_2\kappa^2+q_1)/m(1+\kappa^2)t+\theta)\right]\end{aligned}$$

$$y_2 = y_{2,0} - \tfrac{\dot{x}_{2,0}}{b(q_2\kappa^2+q_1)/m(1+\kappa^2)} + \tfrac{1}{b(q_2\kappa^2+q_1)/m(1+\kappa^2)}$$

$$\begin{aligned}&\left[\dot{x}_{2,0}\cos(b(q_2\kappa^2+q_1)/m(1+\kappa^2)t+\theta)+\right.\\&\left.\dot{y}_{2,0}\sin(b(q_2\kappa^2+q_1)/m(1+\kappa^2)t+\theta)\right]\end{aligned}$$

for all $t \in \mathbb{R}$, where

$x_{2,0} = x_2(v_2)$,

$y_{2,0} = y_2(v_2)$,

$\dot{x}_{2,0} = \kappa\ (\dot{x}_{1,0}\ \cos\theta - \dot{y}_{1,0}\ \sin\theta)$

and $\dot{y}_{2,0} = \kappa\ (\dot{y}_{1,0}\ \cos\theta + \dot{x}_{1,0}\ \sin\theta)$.

Hence, the flow of the d'Alembert-Birkhoff vector field has been determined.

Example 12. Here we present a class of examples of the so called Čaplygin or Principal Kinematic Case (see Bloch et al. (1996)) to which a reduction can be undertaken in such a way that the associated non-holonomic systems are reduced to the base manifold as an unconstrained system with a force (see also Koiller (1992)).

Let us recall the following definitions and results for semisimple Lie algebras (see Helgason (1978)):

1. A Lie algebra $\mathfrak{g}$ is called *semisimple* if the Killing form $\kappa(X, Y) = \operatorname{tr}(\operatorname{ad} X \operatorname{ad} Y)$ on $\mathfrak{g} \times \mathfrak{g}$ is non-degenerate. An analytical Lie group is *semisimple* if its Lie algebra is semisimple.

2. Let $\mathfrak{g}$ be a Lie algebra. Then $\theta \in \operatorname{Aut}(\mathfrak{g})$ is an involution if $\theta^2 = 1$.

3. If $\mathfrak{g}$ is a real semisimple Lie algebra, then an involution θ on $\mathfrak{g}$ is called a Cartan Involution if the symmetric bilinear form

$$\kappa_\theta(X, Y) = -\kappa(X, \theta Y)$$

is positive definite, where κ is the so called Killing form of $\mathfrak{g}$.

4. Every real semisimple Lie algebra has a Cartan involution. Moreover any two Cartan involutions are conjugate via $\operatorname{Int}(\mathfrak{g})$.

5. Any Cartan involution yields a Cartan Decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$, where

$$\mathfrak{k} = \{X \in \mathfrak{g} \mid \theta(X) = X\},$$
$$\mathfrak{p} = \{X \in \mathfrak{g} \mid \theta(X) = -X\},$$

where $\mathfrak{k}$ is a maximal compactly embedded subalgebra of $\mathfrak{g}$.

6. The following properties hold:

   (a) $[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}, \quad [\mathfrak{k}, \mathfrak{p}] \subset \mathfrak{p}, \quad [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k},$

   (b) $\kappa_\theta(\mathfrak{k}, \mathfrak{p}) = \kappa(\mathfrak{k}, \mathfrak{p}) = 0,$

   (c) $\kappa|_{\mathfrak{k}}$ is negative definite, $\kappa|_{\mathfrak{p}}$ is positive definite.

On a semisimple analytical Lie group G with Lie algebra $\mathfrak{g}$, let us consider the left invariant distribution defined by $\Sigma_e = \mathfrak{p}$ and the left invariant metric associated with an arbitrary metric on $\mathfrak{g}$ such that $\mathfrak{p}$ and $\mathfrak{k}$ are orthogonal, for instance, $(X, Y)_{\mathfrak{g}} = a\kappa_\theta(X, Y)$, for all $X, Y \in \mathfrak{g}$, with $a > 0$.

We consider the left action $\Phi : \mathrm{H} \times \mathrm{G} \to \mathrm{G}$ on G of the connected Lie subgroup H defined by the Lie subalgebra $\mathfrak{k} \subset \mathfrak{g}$. Note that the action is free and proper (H is compact) so that $\pi : \mathrm{G} \to \mathrm{G}/\mathrm{H}$ is a principal bundle.

Then, both the metric and the distribution are invariant with respect to the action Φ. Indeed, this follows at once from H being a Lie subgroup of G and both the metric and the distribution being left invariant. Moreover, because the Lie algebra of H is $\mathfrak{k}$, we also have $\mathrm{T}_p\mathrm{M} = \Sigma_p \oplus \mathrm{T}_p(H \cdot p)$, for all $p \in \mathrm{G}$ as required.

As concrete examples, we mention the so called pseudo-rigid bodies (see Oliva (2002)), whose configuration space is $\mathrm{SL}(n)$. Then, for $\theta(X) = -X^\dagger$, for all $X \in \mathfrak{sl}(n)$, we have $\mathrm{H} = \mathrm{SO}(n)$. This class of problems are also interesting from the viewpoint of its dynamical characteristics. Indeed, in Castro, Kobayashi & Oliva (2001) it is shown that for $\mathrm{M} = \mathrm{SO}(n, 1)$ (which is a Lie subgroup of $\mathrm{SL}(n+1)$) together with the above construction, provides a class of constrained mechanical systems whose Σ-geodesic flow, that is, the flow associated to the d'Alembert vector field, is partially hyperbolic and preserves a measure given by a volume on Σ.

7 Future work

As concluding remarks, we outline some possible lines for future work in the area of Birkhoff systems:

Geometrical aspects We have dealt with the inverse problem of Lagrangian mechanics. A natural continuation of the present work would be the study of the inverse problem in the more general context of the calculus of variations. Also, the whole theory of mechanical systems with constraints that are not regular, in the sense that the dimension of the constraint may change from point to point, is still undone. A possible starting point would be the extension of the present theory to handle constraints defined by algebraic varieties.

Dynamical aspects Finally, it would be interesting to study many dynamical properties of Birkhoff flows. To mention just a few: non-uniform hyperbolic properties (Barreira, Katok & Pesin (2004)), ergodic stability (Pugh & Shub (1997)), viscosity solutions (Bardi & Capuzzo-Dolcetta (1997)) and dynamic convexity (Zampieri (2003)).

Acknowledgements.
The authors would like to express their gratitude to Professor Tudor Ratiu for his interest and references provided and to Professor Charles-Michel Marle for his kind suggestion on the Inverse Problem of the Lagrangian Mechanics. We would also like to acknowledge the many fruitful discussions on the subject with our colleagues E. Dias, P.G. Henriques, R. Loja Fernandes, J. Natário, M. Perlmutter and G. Terra. The second author thanks also to CAMGSD and ISR(IST) and the FCT (Portugal) for the support through the Research Units Pluriannual Funding Program.

A Proof of Lemma 4

The first equation (17) is equivalent to the principle of reciprocity. For the remaining equations (18) and (19), let $(q, \dot{q}, \ddot{q})$ denote a local natural coordinate system for $J^2(\mathrm{M})$ associated to a local chart (U, q) of M and let $(\bar{U}, \bar{q})$ denote another local chart on M such that $\bar{U} \cap U \neq \varnothing$, then

$$\bar{q}^a = \bar{q}^a(q^1, \ldots, q^n),\ \dot{\bar{q}}^a = A^a_i(q)\dot{q}^i,\ \ddot{\bar{q}}^a = A^a_i(q)\ddot{q}^i + \frac{\partial^2 \bar{q}^a}{\partial q^j \partial q^k}\dot{q}^j \dot{q}^k$$

where $A^a_i(q) := \frac{\partial \bar{q}^a}{\partial q^i}(q)$. So,

$$\frac{\partial \dot{\bar{q}}^a}{\partial q^i} = \frac{dA^a_i}{dt},\ \frac{\partial \ddot{\bar{q}}^a}{\partial q^i} = \frac{d^2 A^a_i}{dt^2},\ \frac{\partial \ddot{\bar{q}}^a}{\partial \dot{q}^i} = 2\frac{dA^a_i}{dt}.$$

Also, $Q_i = \frac{\partial \bar{q}^a}{\partial q^i} \bar{Q}_a$ and so,

$$\begin{aligned}
\frac{\partial Q_i}{\partial q^j} &= \left(\frac{\partial \bar{Q}_a}{\partial \bar{q}^b} \frac{\partial \bar{q}^b}{\partial q^j} + \frac{\partial \bar{Q}_a}{\partial \dot{\bar{q}}^b} \frac{\partial \dot{\bar{q}}^b}{\partial q^j} + \frac{\partial \bar{Q}_a}{\partial \ddot{\bar{q}}^b} \frac{\partial \ddot{\bar{q}}^b}{\partial q^j} \right) A_i^a + \bar{Q}_a \frac{\partial A_i^a}{\partial q^j} \\
&= \left(\frac{\partial \bar{Q}_a}{\partial \bar{q}^b} A_j^b + \frac{\partial \bar{Q}_a}{\partial \dot{\bar{q}}^b} \frac{dA_j^b}{dt} + \frac{\partial \bar{Q}_a}{\partial \ddot{\bar{q}}^b} \frac{d^2 A_j^b}{dt^2} \right) A_i^a + \bar{Q}_a \frac{\partial A_i^a}{\partial q^j} \\
\frac{\partial Q_i}{\partial \dot{q}^j} &= \left(\frac{\partial \bar{Q}_a}{\partial \dot{\bar{q}}^b} A_j^b + 2 \frac{\partial \bar{Q}_a}{\partial \ddot{\bar{q}}^b} \frac{dA_j^b}{dt} \right) A_i^a \\
\frac{\partial Q_i}{\partial \ddot{q}^j} &= \frac{\partial \bar{Q}_a}{\partial \ddot{\bar{q}}^b} A_i^a A_j^b.
\end{aligned}$$

Now, taking into account (17), the left hand side of (18) changes as

$$\begin{aligned}
\frac{\partial Q_i}{\partial \dot{q}^j} + \frac{\partial Q_j}{\partial \dot{q}^i} &= \left(\frac{\partial \bar{Q}_a}{\partial \dot{\bar{q}}^b} + \frac{\partial \bar{Q}_b}{\partial \dot{\bar{q}}^a} \right) A_i^a A_j^b + 2 \frac{\partial \bar{Q}_a}{\partial \ddot{\bar{q}}^b} \left(\frac{dA_j^b}{dt} A_i^a + \frac{dA_i^a}{dt} A_j^b \right), \\
&= \left(\frac{\partial \bar{Q}_a}{\partial \dot{\bar{q}}^b} + \frac{\partial \bar{Q}_b}{\partial \dot{\bar{q}}^a} \right) A_i^a A_j^b + 2 \frac{\partial \bar{Q}_a}{\partial \ddot{\bar{q}}^b} \frac{d}{dt} (A_i^a A_j^b),
\end{aligned}$$

whereas the right hand side as

$$\frac{d}{dt} \left(\frac{\partial Q_i}{\partial \ddot{q}^j} + \frac{\partial Q_j}{\partial \ddot{q}^i} \right) = 2 \frac{d}{dt} \frac{\partial Q_i}{\partial \ddot{q}^j} = 2 \left(\frac{d}{dt} \frac{\partial \bar{Q}_a}{\partial \ddot{\bar{q}}^b} \right) A_i^a A_j^b + 2 \frac{\partial \bar{Q}_a}{\partial \ddot{\bar{q}}^b} \frac{d}{dt} (A_i^a A_j^b),$$

which proves that (18) can be globalized under the assumption that ω satisfies the principle of reciprocity. The proof of (19) is analogous but we have to use (18).

References

Abraham, R. & Marsden, J. E. (1978), *Foundations of Mechanics*, second edn, Addison-Wesley Publishing Company, Reading, Massachusetts.

Anderson, I. & Thompson, G. (1992), *The Inverse Problem of the Calculus of Variations for Ordinary Differential Equations*, Vol. 473 of *Memoirs of the American Mathematical Society*, American Mathematical Society, Providence, Rhode Island.

Appell, P. (1900*a*), 'Dévelopments sur une forme nouvelle des équations de la dynamique', *Journal de Math'ematiques Pures et Appliquées. Neuvième Sér.* **6**, 5—40.

Appell, P. (1900*b*), 'Sur une forme générale des équations de la dynamique', *Journal für die Reine und Angewandte Mathematik* **121**, 310—319.

Appell, P. (1941), *Traité de Mécanique Rationelle—Tome I, Statique - Dynamique du point*, sixth edn, Gauthier-Villars, Paris.

Arnold, V. I. (1989), *Mathematical Methods of Classical Mechanics*, Graduate Texts in Mathematics, second edn, Springer Verlag, New York.

Bardi, M. & Capuzzo-Dolcetta, I. (1997), *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Systems & Control: Foundations & Applications, Birkhäuser, Boston.

Barreira, L., Katok, A. & Pesin, Y. (2004), *Nonuniformly Hyperbolic Dynamical Systems*, Encyclopedia of Mathematics and Its Applications, to appear in Cambridge University Press, Cambridge, UK.

Benenti, S. (1996), 'Geometrical aspects of the dynamics of non-holonomic systems', *Rendiconti del Seminario Matematico dell' Univivesitá e del Politecnico di Torino* **54**(3), 203—212.

Birkhoff, G. D. (1927), *Dynamical Systems*, Vol. 9 of *Colloquium Publications*, American Mathematical Society.

Bloch, A., Krishnaprasad, P., Marsden, J. & Murray, R. (1996), 'Nonholonomic mechanical systems and symmetry', *Archive for Rational Mechanics and Analysis* **136**, 21—99.

Cannas da Silva, A. (2001), *Lectures on Symplectic Geometry*, Vol. 1764 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin.

Castro, H. M. A., Kobayashi, M. H. & Oliva, W. M. (2001), 'On partial hyperbolic flows induced by Σ-geodesic flows', *Journal of Differential Equations* **169**, 142—168.

Castro, H. M. A. & Oliva, W. M. (1999), 'Anosov flows induced by partially hyperbolic Σ-geodesic flows', *Resenhas do Instituto de Matemática e Estatística da Universidade de São Paulo* **4**(2), 227—246.

Davydov, A. (1995), 'Local controllabitlity of typical dynamical inequalities on surfaces', *Proceedings of the Steklov Institute of Mathematics* **209**, 73—106.

Gallavotti, G. (1983), *The Elements of Mechanics*, Texts and Monographs in Physics, Springer-Verlag, New York.

Gallavotti, G. & Ruelle, D. (1997), 'SRB states and nonequilibrium statistical mechanics close to equilibrium', *Communication in Mathematical Physics* **190**, 279—285.

Gibbs, J. W. (1879), 'On the fundamental formulae of dynamics', *American Journal of Mathematics* **II**, 49—64.

Godbillon, C. (1969), *Géométrie Différentielle et Mécanique Analytique*, Collection Méthodes, Hermann, Paris.

Goldschmidt, H. L. (1967), 'Integrability criteria for systems of nonlinear partial differential equations', *Journal of Differential Geometry* **1**, 269—307.

Gorni, G. & Zampieri, G. (2000), 'Time reversibility and energy conservation for Lagrangian systems with non-linear nonholonomic constraints', *Reports on Mathematical Physics* **45**(2), 217—227.

Helgason, S. (1978), *Differential Geometry, Lie Groups, and Symmetric Spaces*, Vol. 80 of *Pure and Applied Mathematics*, Academics Press, San Diego.

Hoover, W. (1986), *Molecular Dynamics*, Vol. 258 of *Lecture Notes in Physics*, Springer-Verlag, New York.

Koiller, J. (1992), 'Reduction of some classical non-holonomic systems with symmetry', *Archive for Rational Mechanics and Analysis* **118**, 113—148.

Kupka, I. & Oliva, W. M. (2001), 'The non-holonomic mechanics', *Journal of Differential Equations* **169**, 169—189.

Lang, S. (1995), *Differential and Riemannian Manifolds*, Vol. 160 of *Graduate Texts in Mathematics*, third edn, Springer Verlag, New York.

LeMasurier, M. (2001), 'Singularities of second order implicit differential equations: a geometric approach', *Journal of Dynamical and Control Systems* **7**(2), 277—298.

Lewis, A. D. (1996), 'The geometry of the Gibbs-Appell equations and Gauss's principle of least constraint', *Reports on Mathematical Physics* **38**(1), 11—28.

Marle, C.-M. (1995), 'Reduction of constrained mechanical systems and stability of relative equilibria', *Communications in Mathematical Physics* **174**, 295—318.

Marle, C.-M. (2002), 'On symmetries and constants of motion in Hamiltonian systems with non-holonomic constraints', *preprint Université Pierre et Marie Curie, see also http:\\www.math.jussieu.fr\ ~marle* .

Marsden, J. E. (1992), *Lectures on Mechanics*, Vol. 174 of *London Mathematical Society Lecture Notes Series*, Cambridge University Press, Cambridge.

Marsden, J. E. & Ratiu, T. S. (1999), *Introduction to Mechanics and Symmetry*, Vol. 17 of *Texts in Applied Mathematics*, second edn, Springer-Verlag, New York.

Massa, E. & Pagani, E. (1994), 'Jet bundle geometry, dynamical connections, and the inverse problem of Lagrangian mechanics', *Annales de l'Intitut Henri Poincaré, physique théorique* **61**(1), 17—62.

Oliva, W. M. (1970), 'Lagrangian systems on manifolds, I', *Celestial Mechanics* **1**, 491—511.

Oliva, W. M. (2002), *Geometric Mechanics*, Vol. 1798 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin.

Palais, R. (1968), *The Foundations of Global Non-Linear Analysis*, Benjamin-Cummings, Reading, Mass.

Pars, L. A. (1965), *A Treatise on Analytical Dynamics*, Heinemann Educational Books, London.

Piccione, P. & Tausk, D. V. (2001), 'On the Banach differential structure for sets of maps on non-compact domains', *Journal of Non-Linear Analysis. Series A: Theory and Methods* **46**, 245—265.

Piccione, P. & Tausk, D. V. (2002), 'Lagrangian and Hamiltonian formalism for constrained variational problems', *to appear in Royal Society of Edinburg Proceeding A* .

Pugh, C. & Shub, M. (1997), 'Stably ergodic dynamical systems and partial hyperbolicity', *Journal of Complexity* **13**(1), 125—179.

Rosenberg, R. (1977), *Analytical Dynamics of Discrete Systems*, Plenum Press, New York.

Ruelle, D. (1999), 'Smooth dynamics and new theoretical ideas in nonequilibrium statistical mechanics', *Journal of Statistical Mechanics* **95**, 393—468.

Smale, S. (1970*a*), 'Topology and mechanics I', *Inventiones Mathematicae* **10**, 305—331.

Smale, S. (1970*b*), 'Topology and mechanics II', *Inventiones Mathematicae* **11**, 45—64.

Tavares, J. N. (2001), 'About Cartan geometrization of non holonomic mechanics', *preprint CMUP—2001-22, see also http://www.fc.up.pt/cmup/cmup-pt.html.*

Wang, L.-S., Krishnaprasad, P. & Maddocks, J. (1991), 'Hamiltonian dynamics of a rigid body in a central gravitational field', *Celestial Mechanics and Dynamical Astronomy* **50**, 349—386.

Weber, R. W. (1986), 'Hamiltonian systems with constraints and their meaning in mechanics', *Archive for Rational Mechanics and Analysis* **91**, 309—335.

Wojtkowski, M. P. (2000), 'Magnetic flows and Gaussian thermostats on manifolds of negative curvature', *Fundamenta Mathematicae* **163**, 177—191.

Yosida, K. (1980), *Functional Analysis*, Vol. 123 of *Grundlehren der Mathematischen Wissenschaften*, Springer-Verlag, Berlin.

Zampieri, G. (2003), 'Dynamic convexity for natural thermostatted systems', *to appear in Journal of Differential Equations* .

Marcelo H Kobayashi
Un. of Hawaii at Manoa
Department of Mechanical Engineering
2540 Dole Street- Holmes Hall 302
Honolulu HI 96822
USA
email: marcelok@hawaii.edu

Waldyr M Oliva
Instituto Superior Técnico
Instituto de Sistemas e Robótica
Av. Rovisco Pais 1, 1049-001 Lisbon
Portugal
email: wamoliva@math.ist.utl.pt