R. CUSHMAN
Mathematics Institute, University of Utrecht, 3508TA
Utrecht, The Netherlands
E-mail: cushman@math.uu.nl

J. ŚNIATYCKI
Department of Mathematics and Statistics,
University of Calgary, 2500
University Dr. N.W., Calgary, Alberta, T2N 1N4 Canada
E-mail: sniat@math.ucalgary.ca

NONHOLONOMIC REDUCTION FOR FREE AND PROPER ACTIONS

Received January 09, 2002
DOI: 10.1070/RD2002v007n01ABEH000196

We study a nonholonomically constrained Hamiltonian system with a symmetry group which acts properly and freely on a constraint distribution. We show that the reduced dynamics is described by a generalized distributional Hamiltonian system. The general theory is illustrated by the example of Chaplygin's skate.

1. Introduction

Consider a mechanical Hamiltonian system, with configuration space Q, subjected to a nonholonomic constraint described by a distribution D on Q. The constrained dynamics can be described by a distributional Hamiltonian system (H, ϖ, h), where (H, ϖ) is a symplectic distribution on D and h is the energy function. Motions of the constrained system are integral curves of the distributional Hamiltonian vector field Y_h of h, which is defined as the unique vector field with values in H such that

$$Y_h \lrcorner \, \varpi = \partial_H h.$$

Here $\lrcorner$ denotes the left interior product of vectors and forms, and $\partial_H h$ is the restriction of $\mathrm{d}h$ to H [2].

Let G be the symmetry group of the system which acts freely and properly on the constraint distribution D. Under additional simplifying assumptions, the reduced dynamics is described by a distributional Hamiltonian system $(\overline{H}, \overline{\varpi}, \overline{h})$ on the space $\overline{D}$ of G orbits in D [2]. In this paper, we drop the assumptions made in [2] and show that the reduced dynamics is described by a generalized distributional Hamiltonian system $(\overline{H}, \overline{\varpi}, \overline{h})$ on $\overline{D}$. In other words, $\overline{H}$ is locally spanned by smooth vector fields. We show that $\overline{H}$ is the projection to $\overline{D}$ of a generalized distribution E on D locally spanned by distributional Hamiltonian vector fields of G-invariant functions.

We illustrate the general theory with the example of Chaplygin's skate. Our treatment corrects some errors in [1].

2. General theory

In this section we discuss the distributional approach to the dynamics of nonholonomically constrained systems and the reduction of symmetries for free and proper actions.

Mathematics Subject Classification 70E18, 70E40

2.1. Nonholonomic dynamics

Let Q be the configuration space of the system and k the kinetic energy metric, that is, the kinetic energy corresponding to velocity $v \in TQ$ is $\frac{1}{2}\mathrm{k}(v, v)$.

Let $t \mapsto q(t)$ be a smooth curve in Q which describes a motion of our system. Denote by $\dot{q}(t) \in T_{q(t)}Q$ the velocity at time t and by

$$\ddot{q}(t) = \frac{D\dot{q}(t)}{dt} \tag{2.1}$$

the covariant derivative of $\dot{q}(t)$ with respect to the Levi-Civita connection of the kinetic energy metric k. We interpret $\ddot{q}(t)$ as the acceleration of the motion. A connection is needed to make the acceleration a tangent vector.

Forces are covectors on Q. The Legendre transformation $\mathrm{k}^\flat : TQ \to T^*Q$ corresponding to the kinetic energy metric k is defined as follows. For every u and v in the same fibre of of the tangent bundle projection $\tau_Q : TQ \to Q$,

$$\langle \mathrm{k}^\flat(u) \mid v \rangle = \mathrm{k}(u, v). \tag{2.2}$$

Here $\langle \cdot \mid \cdot \rangle$ denotes the evaluation of forms on vectors. We assume that the system is acted on by a force coming from a known potential V and that an unknown force φ_{constr} constrains the velocity of the system to lie in a distribution D on Q. Under these assumptions, the motion of the constrained system satisfies the conditions

$$\mathrm{k}^\flat(\ddot{q}(t)) = -\,\mathrm{d}V + \varphi_{constr}, \tag{2.3}$$

and

$$\dot{q}(t) \in D. \tag{2.4}$$

Equation (2.3) is Newton's equation of motion and (2.4) is the constraint condition. In addition, we assume that the work of the constraint force on virtual displacements in D vanishes, that is,

$$\langle \varphi_{constr} \mid u \rangle = 0 \quad \text{for all } u \in D. \tag{2.5}$$

In the Hamiltonian formalism, the equations of motion are studied on T^*Q. However, the constraint D is a subbundle of TQ. In order to get the Hamiltonian description of the constrained motion we could use the Legendre transformation $\mathrm{k}^\flat : TQ \to T^*Q$ to map D to a subbundle of T^*Q as in [2]. A computationally simpler alternative is to pull back the symplectic form to TQ. See [5]. We adopt this approach in the present paper.

The *canonical* 1*-form* θ_0 on T^*Q is defined by

$$\langle \theta_0 \mid u \rangle = \langle p \mid T\pi_Q(u) \rangle \quad \text{for all } u \in T_p(T^*Q), \tag{2.6}$$

where $\pi_Q : T^*Q \to Q$ is the cotangent bundle projection. The *canonical symplectic form* on T^*Q is $\omega_0 = -\,\mathrm{d}\theta_0$. Using the Legendre transformation we pull back θ_0 and ω_0 to forms on TQ, obtaining $\theta = (\mathrm{k}^\flat)^*\theta_0$ and $\omega = (\mathrm{k}^\flat)^*\omega_0$, respectively. Since $\mathrm{k}^\flat$ is a diffeomorphism, it follows that ω is a symplectic form on TQ. Moreover, $\omega = -\,\mathrm{d}\theta$.

Let F be a distribution on TQ consisting of vectors in $T(TQ)$ which project to D under the map $T\tau_Q : T(TQ) \to TQ$, that is,

$$F = \{w \in T(TQ) \mid T\tau_Q(w) \in D\}. \tag{2.7}$$

An essential role in the Hamiltonian formulation of the constrained equations of motion is played by the distribution

$$H = F \cap TD \tag{2.8}$$

on the manifold D. Let ϖ be the restriction of the 2-form ω to vectors in H. In other words,

$$\varpi(w_1, w_2) = \omega(w_1, w_2) \tag{2.9}$$

for every pair of vectors w_1, w_2 in the same fibre of H. In [2] it was shown that ϖ is nondegenerate. In other words, (H, ϖ) is a symplectic distribution on D.

For $f \in C^\infty(D)$, we denote by $\partial_H f$ the 1-form on H obtained by restricting $\mathrm{d}f$ to vectors in H, that is,

$$\langle \partial_H f | w \rangle = \langle \mathrm{d}f | w \rangle \quad \text{for all } w \in H. \tag{2.10}$$

The *distributional Hamiltonian vector field* of f, relative to the symplectic distribution (H, ϖ), is the unique vector field Y_f on D with values in H such that

$$Y_f \mathbin{\lrcorner} \varpi = \partial_H f. \tag{2.11}$$

Let h be the energy function on D. For each $u \in D$,

$$h(u) = \tfrac{1}{2}\,\mathrm{k}(u, u) + V(\tau_Q(u)). \tag{2.12}$$

Theorem 1. *Under condition* (2.5)*, a curve $t \mapsto q(t)$ satisfies Newton's equation* (2.3) *with constraint* (2.4) *if and only if it is an integral curve of the distributional Hamiltonian vector field Y_h of the energy function h.*

Proof. See [2].

It follows from theorem 1 that nonholonomic dynamics can be described by a *distributional Hamiltonian system* (H, ϖ, h) on D, where (H, ϖ) is a symplectic distribution on D and $h \in C^\infty(D)$ is the energy function. This formulation of constrained dynamics enables us to employ most of techniques used for Hamiltonian systems. It should be noted, however, that if $h' \in C^\infty(D)$ is another energy function such that $h'(u) = \frac{1}{2}\,\mathrm{k}'(u, u) + V'(\tau_Q(u))$ for all $u \in D$, where $\mathrm{k}' \neq \mathrm{k}$, then integral curves of $Y_{h'}$ need not satisfy the corresponding constrained Newton's equations because the 2-form ϖ is defined in terms of the metric k and not in terms of the metric k$'$.

2.2. Symmetries

Let G be a Lie group which acts freely and properly on D. We assume that its action

$$\Phi : G \times D \to D : (g, u) \mapsto \Phi(g, u)$$

preserves H, ϖ, and h. Denote by $\overline{D} = D/G$ the space of G-orbits in D and by $\rho : D \to \overline{D}$ the G-orbit map. Since the action of G on D is free and proper, D has a structure of a (left) principal G-bundle over $\overline{D}$ with the projection map ρ.

In [2] we introduced the spaces

$$V = \ker T\rho, \tag{2.13}$$

and

$$U = \{w \in H \mid \varpi(w, v) = 0 \quad \forall\, v \in V \cap H\}. \tag{2.14}$$

We proved that the distributional Hamiltonian vector field of a G-invariant function has values in U. For each $u \in D$, the subspace U_u projects to a subspace $T\rho(U_u)$ of $\overline{D}_{\overline{u}}$, where $\overline{u} = \rho(u)$. Since U is G-invariant, it follows that

$$T\rho(U_u) = T\rho(U_{u'}) \quad \text{whenever } \rho(u) = \rho(u').$$

In this way we obtain a subset $\overline{H}$ of $T\overline{D}$ such that for every $\overline{u} \in \overline{D}$

$$\overline{H}_{\overline{u}} = T\rho(U_u), \quad \text{for any } u \in \rho^{-1}(\overline{u}). \tag{2.15}$$

Moreover, the restriction of ϖ to U_u pushes forward to a 2-form $\overline{\varpi}(\overline{u})$ on $\overline{H}_{\overline{u}}$ such that for every $u \in D$ and every $w_1, w_2 \in U_u$

$$\overline{\varpi}(\overline{u})(T\rho(w_1), T\rho(w_2)) = \varpi(u)(w_1, w_2). \tag{2.16}$$

The distributional Hamiltonian vector field Y_h of h pushes forward to a vector field $\overline{Y}_{\overline{h}}$ on $\overline{D}$ describing the reduced evolution. Since h is G-invariant it pushes forward to a function $\overline{h}$ on $\overline{D}$. For every $\overline{u} \in \overline{D}$,

$$\overline{Y}_{\overline{h}}(\overline{u}) \lrcorner \, \overline{\varpi}(\overline{u}) = \partial_{\overline{H}_{\overline{u}}} \overline{h}. \tag{2.17}$$

In [2] we explicitly (or implicitly) made the following additional assumptions: 1) $V \cap H$ is a distribution, 2) U is a distribution, and 3) $\overline{H} = T\rho(U)$ is a distribution. Under these assumptions, $(\overline{H}, \overline{\varpi})$ is a symplectic distribution and the reduced dynamics is given by a distributional Hamiltonian system $(\overline{H}, \overline{\varpi}, \overline{h})$ on $\overline{D}$. Examples in which some of the above assumptions do not hold are discussed in [1]. The aim of this paper is to discuss nonholonomic reduction for free and proper action of the symmetry group without assumptions 1) – 3).

Lemma 1. *For each $\overline{u} \in \overline{D}$, the 2-form $\overline{\varpi}(\overline{u})$ on $\overline{H}_{\overline{u}}$ defined by equation (2.16) is symplectic.*

Proof.
The proof is given in [2]. We repeat it here for the sake of completeness. It is clear from the definition that $\overline{\varpi}(\overline{u})$ is bilinear and antisymmetric. It remains to show that it is nondegenerate. Consider $u \in \rho^{-1}(\overline{u})$. By definition, U_u is the symplectic annihilator of $V_u \cap H_u$ in $(H_u, \varpi(u))$. Hence $U_u \cap (V_u \cap H_u)$ is the kernel of the restriction of $\varpi(u)$ to U_u. Thus the quotient space $U_u/U_u \cap (V_u \cap H_u)$ is symplectic [4]. But $\overline{H}_{\overline{u}}$ is isomorphic to $U_u/U_u \cap (V_u \cap H_u)$ and $\overline{\varpi}(\overline{u})$ is the push forward of the symplectic form on $U_u/U_u \cap (V_u \cap H_u)$ under the isomorphism $U_u/U_u \cap (V_u \cap H_u) \to \overline{H}_{\overline{u}}$. Thus $\overline{\varpi}(\overline{u})$ is symplectic. ■

2.3. Generalized distributions

A generalized distribution on a smooth manifold M is a smooth distribution N in the sense of Sussmann [11]. In other words, N is locally spanned by smooth vector fields. A generalized distribution is a distribution (in the usual sense) if and only if it has constant rank.

Let $C^\infty(D)^G$ be the ring of smooth G-invariant functions on D and let

$$E = \{Y_f(u) \mid f \in C^\infty(D)^G \text{ and } u \in D\}. \tag{2.18}$$

Since H is finite dimensional and the vector fields Y_f have values in H, for each $u \in D$ there exists a finite number k of vectors $Y_{f_1}(u), ..., Y_{f_k}(u)$ which span E_u. The vector fields $Y_{f_1}, ..., Y_{f_k}$ span E in a neighbourhood O of u. Hence, E is a generalized distribution on D, which is locally spanned by the distributional Hamiltonian vector fields of G-invariant functions. From the results of [2] it follows that $E \subseteq U$.

Theorem 2. *If the action of the group G on D is free and proper, then $\overline{H}$ is a generalized distribution locally spanned by the projection to $\overline{D}$ of the distributional Hamiltonian vector fields of G-invariant functions under the tangent to the G-orbit map.*

Proof. For $u \in D$, we have subspaces H_u, V_u, and U_u of T_uD. Since the action of G on D is free and proper, there exists a G-invariant metric g on D. Let A_u be the g-orthogonal complement of U_u in $H_u + V_u$ and let B_u be the g-orthogonal complement of $H_u + V_u$ in T_uD. We have the direct sum decomposition

$$T_uD = U_u \oplus A_u \oplus B_u.$$

Consider $w \in U_u$. We want to show that there exists a G-invariant function f on D such that $T\rho(w) = T\rho(Y_f(u))$. Since $w \in H_u$ and ϖ_u is a symplectic form on H_u, there exists a linear form

$\alpha \in H_u^*$ such that $w \lrcorner \varpi = \alpha$. Since $u \in U_u$, it follows from (2.14) that $\langle \alpha | v \rangle = 0$ for every $v \in V_u \cap H_u$. Hence, we can extend α to a linear form β on $T_u D = U_u \oplus A_u \oplus B_u$ such that $\langle \beta | v \rangle = 0$ for every $v \in V_u \oplus B_u$. Since $V_u = T_u D \cap \ker T_u \rho$ and β annihilates V_u, it follows that there exists $\gamma \in T^*_{\rho(u)} \overline{D}$ such that $\beta = \gamma \circ T_u \rho$. However, there exists a smooth function $\overline{f}$ on $\overline{D}$ such that $\gamma = d\overline{f}_{\rho(u)}$. Let $f = \overline{f} \circ \rho \in C^\infty(D)$. Then f is G-invariant and $Y_f(u) \in U_u \subseteq H_u$. For every $w' \in U_u$, we have

$$\begin{aligned} \varpi(Y_f(u), w') &= \langle df_u \mid w' \rangle = \langle d\overline{f}_{\rho(u)} \circ T\rho \mid w' \rangle = \langle \gamma \circ \rho | w' \rangle \\ &= \langle \beta | w' \rangle = \langle \alpha \mid w' \rangle = \varpi(w, w'). \end{aligned}$$

Hence

$$\overline{\varpi}(T\rho(Y_f(u)), \overline{w}')) = \overline{\varpi}(T\rho(w), \overline{w}')$$

for all $\overline{w}' \in \overline{H}_{\overline{u}}$. Since $\overline{\varpi}$ is symplectic on $\overline{H}_{\overline{u}}$ it follows that $T\rho(w) = T\rho(Y_f(u))$.

We have shown that every vector $\overline{w} \in \overline{H}_{\overline{u}}$ is the value at $\overline{u}$ of the projection to $\overline{D}$ of a distributional Hamiltonian vector field Y_f of some G-invariant function f. Since H has finite dimension, it follows that $\overline{H}$ is locally spanned by a finite number of projections to D of distributional Hamiltonian vector fields invariant functions. Hence $\overline{H}$ is a generalized distribution. □

If $\overline{H}$ is a distribution, then it is a manifold and $\overline{\varpi}$ is smooth as a bilinear map $\overline{H} \times \overline{H} \to \mathbb{R}$. For a generalized distribution $\overline{H}$, this notion of smoothness of $\overline{\varpi}$ is not applicable. We now discuss a suitable replacement. For every 1-form $\overline{\phi}$ on $\overline{D}$, we denote by $\overline{\phi}_{\overline{H}}$ the restriction of $\overline{\phi}$ to vectors in $\overline{H}$. Since $\overline{\varpi}_{\overline{u}}$ is symplectic for every $\overline{u} \in \overline{D}$, there exists a unique vector field $\overline{Y}_{\overline{\phi}}$ on $\overline{D}$ with values in $\overline{H}$ such that

$$\overline{Y}_{\overline{\phi}} \lrcorner \overline{\varpi} = \overline{\phi}_{\overline{H}}. \tag{2.19}$$

Theorem 3. *For every smooth* 1*-form* $\overline{\phi}$ *on* $\overline{D}$*, the vector field* $\overline{Y}_{\overline{\phi}}$ *defined by equation* (2.19) *is smooth.*

Proof. Locally, we can write $\overline{\phi}$ in the form

$$\overline{\phi} = \overline{h}_1 \, \mathrm{d}\overline{f}_1 + ... + \overline{h}_n \, \mathrm{d}\overline{f}_n,$$

where $\overline{h}_1, ..., \overline{h}_n, \overline{f}_1 ... \overline{f}_n$ are smooth functions on $\overline{D}$. Then $h_i = \overline{h}_i \circ \rho$ and $f_i = \overline{f}_i \circ \rho$ are smooth G-invariant functions on D and

$$\overline{Y}_{\overline{\phi}} \circ \rho = T\rho \circ (h_1 Y_{f_1} + ... + h_n Y_{f_n}).$$

This implies that $\overline{Y}_{\overline{\phi}}$ is smooth. □

Summarizing, we have shown that for a free and proper action of the symmetry group, the process of reduction leads from a symplectic distribution (H, ϖ) on D to a generalized symplectic distribution distribution $(\overline{H}, \overline{\varpi})$ on $\overline{D}$. An alternative, equivalent description of the geometry of the reduced space $\overline{D}$ can be given in terms of the almost Poisson structure on induced by $(\overline{H}, \overline{\varpi})$. See [8] and [10].

3. Chaplygin's skate

In this section we use the example of Chaplygin's skate to illustrate the theory of reduction treated above.

3.1. The skate

Chaplygin's skate (or a dynamic hacket planimeter) is a planar rigid body with a chisel pointing downward at one end of a stick and a rod with a sharp tip pointing downward at the other end. The skate moves on a horizontal plane under the influence of a constant vertical gravitational force so that the edge of the chisel moves only in the direction of the center of mass of the skate. In other words, the chisel edge follows a pursuit curve relative to the center of mass. There are several classic sources which treat Chaplygin's skate. We mention Hamel [6, p.465-470] and Rosenberg [9, p.334]. For a more modern treatment see Bates [1], Koiller [7, p.126] or Cantrijn et al. [3].

3.2. A Lie group model

In this subsection we give a Lie group model of Chaplygin's skate, which we use throughout the rest of our discussion.

Suppose that x is a point in $\mathbb{R}^2$ on the reference skate $\mathcal{S}$. Then a point $x(t)$ on the moving skate is given by applying a 2-dimensional Euclidean motion $(A, a) \in \mathrm{E}(2) = \mathrm{SO}(2) \times \mathbb{R}$ to the point x, that is,

$$x(t) = A(t)x + a(t).$$

Thus the configuration space of Chaplygin's skate is the Lie group $\mathrm{E}(2)$. As a manifold $\mathrm{E}(2)$ is $\mathrm{SO}(2) \times \mathbb{R}$, whereas as a group its multiplication $\cdot$ is given by

$$(A', a') \cdot (A, a) = (A'A, A'a + a').$$

The Lie algebra $\mathrm{e}(2)$ of $\mathrm{E}(2)$ is $\mathrm{so}(2) \times \mathbb{R}$ with Lie bracket

$$[(\Omega', b'), (\Omega, b)] = (0, \Omega' b - \Omega b').$$

Denoting differentiation with respect to time by $\dot{}$, we observe that $\dot{x} = \dot{A}x + \dot{a}$ is the velocity of a point on the moving skate with repect to the fixed frame of the reference skate. Thus the kinetic energy of the unconstrained moving skate, whose mass distribution is the finite measure dm, is

$$T = \tfrac{1}{2} \int_{\mathcal{S}} \langle \dot{x}, \dot{x} \rangle \, dm \; = \; \tfrac{1}{2} \int_{\mathcal{S}} \langle \dot{A}x + \dot{a}, \dot{A}x + \dot{a} \rangle \, dm.$$

Here $\langle \, , \, \rangle$ is the Euclidean inner product. Because we have chosen the origin to be the center of mass of the reference skate, that is, $\int_{\mathcal{S}} x \, dm = 0$, it follows that $T = T_{\mathrm{rot}} + T_{\mathrm{trans}}$. The rotational kinetic energy about the center of mass of the unconstrained skate is

$$\begin{aligned} T_{\mathrm{rot}} &= \tfrac{1}{2} \int_{\mathcal{S}} \langle \dot{A}x, \dot{A}x \rangle \, dm \; = \; \tfrac{1}{2} \int_{\mathcal{S}} \langle A^{-1}\dot{A}x, A^{-1}\dot{A}x \rangle \, dm \\ &= \tfrac{1}{2} \int_{\mathcal{S}} \langle \Omega x, \Omega x \rangle \, dm \\ &= (\tfrac{1}{2} \int_{\mathcal{S}} \langle x, x \rangle \, dm) \, i(\Omega)^2 \; = \; \langle I(i(\Omega)), i(\Omega) \rangle, \end{aligned}$$

where I is the moment of inertia of the skate about its center of mass and $\Omega = A^{-1}\dot{A} = i(\Omega)E \in \mathrm{so}(2)$. The matrix $E = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is a basis of $\mathrm{so}(2)$. The mapping $i : \mathrm{so}(2) \to \mathbb{R} : X = xE \mapsto x$ is an isometry from $(\mathrm{so}(2), \langle\langle \, , \, \rangle\rangle)$, where $\langle\langle X, Y \rangle\rangle = -\frac{1}{2} \mathrm{tr} XY$, to $(\mathbb{R}, \langle \, , \, \rangle)$, where $\langle x, y \rangle = xy$. The translational kinetic energy of the center of mass of the unconstrained skate is

$$\begin{aligned} T_{\mathrm{trans}} &= \tfrac{1}{2} \int_{\mathcal{S}} \langle \dot{a}, \dot{a} \rangle \, dm \; = \; \tfrac{1}{2} \int_{\mathcal{S}} \langle A^{-1}\dot{a}, A^{-1}\dot{a} \rangle \, dm \\ &= (\tfrac{1}{2} \int_{\mathcal{S}} dm) \, \langle b, b \rangle, \quad \text{where } b = A^{-1}\dot{a} \\ &= \tfrac{1}{2} \, \mathrm{m} \, \langle b, b \rangle. \end{aligned}$$

Here m is the mass of the skate. Thus the kinetic energy T is a homogeneous quadratic function on e(2) (with coordinates (Ω, b)) given by

$$T = T(\Omega, b) = \tfrac{1}{2}\langle\langle \mathcal{I}(\Omega), \Omega\rangle\rangle + \tfrac{1}{2}\,\mathrm{m}\,\langle b, b\rangle,$$

where $\mathcal{I} : \mathrm{so}(2) \rightarrow \mathrm{so}(2)$ is the generalized moment of inertia map defined by $\mathcal{I} = i^{-1}(I\,i)$. Hence T determines an inner product on e(2)

$$\mathrm{k} : \mathrm{e}(2) \times \mathrm{e}(2) \rightarrow \mathbb{R} : ((\Omega, b), (\Omega', b')) \mapsto \langle\langle \mathcal{I}(\Omega), \Omega\rangle\rangle + \mathrm{m}\,\langle b, b\rangle,$$

which extends to a unique left invariant Riemannian metric on E(2).

Since the center of mass of the skate has constant height above the horizontal plane, the potential energy of the skate is constant. Hence the motion of the unconstrained skate is determined by its kinetic energy.

We now impose the nonholonomic constraint on the skate so that the edge of the chisel follows a pursuit curve relative to its center of mass. In mathematical terms this states that the chisel edge lies in the kernel of the 1-form ϕ on E(2) defined by

$$\phi(A, a) = \langle\langle A^{-1}\,\mathrm{d}A, E\rangle\rangle - \langle A^{-1}\,\mathrm{d}a, e_2\rangle. \tag{3.1}$$

Here $\{e_1, e_2\}$ is the standard basis for $\mathbb{R}^2$. To see what formula (3.1) means, let $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in \mathrm{SO}(2)$, where $\theta \in \mathbb{R}/(2\pi\,\mathbb{Z})$, and let $a = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$. Then

$$\begin{aligned} \phi \;=\; & \left\langle\!\!\left\langle \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} -\sin\theta & -\cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix} \mathrm{d}\theta, E \right\rangle\!\!\right\rangle - \\ & - \left\langle \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \mathrm{d}a_1 \\ \mathrm{d}a_2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \\ =\; & \langle\langle E, E\rangle\rangle\,\mathrm{d}\theta - \left\langle \begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \end{pmatrix}, \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} \right\rangle = \mathrm{d}\theta + \sin\theta\,\mathrm{d}x - \cos\theta\,\mathrm{d}y. \end{aligned}$$

As a submanifold of $\mathrm{E}(2) \times \mathrm{e}(2)$ (with coordinates (A, a, Ω, b)), the constraint distribution $D = \ker\phi$ is defined by

$$0 = \langle\langle A^{-1}\dot{A}, E\rangle\rangle - \langle A^{-1}\dot{a}, e_2\rangle = \langle\langle \Omega, E\rangle\rangle - \langle b, e_2\rangle = i(\Omega) - b_2.$$

Thus

$$D = \{(A, a, b_2 E, b) \in \mathrm{E}(2) \times \mathrm{e}(2) \,\big|\, (A, a, b) \in \mathrm{E}(2) \times \mathbb{R}^2\}.$$

3.3. The equations of motion

In this subsection we derive the equations of motion of Chaplygin's skate in distributional Hamiltonian form.

First we determine the distribution

$$H = TD \cap \ker \pi^*\phi$$

on the constraint manifold D. Here π is the mapping

$$\pi : \mathrm{E}(2) \times \mathrm{e}(2) \rightarrow \mathrm{E}(2) : (A, a, \Omega, b) \mapsto (A, a).$$

Let $c = (A, a, b_2 E, b) \in D$ and let $(Y, y), (Z, z) \in \mathrm{e}(2)$. For $v_c = (AY, Ay, Z, z) \in T_c(\mathrm{E}(2) \times \mathrm{e}(2))$ to lie in $T_c D$, we must have

$$0 = \langle\langle \dot{\Omega}, E\rangle\rangle - \langle \dot{b}, e_2\rangle = \langle\langle Z, E\rangle\rangle - \langle z, e_2\rangle = i(Z) - z_2,$$

while v_c lies in $\ker \pi^*\phi(c)$ if and only if

$$0 = \langle\langle Y, E\rangle\rangle - \langle y, e_2\rangle \;=\; i(Y) - y_2.$$

Thus

$$H_c = \{(A(y_2\, E), Ay, z_2\, E, z) \in T_c(\mathrm{E}(2)\times \mathrm{e}(2)) \big|\, y, z\in \mathbb{R}^2\},$$

which is 4-dimensional.

Next we find a symplectic form ϖ_H on H. The pull back of the canonical 1-form θ_0 on $T^*\,\mathrm{E}(2)$ by the $\mathrm{E}(2)$-invariant metric is a 1-form θ on $T\,\mathrm{E}(2)$. Pulling θ back by the left trivialization

$$\lambda : \mathrm{E}(2)\times \mathrm{e}(2) \to T\,\mathrm{E}(2) : (A, a, \Omega, b) = (A, a, A^{-1}\dot{A}, A^{-1}\dot{a}) \mapsto (A, a, \dot{A}, \dot{a}),$$

gives the 1-form $\widetilde{\theta}$ on $\mathrm{E}(2)\times\mathrm{e}(2)$

$$\widetilde{\theta}(A, a, \Omega, b)(AY, Ay, Z, z) = \mathrm{k}((\Omega, b), (Y, y)).$$

Thus $\widetilde{\omega} = -\,\mathrm{d}\widetilde{\theta}$ is a symplectic form on $\mathrm{E}(2)\times\mathrm{e}(2)$. Explicitly,

$$\begin{aligned}\widetilde{\omega}(A, a, \Omega, b)\big((AY, Ay, Z, z), (AY', Ay', Z', z')\big) \quad &= \\ = -\mathrm{k}((Z, z), (Y', y')) + \mathrm{k}((Z', z'), (Y, y)) + \mathrm{k}((\Omega, b), [(Y, y), (Y', y')]) &= \\ = \langle I(i(Z'), i(Y)\rangle - \langle I(i(Z)), i(Y')\rangle + \mathrm{m}\,(\langle z', y\rangle - \langle z, y'\rangle) &+ \\ + \mathrm{m}\,\langle b, Yy' - Y'y\rangle. &\end{aligned} \tag{3.2}$$

Consequently, $\widetilde{\omega}(c)$ restricted to $H_c\times H_c$ is the symplectic form ϖ_H on H. In other words, for $c = (A, a, b_2E, b) \in D$ and $v_c = (A(y_2E), Ay, z_2E, z)$, $v'_c = (A(y'_2E), Ay', z'_2E, z') \in H_c$ we have

$$\varpi_H(c)(v_c, v'_c) = (\mathrm{m}+I)(z'_2y_2 - z_2y'_2) + \mathrm{m}\,(z'_1y_1 - z_1y'_1) + \mathrm{m}b_2\,(y_2y'_1 - y'_2y_1).$$

The distributional Hamiltonian equations of motion for Chaplygin's skate are

$$(Y_T \lrcorner\, \varpi_H)(c) = \partial_H T(c), \tag{3.3}$$

for every $c\in D$. Because $Y_T(c)\in H_c$, we may write

$$Y_T(c) = (\beta\, AE, A\Big(\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\Big), \delta\, E, \Big(\begin{smallmatrix}\gamma\\ \delta\end{smallmatrix}\Big)), \tag{3.4}$$

for some $\alpha, \beta, \gamma, \delta \in \mathbb{R}$. On the constraint manifold D the kinetic energy is

$$\begin{aligned} T \quad &= T(A, a, b_2\, E, b) = \tfrac{1}{2}\,\langle \mathcal{I}(\langle b_2\, E), b_2\, E\rangle\rangle + \tfrac{1}{2}\,\mathrm{m}\,\langle b, b\rangle = \\ &= \tfrac{1}{2}\,\mathrm{m}\, b_1^2 + \tfrac{1}{2}\,(\mathrm{m}+I)b_2^2.\end{aligned}$$

Hence

$$\mathrm{d}T = \mathrm{m}b_1\;\mathrm{d}b_1 + (\mathrm{m}+I)b_2\;\mathrm{d}b_2. \tag{3.5}$$

Evaluating (3.3) at $(\beta' AE, A\Big(\begin{smallmatrix}\alpha'\\ \beta'\end{smallmatrix}\Big), \delta' E, \Big(\begin{smallmatrix}\gamma'\\ \delta'\end{smallmatrix}\Big))$ in H_c and using (3.4) and (3.5) we obtain

$$\mathrm{m}b_1\,\gamma' + (\mathrm{m}+I)b_2\,\delta' = (\mathrm{m}+I)(\delta'\beta - \delta\beta') + \mathrm{m}(\gamma'\alpha - \gamma\alpha') + \mathrm{m}b_2\,(\beta\alpha' - \beta'\alpha),$$

for every $\alpha', \beta', \gamma', \delta' \in \mathbb{R}$. Successively setting all but one of the variables $\alpha', \beta'\gamma', \delta'$ in the above equation equal to 0 and the remaining variable equal to 1 gives

$$\alpha = b_1,\;\; \beta = b_2,\;\; \gamma = b_2^2,\;\text{ and }\;\delta = -\frac{\mathrm{m}}{(\mathrm{m}+I)}\,b_1b_2.$$

Hence a motion of Chaplgyin's skate is an integral curve of the vector field

$$Y_T(A, a, b_2\, E, b) = \Big(A(b_2 E), Ab, -\frac{\mathrm{m}}{(\mathrm{m}+I)}\, b_1 b_2\, E, \begin{pmatrix} b_2^2 \\ -\frac{\mathrm{m}}{(\mathrm{m}+I)}\, b_1 b_2 \end{pmatrix}\Big)$$

on D, that is, it satisfies

$$\begin{cases} \dot{A} & = A(b_2 E) \\ \dot{a} & = Ab \\ \begin{pmatrix} \dot{b}_1 \\ \dot{b}_2 \end{pmatrix} & = \begin{pmatrix} b_2^2 \\ -\frac{\mathrm{m}}{(\mathrm{m}+I)}\, b_1 b_2 \end{pmatrix}. \end{cases} \tag{3.6}$$

We have omitted the equation $\dot{b}_2\, E = -\frac{\mathrm{m}}{(\mathrm{m}+I)}\, b_1 b_2\, E$ from (3.6) because it is a consequence of the third equation.

3.4. E(2) symmetry

In this subsection we show that Chaplygin's skate has an E(2) symmetry.

The Euclidean group E(2) acts on the configuration space E(2) of the skate by left multiplication, namely,

$$\Phi : \mathrm{E}(2) \times \mathrm{E}(2) \to \mathrm{E}(2) : ((A', a'), (A, a)) \mapsto (A'A, A'a + a').$$

This action lifts to an action on the tangent bundle $T\,\mathrm{E}(2)$ given by

$$\Psi : \mathrm{E}(2) \times T\,\mathrm{E}(2) \to T\,\mathrm{E}(2) : ((A', a'), (A, a, \dot{A}, \dot{a})) \mapsto (A'A, A'a + a', A'\dot{A}, A'\dot{a}).$$

The left trivialization

$$\lambda : \mathrm{E}(2) \times \mathrm{e}(2) \to T\,\mathrm{E}(2) : (A, a, \Omega, b) = (A, a, A^{-1}\dot{A}, A^{-1}\dot{a}) \mapsto (A, a, \dot{A}, \dot{a})$$

intertwines the E(2)-action

$$\begin{aligned} \widetilde{\Psi} : \mathrm{E}(2) \times (\mathrm{E}(2) \times \mathrm{e}(2)) &\to \mathrm{E}(2) \times \mathrm{e}(2) : \\ ((A', a'), (A, a, \Omega, b)) &\mapsto (A'A, A'a + a', \Omega, b) \end{aligned}$$

with the E(2)-action Ψ. The E(2)-action $\widetilde{\Psi}$ preserves the constraint 1-form $\pi^*\phi$ and thus the constraint manifold D. Since it also preserves distribution H, the kinetic energy T, and the symplectic form ϖ_H, it follows that $\widetilde{\Psi}$ preserves the vector field Y_T and hence is a symmetry of Chaplygin's skate.

3.5. The distributions V, $V \cap H$, and U

To reduce the E(2) symmetry of Chaplygin's skate, we follow the pattern of reduction of symmetries for distributional Hamiltonian systems described in section 2. This involves constructing the distributions V, $V \cap H$, and U, which are defined below.

Consider the E(2) action $\widehat{\Psi}$ on the constraint manifold D given by

$$\widehat{\Psi} : \mathrm{E}(2) \times D \to D : ((A', a'), (A, a, b_2 E, b)) \mapsto (A'A, A'a + a', b_2 E, b).$$

Let $\mathcal{O}_c$ be the $\widehat{\Psi}$-orbit through $c = (A, a, b_2 E, b) \in D$. The space of orbits $\overline{D} = D/\,\mathrm{E}(2)$ of the E(2) action $\widehat{\Phi}$ is clearly $\mathbb{R}^2$ with orbit map

$$\rho : D \subseteq \mathrm{E}(2) \times \mathrm{E}(2) \to \overline{D} : c = (A, a, b_2\, E, b) \mapsto b,$$

which is a surjective submersion, that is, the maps ρ and

$$T_c\rho : T_cD \to T_{\rho(c)}\overline{D} : (AY, Ay, z_2\, E, z) \mapsto z$$

are surjective. Differentiating the map

$$\widehat{\Psi}_c : \mathrm{E}(2) \to D : (A', a') \mapsto \widehat{\Psi}((A', a'), c)$$

at the identity element e of $\mathrm{E}(2)$ gives

$$T_e\widehat{\Psi}_c : \mathrm{e}(2) \to T_cD : (\dot{A}', \dot{a}') \mapsto (\dot{A}'A, \dot{A}'a + \dot{a}', 0, 0).$$

Thus we obtain the distribution V on D where

$$V_c = T_e\mathcal{O}_c = \{(\lambda EA, \lambda Ea + \dot{a}', 0, 0) \in T_cD \big| (\lambda E, \dot{a}') \in \mathrm{e}(2)\}.$$

Fix $c = (A, a, b_2E, b) \in D$. The following shows that $\dim V_c \cap H_c = 2$. For $v_c = (Ay_2\, E), y, z_2\, E, z) \in H_c \subseteq T_cD$ we have $T_c\rho(v_c) = z$. Consequently, $T_c\rho(H_c) = T_{\rho(c)}\overline{D}$. But $(\ker T_c\rho) \cap H_c = V_c \cap H_c$. Hence

$$\begin{aligned} 2 &= \dim T_{\rho(c)}\overline{D} = \dim T_c\rho(H_c) = \\ &= \dim H_c - \dim((\ker T_c\rho) \cap H_c) = \dim H_c - \dim V_c \cap H_c. \end{aligned}$$

Since $\dim H_c = 4$, we get $\dim V_c \cap H_c = 2$. We now determine $V_c \cap H_c$. The nonzero vector $w_c = (AY, Ay, Z, z)$ lies in $V_c \cap H_c$ if

$$(\lambda EA, \lambda Ea + \dot{a}', 0, 0) = (AY, Ay, Z, z) \tag{3.7}$$

for some $(Y, y), (Z, z) \in \mathrm{e}(2)$ which are not both zero. Equation (3.7) implies that $(Z, z) = (0, 0)$ and $AY = \lambda EA = \lambda AE$. The second equality in the preceding equation follows because $\mathrm{SO}(2)$ is abelian implies that $\mathrm{Ad}_A = \mathrm{id}$ for every $A \in \mathrm{SO}(2)$ and hence $A^{-1}EA = E$. Consequently, $Y = \lambda E$. In addition, equation (3.7) yields $y = A^{-1}(\lambda Ea + \dot{a}')$. For $(AY, Ay, 0, 0)$, with Y and y given as above, to lie in H_c we must have $\langle\langle Y, E\rangle\rangle = y_2$. This is equivalent to

$$\begin{aligned} \lambda &= \langle \lambda A^{-1}Ea + \dot{a}', e_2\rangle = \lambda\langle EA^{-1}a, e_2\rangle + \langle \dot{a}', Ae_2\rangle = \\ &= -\lambda\langle A^{-1}a, Ee_2\rangle + \langle \dot{a}', Ae_2\rangle, \end{aligned}$$

that is,

$$\langle \dot{a}', Ae_2\rangle = \lambda(1 - \langle a, Ae_1\rangle). \tag{3.8}$$

Let

$$\mathcal{E}_{(A,a)} = \left\{(\omega\, E, \beta) \in \mathrm{e}(2) \big| \langle \beta, Ae_2\rangle = \omega(1 - \langle Ae_1, a\rangle)\right\}.$$

Then $\mathcal{E}_{(A,a)}$ is a codimension 1 subspace of $\mathrm{e}(2)$ which is parametrized by

$$(\lambda\, E, \lambda(1 - \langle Ae_1, a\rangle)Ae_2 + \mu\, Ae_1),$$

for every $\mu, \lambda \in \mathbb{R}$. Consequently $V_c \cap H_c$ is the 2-dimensional subspace of T_cD given by

$$\left\{\Big(A(\lambda E), A(A^{-1}(\lambda E)a + \mu e_1 + \lambda(1 - \langle a, Ae_1\rangle)e_2), 0, 0\Big) \in T_cD \big| \lambda, \mu \in \mathbb{R}\right\}.$$

Thus $V \cap H$ is a distribution on D.

We now determine the space

$$U_c = H_c \cap (V_c \cap H_c)^{\varpi_H(c)},$$

where $(V_c \cap H_c)^{\varpi_H(c)}$ is the symplectic annihilator of $V_c \cap H_c$. The vector $v'_c = (A(y'_2 E), Ay', z'_2 E, z') \in H_c$ lies in U_c if and only if for every $\lambda, \mu \in \mathbb{R}$,

$$\begin{aligned} 0 &= \varpi_H(c)\Big((A(\lambda E), A(A^{-1}(\lambda E)a + \mu\, e_1 + \lambda(1 - \langle a, Ae_1\rangle)e_2), 0, 0), v'_c\Big) = \\ &= (\mathrm{m} + I)\lambda\, y_2 z'_2 + \mathrm{m}(\lambda\, y_1 + \mu)z'_1 + \mathrm{m}b_2\, y_2 y'_1 - \mathrm{m}b_2(\lambda\, y_1 + \mu)y'_2, \end{aligned} \tag{3.9}$$

where

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = A^{-1}Ea + (1 - \langle a, Ae_1\rangle)e_2 = \begin{pmatrix} -\langle a, Ae_2\rangle \\ 1 \end{pmatrix}. \tag{3.10}$$

Setting $\lambda = 0$ and $\mu = 1$ in (3.9) gives $z'_1 = b_2\, y'_2$. Setting $\lambda = 1$ and $\mu = 0$ in (3.9) gives

$$0 = (\mathrm{m} + I)\, y_2 z'_2 + \mathrm{m}\, y_1 z'_1 + \mathrm{m}b_2\, y_2 y'_1 - \mathrm{m}b_2\, y_1 y'_2 \;=\; (\mathrm{m} + I)z'_2 + \mathrm{m}b_2\, y'_1,$$

since $z'_1 = b_2\, y'_2$ and $y_2 = 1$ by (3.10). Thus $z'_2 = -\dfrac{\mathrm{m}}{(\mathrm{m} + I)} b_2\, y'_1$. Consequently

$$U_c = \left\{ \left(A(y'_2 E), Ay', -\frac{\mathrm{m}}{(\mathrm{m} + I)} b_2 y'_1\, E, \begin{pmatrix} b_2 y'_2 \\ -\dfrac{\mathrm{m}}{(\mathrm{m} + I)} b_2 y'_1 \end{pmatrix}\right) \in H_c \,\Big|\, y' \in \mathbb{R}^2 \right\}.$$

Thus U is a distribution on D.

For $c' = (A, a, 0, \begin{pmatrix} b_1 \\ 0 \end{pmatrix}) \in D$, note that

$$U_{c'} = \{(A(y'_2\, E), Ay', 0, 0) \in H_{c'} \,\big|\, y' \in \mathbb{R}^2\} \;=\; V_{c'} \cap H_{c'}.$$

Thus $V_{c'} \cap H_{c'}$ is a Lagrangian subspace of the symplectic vector space $(H_{c'}, \varpi_H(c'))$.

3.6. Reduction of the E(2) symmetry

In this subsection we reduce the E(2) symmetry of Chaplygin's skate.

Since the kinetic energy T of the skate on D is invariant under the E(2)-action $\widehat{\Psi}$, it induces a function

$$\overline{T} = \overline{T}(b) \;=\; \tfrac{1}{2}\,\mathrm{m}\, b_1^2 + \tfrac{1}{2}\,(\mathrm{m} + I) b_2^2$$

on the E(2)-orbit space $\overline{D} = \mathbb{R}^2$ called the E(2)-reduced kinetic energy.

Since the vector field Y_T is also invariant under the E(2)-action $\widehat{\Psi}$, it too induces a vector field

$$Y_{\overline{T}}(b) = Y_{\overline{T}}(\rho(c)) \;=\; T_c\rho(Y_T(c)) \;=\; \begin{pmatrix} b_2^2 \\ -\dfrac{\mathrm{m}}{(\mathrm{m} + I)}\, b_1 b_2 \end{pmatrix}.$$

In other words, on the orbit space $\mathbb{R}^2$ (with coordinates (b_1, b_2)) the integral curves of the E(2)-reduced vector field $Y_{\overline{T}}$ satisfy

$$\begin{aligned} \dot{b}_1 &= b_2^2 \\ \dot{b}_2 &= -\frac{\mathrm{m}}{(\mathrm{m} + I)}\, b_1 b_2. \end{aligned} \tag{3.11}$$

Even though $\overline{T}$ is an integral of $Y_{\overline{T}}$, a straightforward integration of (3.11) shows that every solution is asymptotic in forward and backward time to a point on the line $\{b_2 = 0\}$.

For fixed $c \in D$ the image of U_c under the tangent of the orbit map ρ is

$$\overline{H}_b = \overline{H}_{\rho(c)} \;=\; T_c\rho(U_c) \;=\; \left\{ b_2 \begin{pmatrix} y'_2 \\ -\dfrac{\mathrm{m}}{\mathrm{m} + I}\, y'_1 \end{pmatrix} \in \mathbb{R}^2 \big| y' \in \mathbb{R}^2 \right\}.$$

When $b_2 = 0$, then $\overline{H}_b = \{0\}$. This occurs when $c' = (A, a, 0, \begin{pmatrix} b_1 \\ 0 \end{pmatrix}) \in D$, because then $U_{c'} = V_{c'} \cap H_{c'} \subseteq \ker T_{c'}\rho$. Thus $\overline{H}$ is a *not* a distribution on D. From theorem 2 of section 2 we know that $\overline{H}$ is a *generalized* distribution on $\overline{D}$.

Finally we compute the reduced symplectic form $\varpi_{\overline{H}}$ on $\overline{H}$. Fix $c = (A, a, b_2E, b) \in D$. The form $\varpi_{\overline{H}}$ is defined by

$$\varpi_{\overline{H}}(\rho(c))(T_c\rho(v_c), T_c\rho(v_c')) = \varpi_H(c)(v_c, v_c'),$$

for $v_c, v_c' \in U_c$.

Suppose that $b_2 \neq 0$ and let $v_c = (A(y_2E), Ay, -\frac{\mathrm{m}}{\mathrm{m}+I}y_1\,E, \begin{pmatrix} b_2\,y_2 \\ -\frac{\mathrm{m}}{\mathrm{m}+I}b_2\,y_1 \end{pmatrix})$ and $v_c' = (A(y_2'E), Ay', -\frac{\mathrm{m}}{\mathrm{m}+I}y_1'\,E, \begin{pmatrix} b_2\,y_2' \\ -\frac{\mathrm{m}}{\mathrm{m}+I}b_2\,y_1' \end{pmatrix})$. Let $z_1 = b_2\,y_2$, $z_1' = b_2\,y_2'$, $z_2 = -\frac{\mathrm{m}}{\mathrm{m}+I}b_2\,y_1$ and $z_2' = -\frac{\mathrm{m}}{\mathrm{m}+I}b_2\,y_1'$. Then

$$\begin{aligned} \varpi_{\overline{H}}(b)(z, z') &= (\mathrm{m}+I)(z_1'y_2 - z_2y_2') + \mathrm{m}\,(z_1'y_1 - z_1y_1') + \mathrm{m}b_2\,(y_2y_1' - y_2'y_1) = \\ &= \mathrm{m}b_2\,(y_2y_1' - y_2'y_1) = \frac{(\mathrm{m}+I)}{b_2}\,(z_1z_2' - z_1'z_2). \end{aligned}$$

Hence $\varpi_{\overline{H}}$ is symplectic.

If $b_2 = 0$, then $\overline{H} = \{0\}$. Thus the statement that $\varpi_{\overline{H}}$ is nondegenerate is vacuously satisfied.

Acknowledgement

The first author (R.C.) was supported in part by European Commission functing for the Research Training Network “Mechanics and Symmetry in Europe” (MASIE), Contract No. HPRN-CT-2000-00113. The second author (J.Ś.) was partially supported by the National Science and Engineering Research Council grant RGPIN 8091-00.

References

[1] *L. Bates*. Examples of singular nonholonomic reduction. Rep. Math. Phys. V. 42. 1998. P. 231–247.

[2] *L. Bates and J. Śniatycki*. Non-holonomic reduction. Rep. Math. Phys. V. 32. 1993. P. 99–115.

[3] *F. Cantrijn, J. Corés, M. de León and D. M. de Diego*. On the geometry of general Chaplygin systems. Preprint. 2000.

[4] *R. Cushman and L. Bates*. Global aspects of classical integrable systems. Birkhäuser, Basel. 1997.

[5] *R. Cushman, D. Kemppainen, J. Śniatycki, and L. Bates*. Geometry of nonholonomic constraints. Rep. Math. Phys. V. 36. 1995. P. 275–268.

[6] *G. Hamel*. Theoretische Mechanik. Springer-Verlag, Berlin. 1949.

[7] *J. Koiller*. Reduction of some classical nonholonomic systems with symmetry. Arch. Rat. Mech. Anal. V. 142. 1992. P. 113–148.

[8] *W. S. Koon and J. E. Marsden*. Poisson reduction for nonholonomic mechanical systems with symmetry. Rep. Math. Phys. V. 42. 1998. P. 101–134.

[9] *R. Rosenberg*. Analytical Dynamics. Plenum Press, New York. 1977.

[10] *J. Śniatycki*. Almost Poisson Spaces and Nonholonomic Singular Reduction. Rep. Math. Phys. V. 48. 2001. P. 235–248.

[11] *H. Sussmann*. Orbits of families of vector fields and integrability of distributions. Trans. Amer. Math. Soc. V. 180. 1973. P. 171–188.