IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Towards a Hamilton–Jacobi theory for nonholonomic mechanical systems

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2008 J. Phys. A: Math. Theor. 41 015205

(http://iopscience.iop.org/1751-8121/41/1/015205)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 171.66.16.147
The article was downloaded on 03/06/2010 at 06:35

Please note that terms and conditions apply.

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **41** (2008) 015205 (14pp) doi:10.1088/1751-8113/41/1/015205

Towards a Hamilton–Jacobi theory for nonholonomic mechanical systems*

David Iglesias-Ponte, Manuel de León and David Martín de Diego

Instituto de Matemáticas y Física Fundamental, Consejo Superior de Investigaciones Científicas, Serrano 123, 28006 Madrid, Spain

E-mail: iglesias@imaff.cfmac.csic.es, mdeleon@imaff.cfmac.csic.es and d.martin@imaff.cfmac.csic.es

Received 11 April 2007, in final form 22 October 2007
Published 12 December 2007
Online at stacks.iop.org/JPhysA/41/015205

Abstract
In this paper, we obtain a Hamilton–Jacobi theory for nonholonomic mechanical systems. The results are applied to a large class of nonholonomic mechanical systems, the so-called Čaplygin systems.

PACS number: 45.20.Jj
Mathematics Subject Classification: 70H20, 37J60, 70F25

1. Introduction

The standard formulation of the Hamilton–Jacobi problem for a Hamiltonian system is look for a function $S(t, q^A)$ (called the principal function) such that

$$\frac{\partial S}{\partial t} + H\left(q^A, \frac{\partial S}{\partial q^A}\right) = 0, \tag{1.1}$$

where $H : T^*Q \longrightarrow \mathbb{R}$ is the Hamiltonian function. If one looks for solutions of the form $S(t, q^A) = W(q^A) - tE$, where E is a constant, then W must satisfy

$$H\left(q^A, \frac{\partial W}{\partial q^A}\right) = E, \tag{1.2}$$

where W is called the characteristic function.

Equations (1.1) and (1.2) are indistinctly referred as the Hamilton–Jacobi equation.

The advantage of this method is that, in spite of the difficulties to solve a partial differential equation instead of an ordinary differential one, in many cases it works, being an extremely useful tool, usually more than Hamilton's equations. Indeed, in these cases the method

* This work has been partially supported by MEC (Spain) grants MTM 2004-7832, S-0505/ESP/0158 of the CAM and 'Ingenio Mathematica' (i-MATH) no CSD 2006-00032 (Consolider-Ingenio 2010). D Iglesias acknowledges Ministry of Education and Science for a 'Juan de la Cierva' research contract.

© 2008 IOP Publishing Ltd Printed in the UK

provides an immediate way to integrate the equations of motion. The modern interpretation relating the Hamilton–Jacobi procedure with the theory of Lagrangian submanifolds is an important source of new results and insights [1, 2]. Let us remark that, recently, Cariñena *et al* [6] have developed a new approach to the geometric Hamilton–Jacobi theory.

On the other hand, in the last 15 years there has been a renewed interest in nonholonomic mechanics, that is, those mechanical systems given by a Lagrangian $L = L(q^A, \dot{q}^A)$ subject to constraints $\Phi^i(q^A, \dot{q}^A) = 0$ involving the velocities (see [3] and references therein). A relevant difference with the unconstrained mechanical systems is that a nonholonomic system is not Hamiltonian in the sense that the phase space is just the constraint submanifold and not the cotangent bundle of the configuration manifold; moreover, its dynamics is given by an almost Poisson bracket, that is, a bracket not satisfying the Jacobi identity [5]. In [11], the authors proved that the nonholonomic dynamics can be obtained by projecting the unconstrained dynamics; this will be the point of view adopted in the present paper.

A natural question related with a possible notion of integrability is in what extent one could construct a Hamilton–Jacobi theory for nonholonomic mechanics. Past attempts to obtain a Hamilton–Jacobi theory for nonholonomic systems were non-effective or very restrictive (see [7, 19–22] and also [15]), because, in many of them, they try to adapt the typical proof of the Hamilton–Jacobi equations for systems without constraints (using Hamilton's principle). Usually, the results are valid when the solutions of the nonholonomic problem are also the solutions of the corresponding constrained variational problem (see [10, 17, 18] for a complete discussion).

In our paper, we present an alternative approach based on the geometrical properties of nonholonomic systems (see also [16] for second-order differential equations). The method is applied to a particular class of nonholonomic systems called Čaplygin systems: in such a system the configuration manifold is a fibration over another manifold, and the constraints are given by the horizontal subspaces of a connection on the fibration. In this case, the original nonholonomic system is equivalent to another one whose configuration manifold is the base of the fibration and, in addition, it is subject to an external force [12]. In any case, the equations we have obtained are different that in previous works and may give new insight in this topic. In particular, this theory could give insights in the study of integrability for nonholonomic systems [4] and even in the construction of new geometrical integrators for nonholonomic systems (see [9, 13]).

2. Preliminaries

2.1. Lagrangian and Hamiltonian mechanics

Let $L = L(q^A, \dot{q}^A)$ be a Lagrangian function, where (q^A) are coordinates in a configuration n-manifold Q. Hamilton's principle produces the Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A} = 0, \qquad 1 \leqslant A \leqslant n. \tag{2.1}$$

A geometric version of equation (2.1) (see [14]) can be obtained as follows. Consider the (1,1)-tensor field S and the Liouville vector field Δ locally defined on the tangent bundle TQ of Q by

$$S = \frac{\partial}{\partial \dot{q}^A} \otimes \mathrm{d}q^A, \qquad \Delta = \dot{q}^A \frac{\partial}{\partial \dot{q}^A}.$$

Since the Lagrangian L is a function defined on TQ one can construct the Poincaré–Cartan 1- and 2-forms

$$\alpha_L = S^*(\mathrm{d}L), \qquad \omega_L = -\mathrm{d}\alpha_L,$$

where S^* denotes the adjoint operator of S. The energy is given by $E_L = \Delta(L) - L$. We say that L is regular if the 2-form ω_L is symplectic. In this case, the equation

$$i_X\omega_L = \mathrm{d}E_L \tag{2.2}$$

has a unique solution, $X = \xi_L$, called the Euler–Lagrange vector field; ξ_L is a second-order differential equation (SODE) which means that its integral curves are tangent lifts of their projections on Q (these projections are called the solutions of ξ_L). A direct computation shows that the solutions of ξ_L are just the ones of equations (2.1).

Finally, let us recall that the Legendre transformation $FL : TQ \longrightarrow T^*Q$ is a fibred mapping (that is, $\pi_Q \circ FL = \tau_Q$, where $\tau_Q : TQ \longrightarrow Q$ and $\pi_Q : T^*Q \longrightarrow Q$ denote the canonical projections of the tangent and cotangent bundles of Q, respectively). The regularity of L is equivalent to FL being a local diffeomorphism. Along this paper, we will assume that FL is in fact a global diffeomorphism (in other words, L is hyper-regular) which is the case when L is a Lagrangian of mechanical type, say $L = T - V$, where T is the kinetic energy defined by a Riemannian metric on Q and $V : Q \longrightarrow \mathbb{R}$ is a potential energy.

The Hamiltonian counterpart is developed in the cotangent bundle T^*Q of Q. Denote by $\omega_Q = \mathrm{d}q^A \wedge \mathrm{d}p_A$ the canonical symplectic form, where (q^A, p_A) are the canonical coordinates on T^*Q. The Hamiltonian energy is just $H = E_L \circ FL^{-1}$ and the Hamiltonian vector field is the solution of the symplectic equation

$$i_{X_H}\omega_Q = \mathrm{d}H.$$

As we know, the integral curves $(q^A(t), p_A(t))$ of X_H satisfy the Hamilton equations

$$\left.\begin{aligned} \dot{q}^A &= \frac{\partial H}{\partial p_A} \\ \dot{p}_A &= -\frac{\partial H}{\partial q^A} \end{aligned}\right\}. \tag{2.3}$$

Finally, since $FL^*\omega_Q = \omega_L$ we deduce that ξ_L and X_H are FL-related and, consequently, FL transforms the Euler–Lagrange equation (2.1) into the Hamilton equations (2.3).

2.2. *Nonholonomic mechanical systems*

A nonholonomic mechanical system is given by a Lagrangian function $L = L(q^A, \dot{q}^A)$ subject to a family of constraint functions

$$\Phi^i(q^A, \dot{q}^A) = 0, \qquad 1 \leqslant i \leqslant m \leqslant n = \dim Q.$$

In the sequel, we will assume that the constraints Φ^i are linear in the velocities, i.e., $\Phi^i(q^A, \dot{q}^A) = \Phi^i_A(q)\dot{q}^A$.

Invoking the D'Alembert principle, we derive the nonholonomic equations of motion

$$\left.\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A} = \lambda_i \Phi^i_A(q), && 1 \leqslant A \leqslant n \\ &\Phi^i(q^A, \dot{q}^A) = 0, && 1 \leqslant i \leqslant m \end{aligned}\right\}, \tag{2.4}$$

where $\lambda_i = \lambda_i(q^A, \dot{q}^A)$, $1 \leqslant i \leqslant m$, are Lagrange multipliers to be determined.

In a geometrical setting, L is a function on TQ and the constraints are given by a vector sub-bundle M of TQ locally defined by $\Phi^i = 0$.

Equations (2.4) can be intrinsically (see [12]) rewritten as follows:

$$\left.\begin{array}{l} i_X\omega_L - \mathrm{d}E_L \in S^*((TM)^0) \\ X \in TM. \end{array}\right\}. \tag{2.5}$$

For the formulation of a Hamilton–Jacobi theory we are interested in the 'Hamiltonian version' of the nonholonomic equations. Assuming that the Lagrangian L is hyperregular, then the constraint functions on T^*Q become $\Psi^i = \Phi^i \circ FL^{-1}$, i.e.,

$$\Psi^i(q^A, p_A) = \Phi^i_A(q)\frac{\partial H}{\partial p_A}(q^A, p_A),$$

where the Hamiltonian $H : T^*Q \to \mathbb{R}$ is defined by $H = E_L \circ FL^{-1}$.

The equations of motion for the nonholonomic system on T^*Q can now be written as follows:

$$\left.\begin{array}{l} \dot q^A = \dfrac{\partial H}{\partial p_A} \\ \dot p_A = -\dfrac{\partial H}{\partial q^A} - \bar\lambda_i\Phi^i_A(q) \end{array}\right\} \tag{2.6}$$

together with the constraint equations $\Psi^i(q, p) = 0$.

Let $\bar M$ denote the image of the constraint submanifold M under the Legendre transformation, and let $\bar F$ be the distribution on T^*Q along $\bar M$, whose annihilator is given by

$$\bar F^0 = FL^*(S^*((TM)^0)).$$

Observe that $\bar F^0$ is locally generated by the m independent 1-forms

$$\bar\mu^i = \Phi^i_A(q)\mathrm{d}q^A, \qquad 1 \leqslant i \leqslant m.$$

The nonholonomic Hamilton equations for the nonholonomic system can be then rewritten in intrinsic form as

$$\left.\begin{array}{l} (i_X\omega_Q - \mathrm{d}H)_{|\bar M} \in \bar F^0 \\ X_{|\bar M} \in T\bar M \end{array}\right\}. \tag{2.7}$$

Assume the compatibility condition: $\bar F^\perp \cap T\bar M = \{0\}$, where '$\perp$' denotes the symplectic orthogonal with respect to ω_Q. Observe that, locally, this condition means that the matrix

$$(\bar{\mathcal{C}}^{ij}) = \left(\Phi^i_A(q)\mathcal{H}^{AB}\Phi^j_B(q)\right) \tag{2.8}$$

is regular, where $(\mathcal{H}^{AB}) = (\partial^2 H/\partial p_A\partial p_B)$. The compatibility condition is not too restrictive, since it is trivially verified by the usual systems of mechanical type (i.e. with a Lagrangian of the form kinetic minus potential energy). The compatibility condition guarantees, in particular, the existence of a unique solution of the constrained equations of motion (2.7) which, henceforth, will be denoted by $\bar X_{nh}$.

Moreover, if we denote by X_H the Hamiltonian vector field of H, i.e. $i_{X_H}\omega_Q = \mathrm{d}H$ then, using the constraint functions, we may explicitly determine the Lagrange multipliers λ_i as

$$\bar\lambda_i = \bar{\mathcal{C}}_{ij}X_H(\Psi^j), \tag{2.9}$$

where $(\bar{\mathcal{C}}_{ij})$ is the inverse matrix of $(\bar{\mathcal{C}}^{ij})$.

2.3. *Čaplygin systems*

A Čaplygin system is a nonholonomic mechanical system such that:

(i) the configuration manifold Q is a fibred manifold, say $\rho : Q \longrightarrow N$, over a manifold N;

(ii) the constraints are provided by the horizontal distribution of an Ehresmann connection Γ in ρ;
(iii) the Lagrangian $L : TQ \longrightarrow \mathbb{R}$ is Γ-invariant.

Remark 2.1. A particular case is when $\rho : Q \longrightarrow N = Q/G$ is a principal G-bundle and Γ a principal connection.

Let us recall that the connection Γ induces a Whitney decomposition $TQ = \mathcal{H} \oplus V\rho$, where $\mathcal{H}$ is the horizontal distribution and $V\rho = \ker T\rho$ is the vertical distribution. Take fibred coordinates $(q^A) = (q^a, q^i)$ such that $\rho(q^a, q^i) = (q^a)$; therefore, we can obtain an adapted local basis of vector fields

$$\mathcal{H} = \left\langle \mathcal{H}_a = \frac{\partial}{\partial q^a} - \Gamma^i_a \frac{\partial}{\partial q^i} \right\rangle, \qquad V\rho = \left\langle V_i = \frac{\partial}{\partial q^i} \right\rangle.$$

Here $\mathcal{H}_a = \left(\frac{\partial}{\partial q^a}\right)^{\mathcal{H}} = h\left(\frac{\partial}{\partial q^a}\right)$, where $y^{\mathcal{H}}$ denotes the horizontal lift of a tangent vector y on N to Q, and $h : TQ \longrightarrow \mathcal{H}$ is the horizontal projector; $\Gamma^i_a = \Gamma^i_a(q^A)$ are the Christoffel components of the connection Γ.

The dual local basis of 1-forms is

$$\left\{\eta_a = \mathrm{d}q^a, \eta_i = \mathrm{d}q^i + \Gamma^i_a\, \mathrm{d}q^a\right\}.$$

The curvature of Γ is the (1, 2)-tensor field $R = \frac{1}{2}[h, h]$, where $[\ ,\ ]$ is the Nijenhuis tensor of h, that is

$$R(X, Y) = [hX, hY] - h[hX, Y] - h[X, hY] + h^2[X, Y].$$

Therefore, we have

$$R\left(\frac{\partial}{\partial q^a}, \frac{\partial}{\partial q^b}\right) = R^i_{ab} \frac{\partial}{\partial q^i},$$

where

$$R^i_{ab} = \frac{\partial \Gamma^i_a}{\partial q^b} - \frac{\partial \Gamma^i_b}{\partial q^a} + \Gamma^j_a \frac{\partial \Gamma^i_b}{\partial q^j} - \Gamma^j_b \frac{\partial \Gamma^i_a}{\partial q^j}.$$

The constraints are locally given by $\Phi^i = \dot{q}^i + \Gamma^i_a \dot{q}^a = 0$. In other words, the solutions are horizontal curves with respect to Γ.

Since the Lagrangian L is Γ-invariant, that is, $L((Y^{\mathcal{H}})_{q_1}) = L((Y^{\mathcal{H}})_{q_2})$ for all $Y \in T_yN$, $y = \rho(q_1) = \rho(q_2)$, we can define a function $L^* : TN \longrightarrow \mathbb{R}$ as follows: $L^*(Y_y) = L((Y^{\mathcal{H}})_q)$, where $y = \rho(q)$. Therefore we have

$$L^*(q^a, \dot{q}^a) = L\left(q^a, q^i, \dot{q}^a, -\Gamma^i_a \dot{q}^a\right).$$

Equations (2.5) read now as

$$\left.\begin{array}{l} i_X\omega_L - \mathrm{d}E_L \in S^*((T\mathcal{H})^0) \\ X \in T\mathcal{H} \end{array}\right\}. \tag{2.10}$$

Define a 1-form α^* on TN by putting

$$(\alpha^*)(u)(U) = -(\alpha_L)(x)(\tilde{u}),$$

where $U \in T_u(TN)$, $u \in T_yN$, $\tilde{U} \in T_x(TQ)$ such that $\tilde{U}$ projects onto

$$R\left((u)^{\mathcal{H}}_q, (T\tau_N(U)^{\mathcal{H}}_q)\right) \in T_qQ,$$

$\rho(q) = y$, $x \in \mathcal{H}$, $\tau_Q(x) = q$. In local coordinates we obtain

$$\alpha^* = \left(\frac{\partial L}{\partial \dot{q}^i} \dot{q}^b R^i_{ab}\right) \mathrm{d}q^a.$$

Consider the following equation:

$$i_Y\omega_{L^*} - \mathrm{d}E_{L^*} = \alpha^*. \tag{2.11}$$

A long but straightforward proof shows that L^* is a regular Lagrangian on TN, therefore (2.11) has a unique solution Y^*. Note that the pair (L^*, α^*) can be considered as an unconstrained system subject to an external force α^*. The corresponding equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L^*}{\partial \dot{q}^a}\right) - \frac{\partial L^*}{\partial q^a} = -\frac{\partial L}{\partial \dot{q}^i}\dot{q}^b R^i_{ab}. \tag{2.12}$$

Both systems, the nonholonomic one on Q given by L and the constraints given by Γ, and that given by L^* and α^*, are equivalent. The equivalence is explained in the following.

Γ induces a connection $\bar{\Gamma}$ in the fibred manifold $T\rho : TQ \longrightarrow TN$ along $\mathcal{H}$ by defining its horizontal distribution as follows:

$$\left(\frac{\partial}{\partial q^a}\right)^{\bar{\mathcal{H}}} = \frac{\partial}{\partial q^a} - \Gamma^i_a\frac{\partial}{\partial q^i} - \left(\dot{q}^b\frac{\partial \Gamma^i_b}{\partial q^a} - \Gamma^j_a\frac{\partial \Gamma^i_b}{\partial q^j}\right)\frac{\partial}{\partial \dot{q}^i}$$

$$\left(\frac{\partial}{\partial \dot{q}^a}\right)^{\bar{\mathcal{H}}} = \frac{\partial}{\partial \dot{q}^a} - \Gamma^i_a\frac{\partial}{\partial \dot{q}^i}.$$

Theorem 2.2. *The nonholonomic dynamics X_{nh} is a vector field on $\mathcal{H}$ which is $T\rho$-projectable onto Y^*. Furthermore, X_{nh} is the horizontal lift of Y^* with respect to the induced connection $\bar{\Gamma}$.*

Example 2.3 (Mobile robot with fixed orientation)**.** The body of the robot maintains a fixed orientation with respect to the environment. The robot has three wheels with radius R, which turn simultaneously about independent axes, and perform a rolling without sliding over a horizontal floor.

Let (x, y) denotes the position of the centre of mass, θ the steering angle of the wheel, ψ the rotation angle of the wheels in their rolling motion over the floor. So, the configuration manifold is $Q = S^1 \times S^1 \times \mathbb{R}^2$. The Lagrangian L is

$$L = \tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{2}m\dot{y}^2 + \tfrac{1}{2}J\dot{\theta}^2 + \tfrac{3}{2}J_\omega\dot{\psi}^2,$$

where m is the mass, J is the moment of inertia and J_ω is the axial moment of inertia of the robot.

The constraints are induced by the conditions that the wheels roll without sliding, in the direction in which they point, and that the instantaneous contact point of the wheels with the floor have no velocity component orthogonal to that direction,

$$\dot{x}\sin\theta - \dot{y}\cos\theta = 0, \qquad \dot{x}\cos\theta + \dot{y}\sin\theta - R\dot{\psi} = 0.$$

The Abelian group $G = \mathbb{R}^2$ acts on Q by translations, say

$$((a, b), (\theta, \psi, x, y)) \mapsto (\theta, \psi, a + x, b + y).$$

Therefore, we have a principal G-bundle $\rho : Q \longrightarrow N = Q/G$ with a principal connection given by the connection 1-form

$$\beta = (\mathrm{d}x - R\cos\theta\,\mathrm{d}\psi)e_1 + (\mathrm{d}y - R\sin\theta\,\mathrm{d}\psi)e_2,$$

where $\{e_1, e_2\}$ denotes the standard basis of $\mathbb{R}^2$. The constraints are given by the horizontal subspaces of β. If we apply the above reduction procedure we deduce $\alpha^* = 0$.

3. Geometric Hamilton–Jacobi theory

The following result is a geometric version of the standard formulation of the Hamilton–Jacobi problem [1].

Theorem 3.1. *Let γ be a closed 1-form on Q. Then the following conditions are equivalent:*

(i) for every curve $\sigma : \mathbb{R} \longrightarrow Q$ such that

$$\dot{\sigma}(t) = T\pi_Q(X_H(\gamma(\sigma(t))))$$

for all t, then $\gamma \circ \sigma$ is an integral curve of X_H.

(ii) $\mathrm{d}(H \circ \gamma) = 0$.

If $\gamma = \mathrm{d}W$ we recover the standard formulation since $\mathrm{d}(H \circ \mathrm{d}W) = 0$ is equivalent to the condition $H \circ \mathrm{d}W = cte$, that is,

$$H\left(q^A, \frac{\partial W}{\partial q^A}\right) = E,$$

where E is a constant.

A interesting new point of view of the geometric Hamilton–Jacobi theory has been recently developed by Cariñena *et al* [6].

Let γ be a closed 1-form as in theorem 3.1. Since FL is a diffeomorphism, we can define a vector field X on Q by

$$X = FL^{-1} \circ \gamma.$$

Therefore, we have

$$0 = \mathrm{d}(H \circ \gamma) = \mathrm{d}(E_L \circ FL^{-1} \circ \gamma) = \mathrm{d}(E_L \circ X)$$

because $H = E_L \circ FL^{-1}$.

Hence, theorem 3.1 can be reformulated as follows.

Theorem 3.2 [6]**.** *Let X be a vector field on Q such that $FL \circ X$ is a closed 1-form. Then the following conditions are equivalent:*

(i) for every curve $\sigma : \mathbb{R} \longrightarrow Q$ such that

$$\dot{\sigma}(t) = T\tau_Q(\xi_L(X(\sigma(t)))),$$

for all t, then $X \circ \sigma$ is an integral curve of ξ_L.

(ii) $\mathrm{d}(E_L \circ X) = 0$.

Definition 3.3. *A vector field X satisfying the conditions of theorem 3.2 will be called a solution for the Hamilton–Jacobi problem given by L.*

3.1. An interlude: mechanical systems with external forces

We shall need the following formulation of the Hamilton–Jacobi theory for mechanical systems with external forces.

A mechanical system with an external force is given by (see [8])

(i) A Lagrangian function $L : TQ \longrightarrow \mathbb{R}$, where Q is the configuration manifold;

(ii) a semibasic 1-form α on TQ.

Since α is semibasic (that means that α vanishes when it is applied to vertical tangent vectors) we have

$$\alpha = \alpha_A(q, \dot{q})\, \mathrm{d}q^A.$$

The Euler–Lagrange equations are then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A} = -\alpha_A, \qquad 1 \leqslant A \leqslant n, \tag{3.1}$$

which correspond to the symplectic equation

$$i_X\omega_L = \mathrm{d}E_L + \alpha. \tag{3.2}$$

Indeed, when L is regular, equation (3.2) has a unique solution $\xi_{L,\alpha}$ which is a second-order differential equation whose solutions are just the ones of (3.1).

Note that $\xi_{L,\alpha} = \xi_L + Z$, where $i_Z\omega_L = \alpha$.

Observe that we can construct the Hamiltonian counterpart using the Legendre transformation, so that we have a Hamiltonian $H = E_L \circ FL^{-1}$ subject to the external force $\beta = (FL^{-1})^*\alpha$ which is again semibasic (i.e. $\beta = \beta_A\, \mathrm{d}q^A$). The equation

$$i_{X_{H,\beta}}\omega_Q = \mathrm{d}H + \beta$$

has a unique solution $X_{H,\beta}$ whose integral curves satisfy the Hamilton equations with external force

$$\left.\begin{aligned} \dot{q}^A &= \frac{\partial H}{\partial p_A} \\ \dot{p}_A &= -\frac{\partial H}{\partial q^A} - \beta_A \end{aligned}\right\}. \tag{3.3}$$

Theorem 3.4. *Let γ be a closed 1-form on Q. Then the following conditions are equivalent:*

(i) for every curve $\sigma : \mathbb{R} \longrightarrow Q$ such that

$$\dot{\sigma}(t) = T\pi_Q(X_{H,\beta}(\gamma(\sigma(t)))), \tag{3.4}$$

for all t, then $\gamma \circ \sigma$ is an integral curve of $X_{H,\beta}$.

(ii) $\mathrm{d}(H \circ \gamma) = -\gamma^*\beta$.

Proof. Since $\gamma = \gamma_A \mathrm{d}q^A$ is closed then

$$\frac{\partial \gamma_A}{\partial q^B} = \frac{\partial \gamma_B}{\partial q^A}.$$

It is easy to show that equation (3.4) is rewritten, in local coordinates, as

$$\dot{\sigma}^A(t) = \frac{\partial H}{\partial p_A}(\sigma^B(t), \gamma_B(\sigma(t))). \tag{3.5}$$

We also have that condition

$$\mathrm{d}(H \circ \gamma) = -\gamma^*\beta$$

is written in local coordinates as

$$\frac{\partial H}{\partial q^A} + \frac{\partial H}{\partial p_B}\frac{\partial \gamma_B}{\partial q^A} = -\beta_A. \tag{3.6}$$

($\Longrightarrow$) Assume that (i) holds. Therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}(\gamma_A(\sigma(t))) = -\frac{\partial H}{\partial q^A}(\gamma(\sigma(t))) - \beta_A(\gamma(\sigma(t))). \tag{3.7}$$

Moreover

$$\begin{aligned}\frac{\partial H}{\partial q^A}+\frac{\partial H}{\partial p_B}\frac{\partial \gamma_B}{\partial q^A} &= \frac{\partial H}{\partial q^A}+\frac{\partial H}{\partial p_B}\frac{\partial \gamma_A}{\partial q^B}\ \text{(since } \gamma \text{ is closed)}\\ &= \frac{\partial H}{\partial q^A}+\dot\sigma^B(t)\frac{\partial \gamma_A}{\partial q^B}\quad \text{(from (3.5))}\\ &= -\beta_A \quad \text{(from (3.7))}.\end{aligned}$$

($\Longleftarrow$) Assume that (ii) holds, that is,

$$\frac{\partial H}{\partial q^A}+\frac{\partial H}{\partial p_B}\frac{\partial \gamma_B}{\partial q^A} = -\beta_A.$$

Now using (3.5) and since γ is closed, then

$$\frac{\partial H}{\partial q^A}+\dot\sigma^B(t)\frac{\partial \gamma_A}{\partial q^B} = -\beta_A.$$

Therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}(\gamma_A(\sigma(t))) = -\frac{\partial H}{\partial q^A}(\gamma(\sigma(t))) - \beta_A(\gamma(\sigma(t))),$$

which proves that $\gamma\circ\sigma$ is an integral curve of $X_{H,\beta}$. □

Therefore we have the Lagrangian version.

Theorem 3.5. *Let X be a vector field on Q such that $FL\circ X$ is a closed 1-form. Then the following conditions are equivalent:*

(i) for every curve $\sigma:\mathbb{R}\longrightarrow Q$ such that

$$\dot\sigma(t) = T\tau_Q(\xi_{L,\alpha}(X(\sigma(t))))$$

for all t, then $X\circ\sigma$ is an integral curve of $\xi_{L,\alpha}$.

(ii) $\mathrm{d}(E_L\circ X) = -X^*\alpha$.

Definition 3.6. *A vector field X satisfying the conditions of theorem 3.5 will be called a solution for the Hamilton–Jacobi problem given by L and α.*

4. Hamilton–Jacobi theory for nonholonomic mechanical systems

Let $L: TQ\longrightarrow\mathbb{R}$ be a Lagrangian function subject to nonholonomic constraints given by a vector subbundle M of TQ, locally defined by the linear constraints $\Phi^i=\Phi^i_A(q)\dot q^A$, $1\leqslant i\leqslant m$. Denote by D the distribution on Q whose annihilator is $D^0=\mathrm{span}\{\mu^i=\Phi^i_A(q)\,\mathrm{d}q^A\}$. Note that $S^*(TM^0)$ is the pullback to TQ of the annihilator D^0 of D.

We assume the compatibility conditions, and consider the Hamiltonian counterpart given by a Hamiltonian function $H: T^*Q\longrightarrow\mathbb{R}$ and a constraint submanifold $\bar M = FL(M)$ as in the precedent sections. X_{nh} and $\bar X_{nh}$ will denote the corresponding nonholonomic dynamics. Given D^0, the annihilator of D, we can form the algebraic ideal $\mathcal{I}(D^0)$ in the algebra $\Lambda^*(Q)$. Therefore, if a k-form $\nu\in\mathcal{I}(D^0)$ then

$$\nu=\beta_i\wedge\mu^i,\qquad \text{where}\qquad \beta_i\in\Lambda^{k-1}(Q),\quad 1\leqslant i\leqslant m.$$

Theorem 4.1. *Let γ be a 1-form on Q such that $\gamma(Q)\subset\bar M$ and $\mathrm{d}\gamma\in\mathcal{I}(D^0)$. Then the following conditions are equivalent:*

(i) for every curve $\sigma : \mathbb{R} \longrightarrow Q$ *such that*

$$\dot{\sigma}(t) = T\pi_Q(X_H(\gamma(\sigma(t)))) \tag{4.1}$$

for all t*, then* $\gamma \circ \sigma$ *is an integral curve of* $\bar{X}_{nh}$.

(ii) $\mathrm{d}(H \circ \gamma) \in D^0$.

Proof. The condition $\mathrm{d}\gamma \in \mathcal{I}(D^0)$ means that

$$\frac{\partial \gamma_A}{\partial q^B} = \frac{\partial \gamma_B}{\partial q^A} + \beta_{iA}\Phi^i_B - \beta_{iB}\Phi^i_A,$$

where $\gamma = \gamma_A \, \mathrm{d}q^A$ and $\beta_i = \beta_{iA}\, \mathrm{d}q^A$. It is easy to show that equation (4.1) is rewritten, in local coordinates, as

$$\dot{\sigma}^A(t) = \frac{\partial H}{\partial p_A}(\sigma^B(t), \gamma_B(\sigma(t))). \tag{4.2}$$

We also have that condition

$$\mathrm{d}(H \circ \gamma) \in D^0$$

is written in local coordinates as

$$\left[\frac{\partial H}{\partial q^A} + \frac{\partial H}{\partial p_B}\frac{\partial \gamma_B}{\partial q^A}\right]\mathrm{d}q^A = \tilde{\lambda}_i \mu^i = \tilde{\lambda}_i \Phi^i_A(q)\,\mathrm{d}q^A \tag{4.3}$$

for some Lagrange multipliers $\tilde{\lambda}_i$'s.

($\Longrightarrow$) Assume that (i) holds. Therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}(\gamma_A(\sigma(t))) = -\frac{\partial H}{\partial q^A}(\gamma(\sigma(t))) - \bar{\lambda}_i \Phi^i_A(\sigma(t)), \tag{4.4}$$

where the $\bar{\lambda}_i$'s are determined using the constraint equations

$$\Psi^i(\sigma(t), \gamma(\sigma(t))) = \frac{\partial H}{\partial p_B}(\sigma(t), \gamma(\sigma(t)))\Phi^i_B(\sigma(t)) = 0.$$

Using the constraint equations we deduce that

$$\begin{aligned}
\frac{\partial H}{\partial q^A} + \frac{\partial H}{\partial p_B}\frac{\partial \gamma_B}{\partial q^A} &= \frac{\partial H}{\partial q^A} + \frac{\partial H}{\partial p_B}\frac{\partial \gamma_A}{\partial q^B} + \frac{\partial H}{\partial p_B}\beta_{iA}\Phi^i_B - \frac{\partial H}{\partial p_B}\beta_{iB}\Phi^i_A \\
&= \frac{\partial H}{\partial q^A} + \dot{\sigma}^B(t)\frac{\partial \gamma_A}{\partial q^B} - \frac{\partial H}{\partial p_B}\beta_{iB}\Phi^i_A \\
&= -\left(\bar{\lambda}_i + \frac{\partial H}{\partial p_B}\beta_{iB}\right)\Phi^i_A \quad \text{(from (4.3))}.
\end{aligned}$$

Therefore, we conclude that $\mathrm{d}(H \circ \gamma) \in D^0$.

($\Longleftarrow$) Assume that (ii) holds, that is,

$$\left[\frac{\partial H}{\partial q^A} + \frac{\partial H}{\partial p_B}\frac{\partial \gamma_B}{\partial q^A}\right]\mathrm{d}q^A = \tilde{\lambda}_i \mu^i.$$

Now using (4.1) and since $\mathrm{d}\gamma \in \mathcal{I}(D^0)$, then

$$\frac{\partial H}{\partial q^A} + \dot{\sigma}^B(t)\left(\frac{\partial \gamma_A}{\partial q^B} - \beta_{iA}\Phi^i_B + \beta_{iB}\Phi^i_A\right) = \tilde{\lambda}_i \Phi^i_A.$$

Therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}(\gamma_A(\sigma(t))) = -\frac{\partial H}{\partial q^A}(\gamma(\sigma(t))) + (\tilde{\lambda}_i - \dot{\sigma}^B(t)\beta_{iB}(\sigma(t)))\Phi^i_A(\sigma(t)). \tag{4.5}$$

Using that $\mathrm{Im}(\gamma) \in \bar{M}$, we deduce that $\bar{\lambda}_i = \frac{\partial H}{\partial p_B}\beta_{iB} - \tilde{\lambda}_i$ along γ. □

Remark 4.2. Suppose that $\gamma = \mathrm{d}S$ where S is a function $S : Q \longrightarrow \mathbb{R}$. In this case, the condition $\mathrm{d}\gamma \in \mathcal{I}(D^0)$ is trivially satisfied. Moreover, we note that in previous approximations to Hamilton–Jacobi theory [7, 15, 17–19] the considered sections are of the form

$$\gamma(q) = \left(q, \frac{\partial S}{\partial q^A} - \tilde{\lambda}_i \mu_A^i\right), \tag{4.6}$$

and the coefficients $\tilde{\lambda}_i$ are determined through the nonholonomic constraint equations

$$\mu_A^i(q)\frac{\partial H}{\partial p_A}(q, \gamma_A(q)) = 0.$$

In general, this type of 1-forms does not satisfy the condition that we initially impose, $\mathrm{d}\gamma \in \mathcal{I}(D^0)$. Observe that in the particular case of holonomic constraints both approaches coincide.

Now, we write a coordinate expression for the Hamilton–Jacobi equation that we have proposed. In order to do it, consider a set of independent vector fields $\left\{Z_a = Z_a^A \frac{\partial}{\partial q^A}\right\}$, $1 \leqslant a \leqslant n - m$, on Q such that $\mu^i(Z_a) = 0$, i.e, $D_q = \mathrm{span}\,\{(Z_a)|_q\}$. Thus a 1-form γ on Q, solution of the nonholonomic Hamilton–Jacobi equation, must verify the condition $\mathrm{d}\gamma \in \mathcal{I}(D^0)$ and, additionally,

$$Z_a^A(q)\left(\frac{\partial H}{\partial q^A}(q, \gamma(q)) + \frac{\partial H}{\partial p_B}(q, \gamma(q))\frac{\partial \gamma_B}{\partial q^A}(q)\right) = 0,$$
$$\mu_A^i(q)\frac{\partial H}{\partial p_A}(q, \gamma(q)) = 0,$$

for the condition $\mathrm{d}(H \circ \gamma) \in D^0$ and for the condition $\gamma(Q) \subset \bar{M}$, correspondingly.

Theorem 4.3. *Let X be vector field on Q such that $X(Q) \subset M$ and $\mathrm{d}(FL \circ X) \in \mathcal{I}(D^0)$. Then the following conditions are equivalent:*

(i) for every curve $\sigma : \mathbb{R} \longrightarrow Q$ such that

$$\dot{\sigma}(t) = T\tau_Q(X_{nh}(X(\sigma(t)))) \tag{4.7}$$

for all t, then $X \circ \sigma$ is an integral curve of X_{nh}.

(ii) $\mathrm{d}(E_L \circ X) \in D^0$.

Definition 4.4. *A vector field X satisfying the conditions of theorem 4.1 will be called a solution for the Hamilton–Jacobi problem given by L and M.*

4.1. An application to Čaplygin systems

Consider now the case of a Čaplygin system (see section 2.3). That is, we have a fibration $\rho : Q \longrightarrow N$, and an Ehresmann connection Γ in ρ, whose horizontal distribution imposes the constraints to a Lagrangian $L : TQ \longrightarrow \mathbb{R}$.

Let $L^* : TN \longrightarrow \mathbb{R}$ be the reduced Lagrangian and α^* the corresponding external force. We denote by X_{nh} the nonholonomic vector field on TQ and by X^* the solution of the reduced Lagrangian system with external force α^*.

Theorem 4.5

(i) Assume that a vector field X on Q is a solution for the Hamilton–Jacobi problem given by L and Γ. If X is ρ-projectable to a vector field Y on N and $\mathrm{d}(FL^* \circ Y) = 0$ then Y is a solution of the Hamilton–Jacobi problem given by L^* and α^*.

(ii) *Conversely, let Y be a vector field which is a solution of the Hamilton–Jacobi problem given by L^* and α^*. Then, if* $\mathrm{d}(FL \circ Y^{\mathcal{H}}) \in \mathcal{I}(\mathcal{H}^0)$, *the horizontal lift $Y^{\mathcal{H}}$ is a solution for the Hamilton–Jacobi problem given by L and ρ.*

Proof. ($\Longrightarrow$)

Assume that a vector field X on Q is a solution for the Hamilton–Jacobi problem given by L and Γ, and that X is ρ-projectable onto a vector field Y on N. We have to prove that Y is then a solution of the Hamilton problem given L^* and α^*. Let μ a curve in N such that

$$\dot{\mu}(t) = T\tau_N(Y^*(Y(\mu(t)))) \tag{4.8}$$

for all t. Take an horizontal lift σ of μ to Q with respect to the connection Γ. A direct computation shows that

$$\dot{\sigma}(t) = T\tau_Q(X_{nh}(X(\sigma(t)))), \tag{4.9}$$

since X_{nh} is the horizontal lift of Y^* with respect to the prolongated connection $\bar{\Gamma}$. Therefore, we have that $X \circ \sigma$ is an integral curve of X_{nh} and, consequently, $Y \circ \mu$ is an integral curve of Y^*.

($\Longleftarrow$)

Assume that Y is vector field on N which is a solution of the Hamilton–Jacobi problem given by L^* and α^*. Take its horizontal lift $X = Y^{\mathcal{H}}$ to Q with respect to Γ. If σ is a curve in Q satisfying

$$\dot{\sigma}(t) = T\tau_Q(X_{nh}(X(\sigma(t)))), \tag{4.10}$$

then the projection $\mu = \rho \circ \sigma$ satisfies (4.8). So, $Y \circ \mu$ is an integral curve of Y^* and, hence $X \circ \sigma$ is an integral curve of X_{nh}. □

Example 4.6 (The mobile robot with fixed orientation revisited)**.** The reduced Lagrangian in this case is

$$L^*(\theta, \psi) = \frac{1}{2}J\dot{\theta}^2 + \frac{mR^2 + 3J_\omega}{2}\dot{\psi}^2,$$

and $\alpha^* = 0$.

Therefore,

$$Y_1 = \frac{\partial}{\partial\theta} \qquad \text{and} \qquad Y_2 = \frac{\partial}{\partial\psi}$$

are solutions of the Hamilton–Jacobi problem given by (L^*, α^*). Calculating the horizontal lifts off both vector fields we have that

$$Y_1^{\mathcal{H}} = \frac{\partial}{\partial\theta} \qquad \text{and} \qquad Y_2^{\mathcal{H}} = \frac{\partial}{\partial\psi} + R\cos\theta\frac{\partial}{\partial x} + R\sin\theta\frac{\partial}{\partial y}.$$

Now

$$\gamma_1 = FL \circ Y_1^{\mathcal{H}} = J\,\mathrm{d}\theta$$
$$\gamma_2 = FL \circ Y_2^{\mathcal{H}} = 3J_\omega\,\mathrm{d}\psi + mR\cos\theta\,\mathrm{d}x + mR\sin\theta\,\mathrm{d}y$$

and

$$\mathrm{d}\gamma_1 = 0 \in \mathcal{I}(\mathcal{H}^0)$$
$$\mathrm{d}\gamma_2 = -mR\,\mathrm{d}\theta \wedge (\sin\theta\,\mathrm{d}x - \cos\theta\,\mathrm{d}y) \in \mathcal{I}(\mathcal{H}^0).$$

Therefore, $Y_1^{\mathcal{H}}$ and $Y_2^{\mathcal{H}}$ are solutions of the Hamilton–Jacobi problem of the nonholonomic problem given by $(L, \mathcal{H})$. Observe that in both cases $\mathrm{d}(H \circ \gamma_i) = 0$, for $i = 1, 2$. In such a case,

$$t \longmapsto (x_0, y_0, t + \theta_0, \psi_0)$$
$$t \longmapsto (tR\cos\theta_0 + x_0, tR\sin\theta_0 + y_0, \theta_0, t + \psi_0)$$

are the solutions of the nonholonomic system $(L, \mathcal{H})$ obtained from $Y_1^{\mathcal{H}}$ and $Y_2^{\mathcal{H}}$, respectively. Both solutions are also solutions of the Lagrangian system determined by L without constraints; indeed, they are solutions of the free system satisfying additionally the nonholonomic constraints.

But taking now the vector field

$$Y_3 = Y_1 + Y_2 = \frac{\partial}{\partial\theta} + \frac{\partial}{\partial\psi},$$

it is obviously a solution of the Hamilton–Jacobi equations for the Lagrangian L^* and its horizontal lift

$$Y_3^{\mathcal{H}} = \frac{\partial}{\partial\theta} + \frac{\partial}{\partial\psi} + R\cos\theta\frac{\partial}{\partial x} + R\sin\theta\frac{\partial}{\partial y}$$

is a solution of the Hamilton–Jacobi equations for the nonholonomic system (L, H),

$$\gamma_3 = FL \circ Y_3^{\mathcal{H}} = J\,\mathrm{d}\theta + 3J_\omega\,\mathrm{d}\psi + mR\cos\theta\,\mathrm{d}x + mR\sin\theta\,\mathrm{d}y$$

and $\mathrm{d}\gamma_3 \in \mathcal{I}(\mathcal{H}^0)$. In such a case, the solution of the nonholonomic problem that we obtain is

$$t \longmapsto (R\sin(t - \theta_0) + x_0 + R\sin\theta_0, -R\cos(t - \theta_0) + y_0 + R\cos\theta_0, t + \theta_0, t + \psi_0),$$

which is a solution of the nonholonomic problem but not of the free system.

Acknowledgments

The authors would like to thank the referees for the careful reading of the paper and the interesting remarks, which have highly improved the content of the paper.

References

[1] Abraham R and Marsden J E 1978 *Foundations of Mechanics* 2nd edn (New York: Benjamin)
[2] Arnold V I 1978 *Mathematical Methods of Classical Mechanics (Graduate Texts in Mathematics* vol 60*)* (Berlin: Springer)
[3] Bloch A M 2003 Nonholonomic Mechanics and Control *Systems and Control (Interdisciplinary Applied Mathematics* vol 24*)* (New York: Springer)
[4] Bates L and Cushman R 1999 What is a completely integrable nonholonomic dynamical system? *Proc. XXX Symp. on Mathematical Physics (Toruń, 1998) (Rep. Math. Phys.* 44*)* 29–35
[5] Cantrijn F, de León M and Martín de Diego D 1999 On almost-poisson structures in nonholonomic mechanics *Nonlinearity* **12** 721–37
[6] Cariñena J F, Gracia X, Marmo G, Martínez E, Muñoz-Lecanda M and Román-Roy N 2006 Geometric Hamilton–Jacobi theory *Int. J. Geom. Methods Mod. Phys.* **3** 1417–58
[7] Eden R J 1951 The Hamiltonian dynamics of non-holonomic systems *Proc. R. Soc.* A **205** 564–83
[8] Godbillon Ch 1969 *Géométrie Différentielle et Mécanique Analytique* (Paris: Hermann)
[9] Hairer E, Lubich C and Wanner G 2002 *Geometric Numerical Integration, Structure-Preserving Algorithms for Ordinary Differential Equations (Springer Series in Computational Mathematics* vol 31*)* (Berlin: Springer)
[10] Kozlov V V 2002 On the integration theory of equations of nonholonomic mechanics *Regular Chaotic Dyn.* **7** 161–76
[11] de León M and Martín de Diego D 1996 Solving non-holonomic Lagrangian dynamics in terms of almost product structures *Extracta Math.* **11** 325–47

[12] de León M and Martín de Diego D 1996 On the geometry of non-holonomic Lagrangian systems *J. Math. Phys.* **37** 3389–414

[13] de León M, Martín de Diego D and Santamaría-Merino A 2004 Geometric integrators and nonholonomic mechanics *J. Math. Phys.* **45** 1042–64

[14] de León M and Rodrigues P R 1989 *Methods of Differential Geometry in Analytical Mechanics* (Amsterdam: North-Holland)

[15] Pavon M 2005 Hamilton–Jacobi equations for nonholonomic dynamics *J. Math. Phys.* **46** 032902

[16] Prince G E, Aldridge J E and Byrnes G B 1999 A universal Hamilton–Jacobi equation for second-order ODEs *J. Phys. A: Math. Gen.* **32** 827–44

[17] Rumyantsev V V 2002 Forms of Hamilton's principle for nonholonomic systems *Mech. Autom. Control Robot.* **2** 1035–48

[18] Sumbatov A S 2002 Nonholonomic systems *Regular Chaotic Dyn.* **7** 221–38

[19] Van Dooren R 1975 The generalized Hamilton–Jacobi method for non-holonomic dynamical systems of Chetaev's type *Z. Angew. Math. Mech.* **55** 407–11

[20] Van Dooren R 1976 Motion of a rolling disc by a new generalized Hamilton–Jacobi method *J. Appl. Math. Phys.* **27** 501–5

[21] Van Dooren R 1978 Second form of the generalized Hamilton–Jacobi method for nonholonomic dynamical systems *J. Appl. Math. Phys.* **29** 828–34

[22] Van Dooren R 1979 On the generalized Hamilton–Jacobi method for nonholonomic dynamical systems *Dienst Anal. Mech.* **Tw** 1–6