Lagrangian systems with constraints: a geometric approach to the method of Lagrange multipliers

1993 J. Phys. A: Math. Gen. 26 1335

(http://iopscience.iop.org/0305-4470/26/6/016)

J. Phys. A: Math. Gen. **26** (1993) 1335–1351. Printed in the UK

Lagrangian systems with constraints: a geometric approach to the method of Lagrange multipliers

José F Cariñena and Manuel F Rañada
Departamento de Física Teórica, Facultad de Ciencias, Universidad de Zaragoza, 50009 Zaragoza, Spain

Received 15 July 1992, in final form 19 October 1992

Abstract. A geometric approach to the method of Lagrange multipliers is presented using the framework of the tangent bundle geometry. The non-holonomic systems with constraint functions linear in the velocities are studied in the first place and then, and using this study of the non-holonomic mechanical systems as a previous result, the holonomic systems are considered. The Lagrangian inverse problem is also analysed and, finally, the theory is illustrated with several examples.

1. Introduction

One of the classical problems of theoretical mechanics is the study of constrained Lagrangian systems. The expression 'constrained systems' has two different meanings in Lagrangian mechanics: it either refers to systems described by a singular Lagrangian for which the Legendre transformation is not a diffeomorphism [1–3] or to the case where there exist some constraints allowing only some particular motions. Concerning the second case (that is, when the Lagrangian L is regular), although it is usually assumed that the Lagrangian function must contain all the relevant information on the system, this is not always the case, and often there are some relations among the coordinates and the velocities which express the presence of forces of constraint on the system. The Lagrangian does not necessarily contain the characteristics of these additional forces and, because of this, it is usually considered that one must leave the Hamilton principle and look for another, more appropriate starting point (d'Alembert principle of virtual works or something similar).

It is known that, when the Lagrangian is regular, one possible way for incorporating the constraint functions in the equations of motion is the use of the so-called Lagrange multipliers. This approach is classical and it provides a method for dealing with systems that has been proved to be successful; however a rigorous study of the underlying mathematical structure has not yet been adequately developed.

The situation can be summarized as follows [4–6]: if $L(q,v)$ is the Lagrangian of the system without constraints, the equations of the motion for the system in presence of the constraint ϕ, are

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^i}\right)-\frac{\partial L}{\partial q^i}=\Lambda_i(q,v,\lambda)\qquad i=1,\ldots,n \tag{1}$$

where the functions $\Lambda_i(q, v, \lambda)$ are

$$\Lambda_i(q, v, \lambda) = \lambda \frac{\partial \phi}{\partial q^i} \tag{2a}$$

when ϕ is holonomic, and

$$\Lambda_i(q, v, \lambda) = \lambda \frac{\partial \phi}{\partial v^i} \tag{2b}$$

when ϕ is non-holonomic. It is important to emphasize that these equations are correct, in the sense that it is assumed that their predictions are in agreement with the physical measures, but they are usually presented as result of a 'recipe' and not as arising as a consequence of basic principles.

From the point of view of the calculus of variations the main point is the question of the existence of an appropriate Lagrangian leading, via the Hamilton principle (or other related variational principle), to the above equations. It has often been claimed that such Lagrangian function must be $L + \lambda\,\phi$. Nevertheless this function poses two problems: the behaviour of λ under the variations is not clear and, although admissible for holonomic systems, its result is clearly incorrect for non-holonomic systems (see [7–13] and references therein). Recent research has considered this matter within the context of differential geometric techniques in theoretical mechanics (see [14–18] for some recent papers).

Nowadays it is known that the symplectic geometry in the tangent bundle of a manifold is the appropriate geometric setting for the description of Lagrangian autonomous systems [19–22]. This geometric approach is considered as more fundamental than the traditional one, mainly for the following two reasons: (1) the theorems and properties are proved using an intrisic or coordinate-free formulation and, because of this, they are valid for topologically non-trivial configuration spaces and are ready for a possible generalization to the infinite-dimensional case; (2) all the previously known properties of the traditional approach are recovered here as particular cases of these new and more general results.

Concerning the problem of the Lagrange multipliers, there are basically two different approaches:

(i) The presence of Lagrange multipliers is related with the existence of a Lagrangian $L' \neq L$ in an extended space, but this new Lagrangian is singular. In geometric terms, this means that it would be necessary to leave the symplectic setting and then to use the tools of the presymplectic geometry [23–27].

(ii) The original (regular) Lagrangian is the appropriate Lagrangian but the presence of constraints introduce new terms in the equations that can be considered as perturbations. This suggests to us that it would be convenient to develop a geometric model for the description of dynamical systems represented by Euler–Lagrange vector fields in tangent bundles that are submitted to perturbations of (probably) non-Lagrangian character.

The purpose of this paper is to carry out the geometric study of the second approach. We will first introduce, in an intrinsic or coordinate-free formulation, some basic principles, and then prove that they (i) lead to the equations of the constrained Lagrangian systems and (ii) give a geometric interpretation for the Lagrange multipliers. One of our aims in this paper is to prove that this geometric method covers, at the same time, both the holonomic and the non-holonomic systems (these two cases have been traditionally considered as completely different). Finally we notice that, in differential geometric terms, it is usual to study the Hamiltonian formalism in the first place and only then go into the Lagrangian approach (see, e.g., [15] and [19]). In this paper the Lagrangian symplectic theory is directly presented (see, e.g., [20,21] and [28,29]) and the properties of the constraints are directly presented inside the framework of the tangent bundle geometry.

The paper is organized as follows: in section 2 we first present the notation that will be used throughout the paper and then we study the non-holonomic case. The properties of the the vector field representing the constrained dynamics are discussed in section 3 and the ensuing sections consider the holonomic case, some examples and, finally, the Hamiltonian formulation.

2. Geometric formalism and non-holonomic constraints

Let Q be a differentiable manifold (of dimension n), TQ its tangent bundle, $\mathcal{X}(TQ)$ the set of smooth vector fields on TQ and $\tau : TQ \to Q$ the canonical tangent bundle projection. We will denote by $\{q^i; i = 1, \ldots, n\}$ a set of local coordinates in Q and by $\{q^i, v^i; i = 1, \ldots, n\}$ the associated coordinates in TQ.

The tangent bundle TQ possesses two important geometrical objects: the Liouville vector field and the vertical endomorphism [21,29]. The Liouville vector field Δ is the vector field $\Delta \in \mathcal{X}(TQ)$ generating dilations along the fibres. The vertical endomorphism S is a $(1,1)$ tensor field $S : \mathcal{X}(TQ) \to \mathcal{X}(TQ)$. In terms of the coordinates $\{q^i, v^i; i = 1, \ldots, n\}$, Δ and S have the form

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

We will use the notation S^* instead of S when it acts on $\bigwedge^1(TQ)$.

Suppose that a Lagrangian is given, i.e. a differentiable function L on TQ. Then one can construct a semibasic 1-form $\theta_L \in \bigwedge^1(TQ)$ (called the Cartan form), an exact 2-form $\omega_L \in \bigwedge^2(TQ)$ and an energy function $E_L \in C^\infty(TQ)$ by

$$\theta_L = S^*(\mathrm{d}L) \qquad \omega_L = -\mathrm{d}\theta_L \qquad E_L = \Delta(L) - L\,.$$

In coordinates they read

$$\theta_L = \frac{\partial L}{\partial v^i}\,\mathrm{d}q^i \qquad \omega_L = \frac{\partial^2 L}{\partial q^j\,\partial v^i}\,\mathrm{d}q^i \wedge \mathrm{d}q^j + \frac{\partial^2 L}{\partial v^j\,\partial v^i}\,\mathrm{d}q^i \wedge \mathrm{d}v^j \qquad E_L = v^i \frac{\partial L}{\partial v^i} - L\,.$$

If the Lagrangian L is regular then ω_L is symplectic. The dynamics is then represented by the flow (on TQ) of the vector field $X_L \in \mathcal{X}(TQ)$ determined by the equation

$$\mathrm{i}(X_L)\omega_L = \mathrm{d}E_L\,.$$

The solution X_L of this equation is uniquely determined—it turns out to be a second-order differential equation (hereafter shortened to SODE) vector field, i.e. $S(X_L) = \Delta$, and the curves on the Q projection of its integral curves satisfy the

Euler–Lagrange equations; because of this X_L is called the Euler–Lagrange vector field. In coordinates X_L takes the form

$$X_L = v^i \frac{\partial}{\partial q^i} + f^i(q,v)\frac{\partial}{\partial v^i} \qquad f^i(q,v) = W^{ij}\left[\frac{\partial L}{\partial q^j} - \frac{\partial^2 L}{\partial q^k \partial v^j}\, v^k\right]$$

where $f^i(q,v)$ are the Lagrangian forces (by contrast with the constraint forces) and W^{kj} is the inverse matrix of the Hessian matrix W_{ij} whose entries are the second derivatives of L with respect to the n velocities v^i, $i = 1, \ldots, n$.

Suppose that a system described by a regular Lagrangian is subjected to a force of constraint represented by the constraint function $\phi \in C^\infty(TQ)$. We distinguish two situations:

(i) General case—ϕ is a velocity-dependent function, i.e. $\phi = \phi(q,v)$. Then the dynamics is restricted to the submanifold $M_\phi \subset TQ$ defined by $M_\phi = \phi^{-1}(0)$. This submanifold, which represents the phase space of the constrained system, is of co-dimension 1 and does not have structure of tangent bundle. The Euler–Lagrange vector field X_L is defined in TQ but the constrained dynamics is represented by a vector field Γ_ϕ in M_ϕ (that is, $\Gamma_\phi \in \mathcal{X}(M_\phi)$). The functional form of ϕ determines the geometric properties of M_ϕ. If ϕ is linear in the fibre coordinates (velocities), then it affects only the fibres and does not restrict the configuration space Q. If ϕ is not only linear but also homogeneous then M_ϕ is a vector sub-bundle of TQ and if ϕ has a non-homogeneous term then $M_\phi \subset TQ$ is an affine sub-bundle. Thus, if ϕ is linear, the image of M_ϕ by the projection τ covers Q, i.e. $\tau(M_\phi) = Q$. Consequently, the curves $\{q = q(t), t \in \mathbb{R}\}$ on the Q projection of the integral curves of Γ_ϕ are defined in the whole configuration space Q.

(ii) Particular case—ϕ is a velocity-independent function, i.e. $\phi = \phi(q)$. Then the constraint determines a submanifold Q_ϕ of Q defined by $Q_\phi = \phi^{-1}(0) \subset Q$, and a primary constraint submanifold $M^{(1)}$ of TQ given by $M^{(1)} = \tau^{-1}(Q_\phi) \subset TQ$. Nevertheless, in this case there is a secondary constraint $D(\phi) = 0$ where D is any SODE vector field and, therefore, the final submanifold $M_\phi \subset M^{(1)}$ takes the form

$$M_\phi = \left\{(q,v) \in TQ \mid \phi(q) = 0 \qquad \frac{\partial \phi}{\partial q^i}\, v^i = 0\right\}$$

that is, $M_\phi = T(Q_\phi)$. Consequently, in this case the phase space of the constrained system is of co-dimension 2 and does have structure of tangent bundle. Moreover, since M_ϕ is a tangent bundle, the image of M_ϕ by the projection τ is the base space Q_ϕ. Consequently the integral curves of Γ determines curves $\{q = q(t), t \in \mathbb{R}\}$ defined only in Q_ϕ.

Remark that in a tangent bundle the property of a function being linear in the fibre coordinates (velocities) is an intrinsic property, that is, it is a property that is preserved under point transformations. Moreover if $\phi \in C^\infty(TQ)$ takes the form $\phi = \phi_i(q)v^i$ in a particular set of local coordinates then, under the change of coordinates $q'^i = q'^i(q)$, we have $\phi_i(q)v^i = \phi'_i(q')(\partial q'^i/\partial q^j)v^j$. Therefore the functions $\phi_i(q)$ transform as the components of an associate basic 1-form $\alpha \in \bigwedge^1(TQ)$, $\alpha = \phi_i(q)\,\mathrm{d}q^i$. There is thus a one-to-one linear correspondence between basic 1-forms and linear homogeneous functions.

In what follows we will denote by $\widehat{\alpha}$ and $\widetilde{h}$ two functions defined in TQ in the following form: if $\alpha \in \bigwedge^1(Q)$ is a 1-form defined in Q then $\widehat{\alpha} \in C^\infty(TQ)$ represents a function linear in the fibres defined by $\widehat{\alpha}(q,v) = \langle \alpha_q, v \rangle$, and if $h \in C^\infty(Q)$ is a function defined in Q then $\widetilde{h} \in C^\infty(TQ)$ represents the pullback of h through the tangent projection, i.e. $\widetilde{h} = \tau^* h$. It must be remarked that all these geometric properties have as physical counterpart the fact that, until now, the non-holonomic mechanical systems known are just linear in the velocities [17].

Suppose that a system described by a regular Lagrangian L on TQ is subjected to a linear constraint represented by the function $\phi = \widehat{\alpha} + \widetilde{h}$. In coordinates

$$\phi(q,v) = \alpha_j(q) v^j + h(q) .$$

When only the h-term is present, i.e. $\phi = \widetilde{h}$, the constraint is holonomic. In this section we will study the non-holonomic case; section 4 will consider the holonomic case.

The dynamics of the constrained system will be represented by a vector field Γ_ϕ in M_ϕ. In order to obtain this vector field we first consider the following system of two equations:

$$\mathrm{i}(X_L)\omega_L = \mathrm{d}E_L \tag{3a}$$

$$\mathrm{i}(Z)\omega_L = \widetilde{\alpha} \tag{3b}$$

where $\widetilde{\alpha} = \tau^* \alpha$.

The function E_L is still the energy function of the Lagrangian L (that we can now call the free Lagrangian) and X_L is its associate ω_L-Hamiltonian vector field. Concerning the second equation, the general case corresponds to α being non-closed and, consequently, to Z being a non-Hamiltonian vector field (neither locally Hamiltonian); only in the particular case of α being closed, which corresponds to an integrable constraint, Z is Hamiltonian. In any case Z is uniquely determined by α.

In the following we will denote by $\mathcal{X}^v(TQ)$ the set of all the vertical vector fields (i.e. fields tangent to the fibres)

$$\mathcal{X}^v(TQ) = \{X \in \mathcal{X}(TQ) \mid S(X) = 0\}$$

by $\mathcal{X}^2(TQ)$ the set of the SODE vector fields

$$\mathcal{X}^2(TQ) = \{D \in \mathcal{X}(TQ) \mid S(D) = \Delta\}$$

and by $\bigwedge^1_{\mathrm{sb}}(TQ)$ the set of the semibasic 1-forms

$$\bigwedge\nolimits^1_{\mathrm{sb}}(TQ) = \{\beta \in \bigwedge\nolimits^1(TQ) \mid S^*(\beta) = 0\} .$$

Notice that in geometric terms the forces are represented by semibasic 1-forms, so α can properly be called the constraint force.

When L is regular, ω_L is non-degenerate and the map $\widehat{\omega}_L : \mathcal{X}(TQ) \to \bigwedge^1(TQ)$, $X \to \widehat{\omega}_L(X) = \mathrm{i}(X)\omega_L$ is a bijection. The important point is that the restriction of $\widehat{\omega}_L$ to $\mathcal{X}^v(TQ)$ determines an isomorphism between $\mathcal{X}^v(TQ)$ and $\bigwedge^1_{\mathrm{sb}}(TQ)$; because of this the field Z is vertical.

In coordinates Z is

$$Z = -z^i(q,v)\frac{\partial}{\partial v^i} \qquad z^i(q,v) = W^{ij}\,\alpha_j .$$

Proposition 2.1. Let X_α be the Hamiltonian vector field of the function $\widehat{\alpha}$; then $Z = -S(X_\alpha)$.

Proof. First notice that $\widehat{\alpha}$ and $\widetilde{\alpha}$ are related by $S^*(\mathrm{d}\widehat{\alpha}) = \widetilde{\alpha}$. Thus

$$S^*[\mathrm{i}(X_\alpha)\omega_L] = \mathrm{i}(Z)\omega_L$$

but S satisfies [29]

$$S^*[\mathrm{i}(Y)\omega_L] + \mathrm{i}(S(Y))\omega_L = 0$$

for any vector field Y. Hence

$$\widetilde{\alpha} = \mathrm{i}(Z)\omega_L = -\mathrm{i}(S(X_\alpha))\omega_L$$

and from this, the equality $Z = -S(X_\alpha)$ follows.

Using this property the function $Z(\widehat{\alpha})$ can be written as

$$Z(\widehat{\alpha}) = \mathrm{i}(Z)\mathrm{i}(X_\alpha)\omega_L = \omega_L(X_\alpha, Z) = \omega_L(S(X_\alpha), X_\alpha)$$

Let λ be a function on TQ and Γ_λ the associated λ-dependent family of vector fields defined by

$$\Gamma_\lambda = X_L - \lambda Z \qquad \lambda \in C^\infty(TQ)$$

that is, Γ_λ is solution of

$$\mathrm{i}(\Gamma_\lambda)\omega_L = \mathrm{d}E_L - \lambda\,\widetilde{\alpha}\,. \tag{4}$$

In coordinates Γ_λ takes the form

$$\Gamma_\lambda = v^i \frac{\partial}{\partial q^i} + F^i(q, v, \lambda)\frac{\partial}{\partial v^i}$$

where

$$F^i(q, v, \lambda) = f^i(q, v) + \lambda\, W^{ij}\, \alpha_j$$

and its integral curves are the solutions of

$$\frac{\mathrm{d}}{\mathrm{d}t}\, q^i = v^i \qquad \frac{\mathrm{d}}{\mathrm{d}t}\, v^i = f^i(q, v) + \lambda\, W^{ij}\, \alpha_j$$

or

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\, q^i = f^i\left(q, \frac{\mathrm{d}q}{\mathrm{d}t}\right) + \lambda\, W^{ij}\, \alpha_j$$

that is, the equations to be satisfied by the integral curves of the λ-family Γ_λ are the second-order differential equations (1), (2*b*) of the 'classical' recipe.

The constrained dynamics is given by the restriction of that vector field Γ in the family Γ_λ that is tangent to the restricted phase space. Consequently $\lambda = \lambda(q, v)$ represents a function to be determined.

Therefore we assume that the dynamics is represented by a non-Lagrangian vector field; nevertheless as Z is vertical, Γ remains SODE. The situation is summarized in the following map:

$$C^\infty(TQ) \times \mathcal{X}^v(TQ) \longrightarrow \mathcal{X}^2(TQ) \qquad (\lambda, Z) \to X_L - \lambda Z\,.$$

Thus the vector field Z can be considered as a non-Lagrangian perturbation of the Euler–Lagrange vector field and the function λ, that couples the two equations (3*a*), (3*b*), represents the 'intensity' of the perturbation. In summary, the two fundamental characteristics are (i) the perturbation is vertical and (ii) its intensity is not given by a small real parameter but by a function, that is, the value of the perturbation depends on the point.

Remarks. (1) Since X_L and Z are defined in the whole of TQ, so is Γ_λ. Later, when we consider the restricted phase space M_α, as this submanifold is not a tangent bundle, the property of a vector field being SODE in M_α must be interpreted as being the restriction to M_α of a vector field tangent to M_α that is SODE in TQ.

(2) Given a regular Lagrangian L the 1-form $\beta \in \bigwedge^1(TQ)$ defined by $\beta = \mathrm{i}(X)\omega_L - \mathrm{d}E_L$ is semibasic if and only if the field X is SODE (this statement can be easily proved in coordinates; for an intrinsic (coordinate-free) proof see [29]; see also [22]). This property has been used in the geometric approach to the study of dissipative systems [30, 31] where the original Lagrangian equation is modified by the addition of a semibasic 1-form μ. Consequently, in geometric terms, the theory of constrained Lagrangians can be related with the theory of dissipative systems, the two main differences being: (i) the phase space of a dissipative system is the whole of TQ and not a submanifold M, (ii) the velocity dependence of the semibasic form for a constrained system is given by the function λ.

(3) The 1-form $\widetilde{\alpha}$ represents the constraint and, therefore, it does not depend on the particular Lagrangian considered. The dependence in L is given by the value of the function λ (later on denoted by Λ). This is the reason why we have considered first the two equations (3*a*, *b*) and not (4) directly. In (4) the 1-form $\mu = \lambda\widetilde{\alpha}$ depends on the particular form of the function L.

The function λ, the Lagrange multiplier in the non-geometric formalism, is uniquely determined by imposing the condition

$$\Gamma_\lambda(\widehat{\alpha} + \widetilde{h}) = 0\,.$$

Remark that $\widehat{\alpha} + \widetilde{h}$ is a constant for Γ, not for the Euler–Lagrange X_L-evolution. We obtain for λ the value $\lambda = \Lambda$,

$$\Lambda = \frac{X_L(\widehat{\alpha} + \widetilde{h})}{Z(\widehat{\alpha})}$$

that in coordinates reads,

$$\Lambda(q, v) = -\frac{1}{\alpha^2}\left(\frac{\partial\alpha_r}{\partial q^j}\, v^r v^j + \frac{\partial h}{\partial q^j}\, v^j + \alpha_j f^j\right)$$

where we have used the notation $\alpha^2 = W^{rs}\,\alpha_r\alpha_s$.

Suppose that L is a Lagrangian of mechanical type [19, 21] $L = \frac{1}{2}g(v, v) - V(q)$, in coordinates

$$L(q, v) = \tfrac{1}{2}\, g_{ij}(q) v^i v^j - V(q)$$

where g is a Riemannian metric on Q. Then the term $Z(\widehat{\alpha}) = \omega_L(S(X_\alpha), X_\alpha)$ turns out to be the norm of the 1-form $\widetilde{\alpha}$ with respect to this metric and, because of this, it is a non-vanishing function. This fact is related with the above remark (3): $\widetilde{\alpha}$ does not depend on L but its norm $Z(\widehat{\alpha})$ is L-dependent.

This property can be generalized to the case of L being a more general function (e.g. a non-polynomical regular Lagrangian). In this case W_{ij} can be velocity-dependent and then it does not define a metric on Q but a Euclidean or pseudo-Euclidean bilinear form on the subspace of $T^*_{(q,v)}(TQ)$ obtained the restriction of

$\bigwedge^1_{\rm sb}(TQ)$ to every point $(q,v) \in TQ$. The term $Z(\widehat{\alpha})$ turns out to be the norm (in every point) of the 1-form $\widetilde{\alpha}$ with respect to this bilinear form and, consequently, the general case corresponds to $Z(\widehat{\alpha}) \neq 0$. Nevertheless there is a very special case in which this function can vanish: it corresponds to (i) the bilinear form defined by W_{ij} being pseudo-Euclidean and (ii) $\widetilde{\alpha}$ being an isotropic (null or lightlike) co-vector for it. Thus, only if L and α are related by these two (unusual) relations (i) and (ii), can the Lagrange function Λ not be obtained.

Finally, the constrained dynamics is given by the restriction to M_α of

$$\Gamma = v^i \frac{\partial}{\partial q^i} + \left\{ f^i(q,v) - \frac{1}{\alpha^2}\left(\frac{\partial \alpha_r}{\partial q^j} v^r v^j + \frac{\partial h}{\partial q^j} v^j + \alpha_j f^j \right) W^{ik} \alpha_k \right\} \frac{\partial}{\partial v^i}$$

and its integral curves are the solutions of

$$\frac{\mathrm{d}}{\mathrm{d}t} q^i = v^i \qquad \frac{\mathrm{d}}{\mathrm{d}t} v^i = f^i(q,v) - \frac{1}{\alpha^2}\left(\frac{\partial \alpha_r}{\partial q^j} v^r v^j + \frac{\partial h}{\partial q^j} v^j + \alpha_j f^j \right) W^{ik} \alpha_k \,.$$

Remark that Γ represents a vector field defined in the whole of TQ that has the linear function $\phi = \widehat{\alpha} + \widetilde{h} \in C^\infty(TQ)$ as a constant of the motion. That is, the submanifolds $\{\phi^{-1}(r),\ r \in \mathbb{R}\}$ are invariant submanifolds for Γ and the constrained phase space M_α is the particular leaf corresponding to $r = 0$, i.e. $M_\alpha = \phi^{-1}(0)$. Consequently the flow in $M_\alpha = \phi^{-1}(0)$ of the restriction $\Gamma_\alpha \in \mathcal{X}(M_\alpha)$ of Γ to M_α defines the dynamics of the constrained 'physical' system; nevertheless it is possible to consider also the flow of Γ in other leaves but then the evolution (that could be called 'unphysical') will not satisfy the constraint. Moreover Γ can be modified in such a way that its restriction to M_α is preserved. This can be done by considering the above condition of tangency in the more general form $\Gamma_\lambda(\widehat{\alpha} + \widetilde{h}) = p$ where p is a function vanishing on M_α, e.g. $p = (\widehat{\alpha} + \widetilde{h})^n$. The new vector field $\overline{\Gamma}$ turns out to be

$$\overline{\Gamma} = \Gamma + \left(\frac{p}{\alpha^2} W^{ik} \alpha_k \right) \frac{\partial}{\partial v^i} \,.$$

Obviously $\overline{\Gamma} \neq \Gamma$ but $\overline{\Gamma}|M_\alpha = \Gamma|M_\alpha$.

We conclude this section generalizing all these results to the more general case of several constraints.

Let L be subject to a system $\{\phi_a = \widehat{\alpha}_a + \widetilde{h}_a;\ a = 1,\dots,A\}$ of $A < n$ lineal constraints; in coordinates $\phi_a(q,v) = \alpha_{aj}(q) v^j + h_a(q)$. We first recall two basic properties (see for example [20, 32]):

(i) The property of independence for the system $\{\alpha_a;\ a = 1,\dots,A\}$ is given by the condition $\Omega = \alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_A \neq 0$.

(ii) If $\Omega \wedge \mathrm{d}\alpha_a = 0$, $a = 1,\dots,A$, then the intersection of the kernels of the α_a is an involutive distribution and the system of constraints is holonomic. In this case there exist functions f_a and g^b_a, $a,b = 1,\dots,A$, such that $\alpha_a = g^b_a\,\mathrm{d}f_b$.

The dynamics is restricted to the submanifold $M \subset TQ$, $M = \bigcap M_a$, $M_a = {\phi_a}^{-1}(0)$. Every constraint determines its associated vertical vector field $Z_a \in \mathcal{X}^v(TQ)$ and Γ_λ takes the form $\Gamma_\lambda = X_L - \lambda^a Z_a$, $\lambda^a \in C^\infty(TQ)$.

Let $\{X_a;\ a = 1,\dots,A\}$ be the Hamiltonian vector fields of the functions $\{\widehat{\alpha}_a;\ a = 1,\dots,A\}$. Then Z_a and X_a are related by $Z_a = -S(X_a)$ and the functions $Z_a(\widehat{\alpha}_b)$ become

$$Z_a(\widehat{\alpha}_b) = \omega_L(X_b, Z_a) = \omega_L(S(X_a), X_b)\,.$$

Proposition 2.2. Let L be a Lagrangian of mechanical type. Then the matrix M defined by $M_{ab} = Z_a(\widehat{\alpha}_b)$, $a, b = 1, \ldots, A$ is non-singular.

Proof. First notice that $Z_a(\widehat{\alpha}_b) = W^{ij}\alpha_{ai}\alpha_{bj} = W_{ij}z_a{}^i z_b{}^j$, $z_a{}^i = W^{ij}\alpha_{aj}$, and that the A-independent 1-forms $\{\alpha_a;\ a = 1, \ldots, A\}$ determine in every point $q \in Q$ an A-dimensional subspace V_q^* of T_q^*Q.

If the Lagrangian L is of mechanical type then it defines a Riemannian metric g on Q whose components are the matrix elements W_{ij} of the Hessian W. Hence, by duality, the A vector fields $\{Y_a \in \mathcal{X}(Q);\ a = 1, \ldots, A\}$ defined by $Y_a = z_a{}^i \partial/\partial q^i$, determine an A-dimensional subspace $V_q \subset T_qQ$; moreover, using this metric, $Z_a(\widehat{\alpha}_b)$ can be interpreted as the function $g(Y_a, Y_b)$. Let $\{e_{qa}\}$ be an orthonormal basis for V_q, i.e. $g_q(e_{qa}, e_{qb}) = \delta_{ab}$. Then M can be writen as $M = BgB^t$ where B is the matrix expressing the vectors $\{Y_a\}$ in terms of the orthonormal basis. Thus $\det M = \det(BgB^t) = (\det B)^2 \neq 0$.

This result can be extended to regular Lagrangians defining a pseudo-Riemannian metric, the only additional condition being that the subspace V_q must admit a pseudo-orthonormal basis, i.e. $g_q(e_{qa}, e_{qb}) = \pm\delta_{ab}$.

In the set $\{\Gamma_\lambda,\ \lambda^a \in C^\infty(TQ)\}$ there is a unique vector field Γ that is tangent to M. It is determined by solving for the λ^a the A equations $\Gamma_\lambda(\widehat{\alpha}_b + \widetilde{h}_b) = 0$, $b = 1, \ldots, A$. We obtain

$$X_L(\widehat{\alpha}_b + \widetilde{h}_b) = \lambda^a\, Z_a(\widehat{\alpha}_b)$$

whose solutions are

$$\Lambda^a = M^{ab}\, X_L(\widehat{\alpha}_b + \widetilde{h}_b)$$

where $[M^{ab}]$ is the inverse matrix of $[M_{ab}]$.

Finally, the constrained dynamics is given by the restriction to M of

$$\Gamma = v^i \frac{\partial}{\partial q^i} + \left\{ f^i(q, v) - M^{ab}\, X_L(\widehat{\alpha}_b + \widetilde{h}_b) W^{ij}\, \alpha_{aj} \right\} \frac{\partial}{\partial v^i}$$

and its integral curves are the solutions of

$$\frac{\mathrm{d}}{\mathrm{d}t}\, q^i = v^i \qquad \frac{\mathrm{d}}{\mathrm{d}t}\, v^i = f^i(q, v) - M^{ab}\, X_L(\widehat{\alpha}_b + \widetilde{h}_b) W^{ij}\, \alpha_{aj}\,.$$

3. Constants of the motion and other properties

In this section we study the properties characterizing the fields Γ and Γ_α.

Proposition 3.1. The vector field Z satisfies $\mathcal{L}_Z \theta_L = -\widetilde{\alpha}$.

Proof. Since $\omega_L = -\mathrm{d}\,\theta_L$ we have

$$\mathrm{i}(Z)\mathrm{d}\,\theta_L = -\widetilde{\alpha}$$

and therefore

$$\mathcal{L}_Z\theta_L - \mathrm{d}[\mathrm{i}(Z)\theta_L] = -\widetilde{\alpha}\,.$$

As Z is vertical and θ_L semibasic we have $\mathrm{i}(Z)\theta_L = 0$. Hence

$$\mathcal{L}_Z\theta_L = -\widetilde{\alpha}\,.$$

Using this property and recalling that the Euler–Lagrange vector field X_L satisfies the equation $\mathcal{L}_{X_L}\theta_L = \mathrm{d}L$, we obtain that the equation for Γ can also be written in geometric terms as

$$\mathcal{L}_\Gamma\theta_L = \mathrm{d}L + \Lambda\,\widetilde{\alpha}\,.$$

Proposition 3.2. Let $g \in C^\infty$ be a constant of the motion for L, and Y its Hamiltonian vector field. If $Y \in \ker(\widetilde{\alpha})$ then g is also an integral for Γ.

Proof. The vector field $Y \in \mathcal{X}(TQ)$ is defined by $\mathrm{i}(Y)\omega_L = \mathrm{d}g$, therefore

$$\Gamma(g) = \mathrm{i}(\Gamma)\{\mathrm{i}(Y)\omega_L\} = -\mathrm{i}(Y)\{\mathrm{d}E_L - \Lambda\,\widetilde{\alpha}\}\,.$$

If $X_L(g) = 0$ then $Y(E_L) = 0$; thus

$$\Gamma(g) = \Lambda(\mathrm{i}(Y)\widetilde{\alpha})$$

and consequently if $Y \in \ker(\widetilde{\alpha})$ then $\Gamma(g) = 0$.

In the particular case of $g = E_L$ we obtain

$$\Gamma_\alpha(E_L) = \Lambda(\mathrm{i}(X_L)\widetilde{\alpha}) = \Lambda\,\widehat{\alpha}\,.$$

If the constraint ϕ is homogeneous (that is, $\phi = \widehat{\alpha}$), then the Energy function E_L is a constant of the motion for the constrained system $\Gamma_\alpha = \Gamma|M_\alpha$. If the constraint is non-homogeneous, (that is, $\phi = \widehat{\alpha} + \widetilde{h}$) then E_L is not an integral for the constrained motion and the value of $\Lambda\,\widehat{\alpha}$ in M_α (that is, $\Lambda\,\widehat{\alpha} = -\Lambda\,\widetilde{h}$) represents the dissipative term.

We recall [19, 29, 32] that given a symmetric connection on a manifold Q a vector field X on the tangent bundle TQ is called a geodesic spray if the integral curves of X consist precisely of the natural lifts of the geodesics of the connection. It can be proved that if a SODE field $X \in \mathcal{X}^2(TQ)$ satisfies the homogeneity condition $X = [\Delta, X]$ then it is the geodesic spray of a symmetric connection.

Suppose that L is the kinetic Lagrangian $L = \frac{1}{2}g_{ij}v^iv^j$ associated to a Riemannian metric g defined on Q. In this case X_L is the geodesic spray

$$X_L = v^i\frac{\partial}{\partial q^i} + f^i(q,v)\frac{\partial}{\partial v^i} \qquad f^i(q,v) = -\Gamma^i_{rs}\,v^rv^s$$

where Γ^i_{rs} are the coefficients of the Levi-Civita connection determined by g.

Proposition 3.3. If the Euler–Lagrange field X_L is a geodesic spray and the constraint ϕ is linear and homogeneous then Γ is a spray.

Proof. If X_L is a geodesic spray and the constraint function ϕ takes the form $\phi = \widehat{\alpha}$ then the vector field Γ is

$$\Gamma = v^i \frac{\partial}{\partial q^i} + F^i(q,v) \frac{\partial}{\partial v^i} \qquad F^i(q,v) = -M^i_{rs}\, v^r v^s$$

where the basic functions M^i_{rs} are

$$M^i_{rs} = \Gamma^i_{rs} + \frac{1}{\alpha^2}\, g^{ik}\, \alpha_k \left(\frac{\partial \alpha_r}{\partial q^s} - \alpha_j \Gamma^j_{rs} \right).$$

Thus Γ, besides being SODE, is also a spray—$\Gamma = [\Delta, \Gamma]$.

If the 1-form α is exact then the functions M^i_{rs} are symmetric and they can be considered as the coefficients of a symmetric connection, Γ being the geodesic spray of that connection. If α is not exact then the functions M^i_{rs} are not symmetric and, therefore, they must be symmetrized (that is, $\overline{M}^i_{rs} = \frac{1}{2}(M^i_{rs} + M^i_{sr})$) for obtaining a symmetric connection. The particular case of this connection being of Levi–Civita class corresponds to Γ being a Lagrangian vector field, that is, if $M^i_{rs} = M^i_{sr}$ are not only symmetric but also coefficients of the Levi-Civita connection determined by a Riemannian metric G on Q, then $\Gamma = X_{\mathbb{L}}$ where $X_{\mathbb{L}}$ is the Euler–Lagrange field of the kinetic Lagrangian $\mathbb{L} = \frac{1}{2} G_{ij} v^i v^j$.

The constrained system Γ_α will be said to admit a Lagrangian description with $\mathbb{L} \in C^\infty(TQ)$ as an admissible Lagrangian if the Euler–Lagrange field $X_{\mathbb{L}}$ satisfies $X_{\mathbb{L}}|M_\alpha = \Gamma_\alpha$.

Proposition 3.4. The function $\mathbb{L} = L + \Lambda h$ is an admissible Lagrangian for Γ_α, $\alpha = \mathrm{d}h$, if and only if so is Λ.

Proof. If the function $\mathbb{L}$ takes the form $\mathbb{L} = L + \Lambda h$ then

$$\begin{aligned} \mathcal{L}_\Gamma \theta_{\mathbb{L}} &= \mathcal{L}_\Gamma \theta_L + \mathcal{L}_\Gamma \theta_{\Lambda h} = \mathrm{d}L + \Lambda\, \mathrm{d}h + \Gamma(h)\theta_\Lambda + h \mathcal{L}_\Gamma \theta_\Lambda \\ &= \mathrm{d}\mathbb{L} + h(\mathcal{L}_\Gamma \theta_\Lambda - \mathrm{d}\Lambda) + \Gamma(h)\theta_\Lambda . \end{aligned}$$

On the submanifold M_α the linear function $\Gamma(h)$ vanishes, i.e. $\Gamma(h) = \widehat{\mathrm{d}h} = 0$. Hence L satisfies the equation $\mathcal{L}_\Gamma \theta_{\mathbb{L}} = \mathrm{d}\mathbb{L}$ on M_α if and only if $\mathcal{L}_\Gamma \theta_\Lambda = \mathrm{d}\Lambda$ on M_α.

4. Holonomic constraints

The non-geometric formalism of the theoretical mechanics usually considers first the holonomic systems and only then, and using this study as a previous result, studies the non-holonomic systems. It seems that this fact is motivated because:

(i) The holonomic constraints are represented by functions defined not in the phase space but in the configuration space; because of this, they appear easier to handle.

(ii) It is also assumed that the work to be done consists mainly of generalizing (or extending) the velocity-independent results to the more general case of dependence on the velocities.

The geometric formalism we are presenting here considers the problem in the inverse way. Having solved the non-holonomic case, now we are going to study the case of holonomic constraints. The two basic properties are:

(i) In geometric terms an holonomic constraint has a double value in the sense that it reduces the phase space not in one but in two dimensions.

(ii) *The dynamical vector field is only defined in the tangent bundle* $T(Q_h)$.

Suppose that the Lagrangian system (TQ, ω_L, X_L) is subject to the constraint $\widetilde{h}$. Then the dynamics must be represented by a vector field in $\mathcal{X}(T(Q_h))$. Notice that a point in TQ belongs to $T(Q_h)$ only if its coordinates (q^i, v^i) satisfiy the two equations $h(q^i) = 0$ and $\widehat{\mathrm{d}h}(q^i, v^i) = 0$. Because of this the holonomic case can be considered as a two-step process: first we consider the linear and homogeneous non-holonomic constraint corresponding to $\alpha = \mathrm{d}h$, and then we reduce the configuration space from Q to Q_h.

In geometric terms the basic equations are

$$\mathrm{i}(X_L)\omega_L = \mathrm{d}E_L \tag{5a}$$

$$\mathrm{i}(Z)\omega_L = \mathrm{d}\widetilde{h}\,. \tag{5b}$$

Consequently, if the constraint is holonomic, i.e. $\phi = \widetilde{h}$, then the vector field Z is Hamiltonian and the constraint function $\widetilde{h}$ is the ω_L-Hamiltonian of the perturbation.

The λ-family $\{\Gamma_\lambda,\ \lambda \in C^\infty(TQ)\}$ now satisfies the equation

$$\mathrm{i}(\Gamma_\lambda)\omega_L = \mathrm{d}E_L - \lambda\,\mathrm{d}\widetilde{h}$$

and the dynamics is given by the vector field $\Gamma_h \in \mathcal{X}(T(Q_h))$ obtained as the restriction to $T(Q_h)$ of the vector field Γ in Γ_λ, that is, $\Gamma_h = \Gamma|T(Q_h)$, $\Gamma \in \Gamma_\lambda$.

The coordinate expression for Γ_λ is

$$\Gamma_\lambda = v^i\frac{\partial}{\partial q^i} + F^i(q,v,\lambda)\frac{\partial}{\partial v^i} \qquad F^i(q,v,\lambda) = f^i(q,v) + \lambda\,W^{ij}\frac{\partial h}{\partial q^j}$$

and its integral curves are solutions of

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}q^i = f^i\left(q, \frac{\mathrm{d}q}{\mathrm{d}t}\right) + \lambda\,W^{ij}\frac{\partial h}{\partial q^j}$$

which agree with the equations (1), (2*a*) of the 'traditional' approach.

The function λ, the Lagrange multiplier in the non-geometric formalism, is uniquely determined by imposing the condition

$$\Gamma_\lambda(\widehat{\mathrm{d}h}) = 0\,.$$

Hence

$$X_L(\widehat{\mathrm{d}h}) = \lambda\,Z(\widehat{\mathrm{d}h})\,.$$

In coordinates,

$$X_L(\widehat{\mathrm{d}h}) = \frac{\partial^2 h}{\partial q^k\partial q^j}v^k v^j + \frac{\partial h}{\partial q^k}f^k \qquad Z(\widehat{\mathrm{d}h}) = -W^{kj}\frac{\partial h}{\partial q^k}\frac{\partial h}{\partial q^j} = -h'^2$$

therefore

$$\Lambda(q,v) = -\frac{1}{h'^2}\left(\frac{\partial^2 h}{\partial q^k \partial q^j}\, v^k v^j + \frac{\partial h}{\partial q^k} f^k\right).$$

Hence, the dynamical vector field $\Gamma_h \in \mathcal{X}(T(Q_h))$ takes the form

$$\Gamma_h = v^i \frac{\partial}{\partial q^i} + \left\{ f^i(q,v) - \frac{1}{h'^2}\left(\frac{\partial^2 h}{\partial q^k\, \partial q^j}\, v^k v^j \;+ \frac{\partial h}{\partial q^j} f^j\right) W^{ik} \frac{\partial h}{\partial q^k} \right\} \frac{\partial}{\partial v^i}$$

and the equations of its integral curves are

$$\frac{\mathrm{d}}{\mathrm{d}t}\, q^i = v^i \qquad \frac{\mathrm{d}}{\mathrm{d}t}\, v^i = f^i(q,v) - \frac{1}{h'^2}\left(\frac{\partial^2 h}{\partial q^r\, \partial q^j}\, v^r v^j + \frac{\partial h}{\partial q^j} f^j\right) W^{ik} \frac{\partial h}{\partial q^k}\,.$$

Finally, the classical result concerning the Lagrangian character of the holonomic systems is stated in this geometric formalism as follows.

Proposition 4.1. The function $\mathbb{L} = L + \Lambda h$ is an admissible Lagrangian for Γ_h.

Proof. Before proceeding to the proof notice that here Λ is a function on TQ and does not represent, as in other approaches, a new degree of freedom.

This property follows from proposition 3.4 where we proved that if $\mathbb{L} = L + \Lambda h$ then

$$\mathcal{L}_\Gamma \theta_{\mathbb{L}} = \mathrm{d}\mathbb{L} + h(\mathcal{L}_\Gamma \theta_\Lambda - \mathrm{d}\Lambda) + \Gamma(h)\theta_\Lambda\,.$$

The tangent bundle $T(Q_h)$ is characterized by the vanishing of both h and $\Gamma(h) = \widehat{\mathrm{d}h}$; therefore L satisfies $\mathcal{L}_\Gamma \theta_{\mathbb{L}} = \mathrm{d}\mathbb{L}$ on the restricted phase space.

5. Examples

In the following examples coordinate indices will be written as subscripts.

Example 1. As an example of an α-type constraint we will consider the motion of a rolling disk constrained to remain vertical [6,17]. The configuration space Q is $Q = \mathbb{R}^2 \times S^1 \times S^1$ and the (free) Lagrangian $L \in C^\infty(TQ)$ for this system is

$$L = \tfrac{1}{2}(m v_1^2 + m v_2^2 + I_1 v_3^2 + I_0 v_4^2)$$

where m, I_0 and I_1 are constants.

The constraints for this system are of non-holonomic character and can be expressed by the two equations

$$\widehat{\alpha}_1 = v_1 - (R\cos q_3)v_4 = 0 \qquad \widehat{\alpha}_2 = v_2 - (R\sin q_3)v_4 = 0\,.$$

In this case the 2-form ω_L and the 1-forms α_1, α_2, are

$$\omega_L = m\,\mathrm{d}q_1 \wedge \mathrm{d}v_1 + m\,\mathrm{d}q_2 \wedge \mathrm{d}v_2 + I_1\,\mathrm{d}q_3 \wedge \mathrm{d}v_3 + I_0\,\mathrm{d}q_4 \wedge \mathrm{d}v_4$$

$$\alpha_1 = \mathrm{d}q_1 - (R\cos q_3)\mathrm{d}q_4 \qquad \alpha_2 = \mathrm{d}q_2 - (R\sin q_3)\mathrm{d}q_4$$

and the vector fields X_L and Z_1, Z_2, turn out to be

$$X_L = v_k \frac{\partial}{\partial q_k} \qquad Z_1 = -\frac{1}{m}\frac{\partial}{\partial v_1} + \frac{R}{I_0}\cos q_3 \frac{\partial}{\partial v_4} \qquad Z_2 = -\frac{1}{m}\frac{\partial}{\partial v_2} + \frac{R}{I_0}\sin q_3 \frac{\partial}{\partial v_4}$$

and thus, the λ-dependent vector field Γ_λ takes the form

$$\begin{aligned}\Gamma_\lambda &= X_L - \lambda_1 Z_1 - \lambda_2 Z_2 \\ &= v_k \frac{\partial}{\partial q_k} + \frac{\lambda_1}{m}\frac{\partial}{\partial v_1} - \frac{\lambda_1 R}{I_0}\cos q_3 \frac{\partial}{\partial v_4} + \frac{\lambda_2}{m}\frac{\partial}{\partial v_2} - \frac{\lambda_2 R}{I_0}\sin q_3 \frac{\partial}{\partial v_4}\end{aligned}$$

and its integral curves satisfy

$$\ddot{q}_1 = \frac{\lambda_1}{m} \qquad \ddot{q}_3 = 0 \qquad \ddot{q}_2 = \frac{\lambda_2}{m} \qquad \ddot{q}_4 = -\frac{\lambda_1 R}{I_0}\cos q_3 - \frac{\lambda_2 R}{I_0}\sin q_3 \,.$$

The functions $\Lambda_i = \Lambda_i(q, v)$, $i = 1, 2$, are to be obtained by imposing the conditions

$$\begin{aligned}\Gamma_\lambda(\widehat{\alpha}_1) &= X_L(\widehat{\alpha}_1) - \lambda_1 Z_1(\widehat{\alpha}_1) - \lambda_2 Z_2(\widehat{\alpha}_1) = 0 \\ \Gamma_\lambda(\widehat{\alpha}_2) &= X_L(\widehat{\alpha}_2) - \lambda_1 Z_1(\widehat{\alpha}_2) - \lambda_2 Z_2(\widehat{\alpha}_2) = 0\end{aligned}$$

and thus we obtain

$$\Lambda_1 = -(mR\sin q_3)v_3 v_4 \qquad \Lambda_2 = (mR\cos q_3)v_3 v_4 \,.$$

Finally the dynamics is given by

$$\Gamma = v_k \frac{\partial}{\partial q_k} - (R\sin q_3)v_3 v_4 \frac{\partial}{\partial v_1} + (R\cos q_3)v_3 v_4 \frac{\partial}{\partial v_2} \,.$$

The projection to Q of the integral curves of Γ are

$$q_1(t) = R\frac{v_{40}}{v_{30}}\sin(v_{30}t + q_{30}) + a_1 t + b_1 \qquad q_2(t) = -R\frac{v_{40}}{v_{30}}\cos(v_{30}t + q_{30}) + a_2 t + b_2$$

$$q_3(t) = v_{30}t + q_{30} \qquad q_4(t) = v_{40}t + q_{40}$$

where a_i, b_i, $i = 1, 2$, are constants and $(q_0, v_0) = \{q_{i0}, v_{i0}; i = 1, \ldots, 4\}$ are the initial data. The two functions $\widehat{\alpha}_i$, $i = 1, 2$, are constants of the motion with invariant values $\widehat{\alpha}_i = a_i$, $i = 1, 2$. Consequently if the initial data satisfy the two constraints, i.e. $(q_0, v_0) \in M_\alpha = \widehat{\alpha}_1^{-1}(0) \cap \widehat{\alpha}_2^{-1}(0)$, then $a_1 = a_2 = 0$ and the corresponding functions $q_i = q_i(t)$, $i = 1, \ldots, 4$, represent the evolution of the constrained 'physical' system.

Example 2. As an example of an h-type constraint we will consider the rigid rotor, that is to say, a free point-particle constrained to move on the sphere of radius r. The configuration space Q is $Q = \mathbb{R}^3$, the free Lagrangian $L \in C^\infty(TQ)$ is

$$L = \tfrac{1}{2}m\sum_{k=1}^{3} v_k^2$$

and the constraint for this system can be expressed by the equation $h(q)=0$, where

$$h(q)=\sum_{k=1}^{3} q_k^2 - r^2 .$$

In this case the 'constrained' phase space is the tangent bundle

$$T(Q_h)=\left\{(q,v)\in T(\mathbb{R}^3) \;\middle|\; \sum_{k=1}^{3} q_k^2 - r^2 = 0 \qquad q_k v_k = 0\right\}$$

and the 2-form ω_L and the 1-form α are

$$\omega_L = m\,\mathrm{d}q_k \wedge \mathrm{d}v_k \qquad \alpha = \mathrm{d}h = 2\,q_k\,\mathrm{d}q_k .$$

The vector fields X_L and Z turn out to be

$$X_L = v_i\frac{\partial}{\partial q_i} \qquad Z = -\frac{2q_i}{m}\frac{\partial}{\partial v_i}$$

and hence, the λ-family of vector fields Γ_λ takes the form

$$\Gamma_\lambda = v_i\frac{\partial}{\partial q_i} + \lambda\left(\frac{2q_i}{m}\right)\frac{\partial}{\partial v_i} .$$

The particular value Λ of $\lambda(q,v)$ is obtained by imposing the condition

$$\Gamma_\lambda(\widehat{\mathrm{d}h}) = \Gamma_\lambda(2q_k v_k) = 0$$

and thus we obtain

$$\Lambda = -\frac{m}{2\sum q_j^2}\sum_{k=1}^{3} v_k^2 .$$

Finally the constrained dynamics is given by

$$\Gamma_h = v_i\frac{\partial}{\partial q_i} - \left(\frac{\sum v_k^2}{r^2}\right) q_i \frac{\partial}{\partial v_i} .$$

Remark that as $\Gamma_h \in \mathcal{X}(T(Q_h))$ we have made use of the equation $h(q)=0$. Moreover the second-order equations

$$\ddot{q}_i = -\left(\frac{\sum v_k^2}{r^2}\right) q_i \qquad i=1,2,3$$

are defined not on the whole $Q=\mathbb{R}^3$ but on Q_h.

Example 3. As an example of an $\alpha=\mathrm{d}h$ constraint we will consider the Lagrangian L of the above previous example but now with the constraint $q_k v_k = 0$. The constraint dynamical system is $\Gamma_\alpha = \Gamma|M_\alpha$ with

$$\Gamma = v_i\frac{\partial}{\partial q_i} - \left(\frac{v^2}{q^2}\right) q_i\frac{\partial}{\partial v_i} \qquad v^2=\sum_{k=1}^{3} v_k^2 \qquad q^2 = \sum_{k=1}^{3} q_k^2 .$$

The Euler–Lagrange vector field X_Λ defined as the solution of the equation $\mathrm{i}(X_\Lambda)\omega_\Lambda = \mathrm{d}E_\Lambda$, $E_\Lambda = \Lambda$, is

$$X_\Lambda = v_i\frac{\partial}{\partial q_i} + f_i^\Lambda\frac{\partial}{\partial v_i} \qquad f_i^\Lambda = -\left(\frac{v^2}{q^2}\right)q_i + 2\left(\frac{q_k v_k}{q^2}\right)v_i .$$

Thus $X_\Lambda \neq \Gamma$ but $X_\Lambda|M_\alpha = \Gamma_\alpha$. Hence Λ is a Lagrangian for Γ_α.

6. Hamiltonian formalism

Let $L \in C^\infty(TQ)$ be a regular Lagrangian determining in TQ the Hamiltonian dynamical system $\{TQ, \omega_L, E_L, X_L\}$ and let us denote by $D_L \in \mathrm{Diff}(TQ, T^*Q)$ its associate Legendre transformation; then it is known that D_L determines in T^*Q a new Hamiltonian system $\{T^*Q, \omega_0, H, X_H\}$ by $\omega_0 = D_{L*}(\omega_L)$, $H = D_{L*}(E_L)$, $X_H = D_{L*}(X_L)$, where $D_{L*} = (D_L{}^*)^{-1}$ and $\omega_0 = -\mathrm{d}\theta_0$ is the canonical symplectic form.

Since D_L is fibre-preserving it satisfies $D_{L*}(\mathcal{X}^v(TQ)) = \mathcal{X}^v(T^*Q)$, and therefore the field $Z_* = D_{L*}(Z)$ is vertical in T^*Q. Moreover D_L is also base-invariant, therefore besides preserving the semibasic forms (i.e. $D_L{}^*(\bigwedge^1_{\mathrm{sb}}(T^*Q)) = \bigwedge^1_{\mathrm{sb}}(TQ)$) it reduces to the identity on basic forms. Hence $D_{L*}(\tilde{\alpha}) = \tilde{\alpha}$ (with a slight abuse of notation we also use $\tilde{\alpha}$ for $\pi^*\alpha$, $\pi : T^*Q \to Q$). Thus the image of $\{TQ, \omega_L, \tilde{\alpha}, Z\}$ by D_L is the (non-Hamiltonian) system $\{T^*Q, \omega_0, \tilde{\alpha}, Z_*\}$ where the field Z_* satisfies

$$\mathrm{i}(Z_*)\omega_0 = \tilde{\alpha} .$$

In local coordinates it takes the form

$$Z_* = \alpha_i(q)\frac{\partial}{\partial p_i} .$$

The dynamics is now restricted to the submanifold $N_\alpha \subset T^*Q$ defined by $N_\alpha = D_{L*}(M_\alpha)$ and is given by the restriction to N_α of

$$\Gamma_* = D_{L*}(\Gamma) = X_H - \Lambda_* Z_*$$

where the function $\Lambda_*(q,p) = D_{L*}(\Lambda(q,v))$ takes the form

$$\Lambda_*(q,p) = -\frac{1}{\alpha^2}\left\{\left(\frac{\partial H}{\partial p_i}\right)\frac{\partial}{\partial q^i}(W^{kj}\alpha_k)p_j - \left(\frac{\partial H}{\partial q^i}\right)(W^{ki}\alpha_k) + \left(\frac{\partial H}{\partial p_i}\right)\left(\frac{\partial h}{\partial p_i}\right)\right\}$$

(we have supposed that L is of mechanical type, i.e. $v^i = W^{ij}p_j$, but if L contains magnetic terms linear in the velocities then this relation must be modified by including a non-homogeneous term). Finally, the integral curves of Γ_* in N_α are the solutions of

$$\frac{\mathrm{d}}{\mathrm{d}t}q^r = \frac{\partial H}{\partial p_r}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}p_r = -\frac{\partial H}{\partial q^r} - \frac{1}{\alpha^2}\left\{\left(\frac{\partial H}{\partial p_i}\right)\frac{\partial}{\partial q^i}(W^{kj}\alpha_k)p_j - \left(\frac{\partial H}{\partial q^i}\right)(W^{ki}\alpha_k) + \left(\frac{\partial H}{\partial p_i}\right)\left(\frac{\partial h}{\partial p_i}\right)\right\}\alpha_r .$$

Acknowledgment

Partial financial support by CICYT (Comisión Interministerial Científica y Técnica) is acknowledged.

References

[1] Hanson A, Regge T and Teitelboim C 1976 *Constrained Hamiltonian Systems* (Roma: Accademia dei Lincei)
[2] Sundermeyer K. 1982 *Constrained Dynamics (Lecture Notes in Physics 169)* (Berlin: Springer)
[3] Cariñena J F 1990 Theory of singular Lagrangians *Forschr. Phys.* **38** 641
[4] Saletan E J and Cromer A H 1970 *Am. J. Phys.* **38** 892
[5] Saletan E J and Cromer A H 1971 *Theoretical Mechanics* (New York: Wiley)
[6] Galvao C A and Negri L J 1983 *J. Phys. A: Math. Gen.* **16** 4183
[7] Huston R L and Passerello C E 1976 *Int. J. Non-Linear Mech.* **11** 331
[8] Rumiantsev V V 1978 *Appl. Math. Mech. PMM* **42** 387
[9] Rumiantsev V V 1979 *Appl. Math. Mech. PMM* **43** 583
[10] Rumiantsev V V 1984 *Appl. Math. Mech. PMM* **48** 380
[11] Rumiantsev V V 1983 *Proc. IUTAM Symp. (Modern Developments in Analytical Mechanics)* ed S Benenti, M Francaviglia and A Lichnerowicz
[12] Brendelev V N 1982 *Appl. Math. Mech. PMM* **45** 351
[13] Naziev E Kh 1972 *Appl. Math. Mech. PMM* **36** 1108
[14] Weber R W 1983 *Proc. IUTAM Symp. (Modern Developments in Analytical Mechanics)* ed S Benenti, M Francaviglia and A Lichnerowicz
[15] Weber R W 1985 *Arch. Ration. Mech. Anal.* **91** 309
[16] Cardin F and Zanzotto G 1989 *J. Math. Phys.* **30** 1473
[17] Giachetta G 1992 *J. Math. Phys.* **33** 1652
[18] Koiller J 1992 *Arch. Ration. Mech. Anal.* **118** 113
[19] Abraham R and Marsden J 1978 *Foundations of Mechanics* (Reading, MA: Benjamin)
[20] Marmo G, Saletan E J, Simoni A and Vitale B 1985 *Dynamical Systems: A Differential Geometric Approach* (Chichester, UK: Wiley)
[21] Morandi G, Ferrario C, Lo Vecchio G, Marmo G and Rubano C 1990 The inverse problem in the calculus of variations and the geometry of the tangent bundle *Phys. Rep.* **188** 147
[22] de León M and Rodrigues P R 1989 *Methods of Differential Geometry in Analytical Mechanics* (North-Holland: Amsterdam)
[23] Lichnerowicz A 1975 *C. R. Acad. Sci., Paris* A **280** 523
[24] Gotay M J, Nester J M and Hinds G 1978 *J. Math. Phys.* **19** 2388
[25] Gotay M J and Nester J M 1979 *Ann. Inst. H. Poincaré* A **30** 129
[26] Gotay M J and Nester J M 1980 *Ann. Inst. H. Poincaré* A **32** 1
[27] Cantrijn F, Cariñena J F, Crampin M and Ibort L A 1986 *J. Geom. Phys.* **3** 353
[28] Crampin M 1981 *J. Phys. A: Math. Gen.* **14** 2567
[29] Crampin M 1983 *J. Phys. A: Math. Gen.* **16** 3755
[30] Cantrijn F 1982 *J. Math. Phys.* **23** 1589
[31] Cantrijn F 1984 *J. Math. Phys.* **25** 271
[32] Crampin M and Pirani F A 1988 *Applicable Differential Geometry* (Cambridge: Cambridge University Press)