Available online at www.sciencedirect.com
ScienceDirect
Journal of Geometry and Physics 57 (2006) 209–250
JOURNAL OF GEOMETRY AND PHYSICS
www.elsevier.com/locate/jgp

Dirac structures in Lagrangian mechanics Part II: Variational structures

Hiroaki Yoshimura[a], Jerrold E. Marsden[b,*]

[a] *Department of Mechanical Engineering, Waseda University, Ohkubo, Shinjuku, Tokyo 169-8555, Japan*
[b] *Control and Dynamical Systems, California Institute of Technology, Pasadena, CA 91125, USA*

Received 6 October 2005; received in revised form 17 February 2006; accepted 26 February 2006
Available online 17 April 2006

Abstract

Part I of this paper introduced the notion of implicit Lagrangian systems and their geometric structure was explored in the context of Dirac structures. In this part, we develop the variational structure of implicit Lagrangian systems. Specifically, we show that the implicit Euler–Lagrange equations can be formulated using an extended variational principle of Hamilton called the Hamilton–Pontryagin principle. This variational formulation incorporates, in a natural way, the generalized Legendre transformation, which enables one to treat degenerate Lagrangian systems. The definition of this generalized Legendre transformation makes use of natural maps between iterated tangent and cotangent spaces. Then, we develop an extension of the classical Lagrange–d'Alembert principle called the Lagrange–d'Alembert–Pontryagin principle for implicit Lagrangian systems with constraints and external forces. A particularly interesting case is that of nonholonomic mechanical systems that can have both constraints and external forces. In addition, we define a constrained Dirac structure on the constraint momentum space, namely the image of the Legendre transformation (which, in the degenerate case, need not equal the whole cotangent bundle). We construct an implicit constrained Lagrangian system associated with this constrained Dirac structure by making use of an Ehresmann connection. Two examples, namely a vertical rolling disk on a plane and an L–C circuit are given to illustrate the results.

Keywords: The Hamilton–Pontryagin principle; Degenerate Lagrangians; Constrained Dirac structures; L–C circuits

1. Introduction

In Part II of this paper, we continue to develop the framework of implicit Lagrangian systems and their geometry in the context of Dirac structures, which was begun in Part I. This part focuses on the variational structure of implicit Lagrangian systems. An algebraic theory of Dirac structures associated with formal variational calculus is contained in the work of [17,18] in the Hamiltonian framework of integrable evolution equations. However, it has not been clear how Dirac structures are interrelated with implicit mechanical systems, whether Lagrangian or Hamiltonian, in the context of variational principles. In other words, there has been a gap between Dirac structures and variational

* Corresponding author.
E-mail addresses: yoshimura@waseda.jp (H. Yoshimura), marsden@cds.caltech.edu (J.E. Marsden).

0393-0440/$ - see front matter
doi:10.1016/j.geomphys.2006.02.012

principles in mechanics. In conjunction with Dirac's theory of constraints, we remark that Dirac started off with Hamilton's principle with the aim of investigating degenerate Lagrangian systems, although in the end, he did not focus on the Lagrangian formulation but rather developed the notion of a constrained Poisson structure, or the "Dirac bracket" (see, for instance, [16,24]).

Needless to say, in mechanics, Hamilton's principle is employed to formulate the Euler–Lagrange equations on the Lagrangian side, while Hamilton's phase space principle may be used to derive Hamilton's equations on the Hamiltonian side. As is well known, both formalisms, in the case where a given Lagrangian is hyperregular, are equivalent via the Legendre transformation (see [1,25]). For the case in which a Lagrangian is degenerate, we need a specific treatment to deal with constraints due to the degeneracy, as in Dirac's theory of constraints. For such degenerate Lagrangian systems, Weinstein [40] noted that a Dirac structure on a Lie algebroid may be induced from a Poisson structure on the dual bundle of the Lie algebroid. He also considered the case in which the Lie algebroid is given by a tangent bundle. L–C circuits have not been treated so far in the context of degenerate Lagrangian systems (see, for instance, [12]). A variational principle for L–C circuits was developed using Pontryagin's maximum principle by Moreau and Aeyels [26] and a formulation of implicit Lagrangian systems was developed by Moreau and van der Schaft [27]. Both of these differ from our notion of implicit Lagrangian systems in the sense that they utilize a Dirac structure on a subbundle of the tangent bundle of a configuration manifold, consistent with Weinstein's idea [40]. Furthermore, mechanical systems with nonholonomic constraints have been widely studied (see, for instance, [36,2]), specifically, from the viewpoint of symmetry and reduction by Bloch, Krishnaprasad, Marsden and Murray [8], where the Lagrange–d'Alembert principle played an essential role in formulating the equations of motion and, in addition, the system viewed in terms of a constrained Lagrangian was formulated using an Ehresmann connection. On the Hamiltonian side, constrained Hamiltonian systems were developed from the viewpoint of Poisson structures by van der Schaft and Maschke [34] and then, a notion of implicit Hamiltonian systems was developed in the context of Dirac structures by van der Schaft and Maschke [35] and van der Schaft [33] (see also [3]). Nonconservative systems with external forces that appeared in servomechanisms were also illustrated in the context of the constrained Hamiltonian systems by Marle [22]. The equivalence of the Lagrangian and Hamiltonian formalisms for nonholonomic mechanical systems was demonstrated by Koon and Marsden [19,20] together with their intrinsic expressions. As for the details on nonholonomic mechanics and control, refer also to Bloch [5] and Cendra, Marsden and Ratiu [11].

As in Part I, an implicit Lagrangian system, whose Lagrangian may be degenerate, can be defined by using a Dirac structure on T^*Q that is induced from a constraint distribution $\Delta_Q \subset TQ$. To carry this out, Part I utilized natural symplectomorphisms between the spaces TT^*Q, T^*TQ, and T^*T^*Q. We also developed the Dirac differential of a Lagrangian which, amongst other things, incorporated the Legendre transformation into the context of induced Dirac structures. This procedure is consistent with the idea of a generalized Legendre transformation, which was originally developed by Tulczyjew [31] (see also, for instance [9]). In the present Part II, we establish some basic links between variational principles on the one hand and Dirac structures and implicit Lagrangian systems, including the generalized Legendre transform, on the other.

Another important issue that is relevant for the present paper is Pontryagin's maximum principle in optimal control developed by Pontryagin, Boltyanskiĭ, Gamkrelidze and Mishchenko [29]. It goes without saying that Pontryagin's maximum principle is the machinery that gives necessary conditions for solutions of optimal control problems; we remark that a coordinate-free version of the maximum principle was given by Sussmann [30].

1.1. Variational principles

One of the main goals in this Part II is to provide a link between variational structures, induced Dirac structures, and implicit Lagrangian systems. To do this, we shall develop an extended variational principle called the Hamilton–Pontryagin principle.

The variational principle of Hamilton for classical holonomic mechanical systems is given by the stationary condition of the action functional for a Lagrangian L such that

$$\delta \int_{t_1}^{t_2} L(q, v)\,\mathrm{d}t = 0,$$

which is subject to the second-order condition $\dot{q} = v$ and with the endpoints of $q(t)$ fixed. Regarding the second-order condition $\dot{q} = v$ as a kinematic constraint, we introduce the momentum variable p as a Lagrange multiplier for this constraint, and then we rewrite Hamilton's principle as

$$\delta \int_{t_1}^{t_2} \{L(q, v) + p \cdot (\dot{q} - v)\} \, \mathrm{d}t = 0. \tag{1.1}$$

We will call this form of the basic variational principle of holonomic mechanics the *Hamilton–Pontryagin principle* because of its close relation with the classical Pontryagin principle.[1] This principle is also closely related to what in elasticity theory is called the Hu–Washizu principle (see, for instance, [23,37]), that is so important in the discontinuous Galerkin method. In the form (1.1), this principle seems to be due to Livens [21], and it also appears in Pars [28], Section 26.2.

The Hamilton–Pontryagin principle in Eq. (1.1) may also be stated in the following equivalent form, by introducing an extended, or generalized, energy $E(q, v, p) = p \cdot v - L(q, v)$, so that

$$\delta \int_{t_1}^{t_2} \{p \cdot \dot{q} - E(q, v, p)\} \, \mathrm{d}t = 0.$$

Note that the Hamilton–Pontryagin variational principle gives us the second-order condition, the Legendre transformation, and the Euler–Lagrange equations:

$$\dot{q} = v, \quad p = \frac{\partial L}{\partial v}, \quad \dot{p} = \frac{\partial L}{\partial q}. \tag{1.2}$$

While this is a special case of the system that we wish to develop, it does provide a point of view for generalizing this procedure to the case of implicit Lagrangian systems. Note that this Eq. (1.2) includes the case of degenerate Lagrangians, and that these equations may be implicit in the sense that the Euler–Lagrange equations could be "hidden". Of course the degenerate case of this problem and its transformation to the Hamiltonian side were the subject of the important work of Dirac [15]. For these reasons, we refer to equations Eq. (1.2) as the *implicit Euler–Lagrange equations.*

Consistent with the above discussion of the Hamilton–Pontryagin principle, for the general case in which a constraint distribution $\Delta_Q \subset TQ$ is given, we will show how to formulate an implicit Lagrangian system (L, Δ_Q, X) in terms of an extended Lagrange–d'Alembert principle that will be called the Lagrange–d'Alembert–Pontryagin principle.

Furthermore, we shall develop a constrained Dirac structure on the constraint momentum space $P = \mathbb{F}L(\Delta_Q) \subset T^*Q$, as well as the implicit constrained Lagrangian system associated with a constrained Lagrangian $L_c = L|\Delta_Q$ in the variational context.

1.2. Outline of the paper

Part II of the paper is constructed as follows. In Section 2, we give a brief review of the generalized Legendre transformation, which makes use of iterated tangent and cotangent bundles. In Section 3, the Hamilton–Pontryagin principle and the associated implicit Euler–Lagrange equations are developed in detail along with the geometry of iterated tangent and cotangent spaces. Extending this analysis to the case of nonholonomic constrained systems, it is shown that more general implicit Lagrangian systems can be intrinsically formulated in terms of the Lagrange–d'Alembert–Pontryagin principle. Furthermore, we also elucidate implicit Hamiltonian systems in the variational context for the case of regular Lagrangians. In Section 4, we demonstrate that nonholonomic systems with external force fields can be incorporated into the framework of implicit Lagrangian systems by employing the Lagrange–d'Alembert–Pontryagin principle with external forces. In Section 5, we develop the constrained Dirac structure D_P that is induced on the constraint momentum space P by utilizing an Ehresmann connection; and we construct an implicit constrained Lagrangian system (L_c, Δ_Q, X_P) associated with the constrained Lagrangian L_c and the constrained vector field X_P on P in the variational context. In Section 6, the two examples, namely a vertical rolling disk on a plane and an L–C circuit, are illustrated in the context of implicit constrained Lagrangian systems. In Section 7, concluding remarks and future directions are given.

[1] We thank Matt West, Tom Hughes, and David Gao for helpful comments on this principle.

1.3. Summary of the main results

- The Hamilton–Pontryagin principle for holonomic but possibly degenerate Lagrangians, and its relation to implicit Lagrangian systems and Dirac structures on the cotangent bundle are developed.
- Using the geometry of iterated tangent and cotangent bundles and the Pontryagin bundle, we develop the intrinsic form of implicit Lagrangian systems in the variational context. Intrinsic implicit Hamiltonian systems are also developed.
- Nonholonomic systems with external force fields that appear in controlled mechanical systems such as robots are developed in the context of implicit Lagrangian systems and we show that the equations of motion of such systems can be formulated in terms of the Lagrange–d'Alembert–Pontryagin principle.
- We construct a constrained Dirac structure on the constraint momentum space and show how an implicit constrained Lagrangian system can be formulated using an Ehresmann connection in the variational context.
- Two examples are presented. A vertical rolling disk illustrates implicit nonholonomic constrained Lagrangian systems. An L–C circuit is presented as a typical example of a constrained system with a degenerate Lagrangian.

2. The generalized Legendre transform

As illustrated in Part I, the spaces TT^*Q, T^*TQ, T^*T^*Q are interrelated with each other by two symplectomorphisms $\kappa_Q : TT^*Q \to T^*TQ$ and $\Omega^\flat : TT^*Q \to T^*T^*Q$, which play essential roles in the construction of the generalized Legendre transformation originally developed by Tulczyjew [31]. The link between tangent Dirac structures and the spaces TT^*Q and T^*TQ was investigated by Courant [14]. In this section, we shall review the generalized Legendre transformation before going into the variational framework of implicit Lagrangian systems. As to the details and required mathematical ingredients for the generalized Legendre transform, refer, for instance, to Cendra, Holm, Hoyle and Marsden [9], Weinstein [38,39], Abraham and Marsden [1], and Tulczyjew and Urbański [32].

2.1. Symplectic structure on TT^*Q

Let Q be a manifold, TQ the tangent bundle and T^*Q the cotangent bundle of Q. Let q, $(q, \delta q)$ and (q, p) be local coordinates for Q, TQ and T^*Q respectively. Let $(q, \delta q, \delta p, p)$, $(q, p, \delta q, \delta p)$ and $(q, p, -\delta p, \delta q)$ be local coordinates for T^*TQ, TT^*Q and T^*T^*Q respectively. Let $\pi_Q : T^*Q \to Q;\ (q, p) \mapsto q$ be the cotangent projection and $T\pi_Q : TT^*Q \to TQ;\ (q, p, \delta q, \delta p) \mapsto (q, \delta q)$, be the tangent map of π_Q. Furthermore, $\pi_{TQ} : T^*TQ \to TQ;\ (q, \delta q, \delta p, p) \mapsto (q, \delta q)$ and $\tau_{T^*Q} : TT^*Q \to T^*Q;\ (q, p, \delta q, \delta p) \mapsto (q, p)$. In Part I, we showed that there is a natural diffeomorphism

$$\kappa_Q : TT^*Q \to T^*TQ; \quad (q, p, \delta q, \delta p) \mapsto (q, \delta q, \delta p, p)$$

that is determined by how it intertwines the two sets of maps:

$$\pi_{TQ} \circ \kappa_Q = T\pi_Q \quad \text{and} \quad \pi^1 \circ \kappa_Q = \tau_{T^*Q}.$$

In the above, we recall that the projection $\pi^1 : T^*TQ \to T^*Q;\ (q, \delta q, \delta p, p) \mapsto (q, p)$ is defined, for $\alpha_{v_q} \in T^*_{v_q}TQ$ and $u_q \in T_qQ$, such that $\langle \pi^1(\alpha_{v_q}), u_q \rangle = \langle \alpha_{v_q}, \mathrm{ver}(u_q, v_q) \rangle$, where $\mathrm{ver}(u_q, v_q)$ is the vertical lift of u_q along v_q.

On the other hand, the map $\Omega^\flat : TT^*Q \to T^*T^*Q;\ (q, p, \delta q, \delta p) \mapsto (q, p, -\delta p, \delta q)$ is the natural diffeomorphism associated with the canonical symplectic structure Ω on T^*Q. Recall that the manifold TT^*Q is the symplectic manifold with a particular symplectic form Ω_{TT^*Q} that can be defined by the two distinct but intrinsic one-forms:

$$\lambda = (\kappa_Q)^* \Theta_{T^*TQ} = \delta p\, \mathrm{d}q + p\, \mathrm{d}\delta q,$$
$$\chi = (\Omega^\flat)^* \Theta_{T^*T^*Q} = -\delta p\, \mathrm{d}q + \delta q\, \mathrm{d}p,$$

where Θ_{T^*TQ} is the canonical one-form on T^*TQ and $\Theta_{T^*T^*Q}$ is the canonical one-form on T^*T^*Q. Recall also that the symplectic form Ω_{TT^*Q} is defined by

$$\Omega_{TT^*Q} = -\mathbf{d}\lambda = \mathbf{d}\chi = \mathrm{d}q \wedge \mathrm{d}\delta p + \mathrm{d}\delta q \wedge \mathrm{d}p.$$

Let us see how the Lagrangian and Hamiltonian may be interrelated with each other throughout the symplectic structure Ω_{TT^*Q}.

2.2. Lagrangian constraints

Let L be a Lagrangian on $\Delta_Q \subset TQ$. The symplectic manifold $(TT^*Q, \Omega_{TT^*Q} = -\mathbf{d}\lambda)$ is defined by the quadruple $(TT^*Q, TQ, T\pi_Q, \lambda)$ and the set

$$\begin{aligned} N = \{x \in TT^*Q \mid\, & T\pi_Q(x) \in \Delta_Q, \lambda_x(w) = \big\langle \mathbf{d}L(T\pi_Q(x)), T_x T\pi_Q(w)\big\rangle \\ & \text{for all } w \in T_x(TT^*Q) \text{ such that } T_x T\pi_Q(w) \in T_{T\pi_Q(x)}\Delta_Q\} \end{aligned} \tag{2.1}$$

is a Lagrangian submanifold of $(TT^*Q, \Omega_{TT^*Q} = -\mathbf{d}\lambda)$ with $\frac{1}{2}\dim TT^*Q$, where the submanifold

$$\Delta_Q = T\pi_Q(N) \subset TQ$$

is the constraint distribution on Q called a *Lagrangian constraint*. Hence, the Lagrangian L is a generating function of N, since $N \subset TT^*Q$ is the graph of $(\kappa_Q)^{-1}(\mathbf{d}L)$.

2.3. Hamiltonian constraints

Let H be a Hamiltonian on $P \subset T^*Q$ and the symplectic manifold $(TT^*Q, \Omega_{TT^*Q} = \mathbf{d}\chi)$ be defined by the quadruple $(TT^*Q, T^*Q, \tau_{T^*Q}, \chi)$. The set

$$\begin{aligned} N = \{x \in TT^*Q \mid\, & \tau_{T^*Q}(x) \in P, \chi_x(w) = \big\langle \mathbf{d}H(\tau_{T^*Q}(x)), T_x\tau_{T^*Q}(w)\big\rangle \\ & \text{for all } w \in T_x(TT^*Q) \text{ such that } T_x\tau_{T^*Q}(w) \in T_{\tau_{T^*Q}(x)}P\} \end{aligned} \tag{2.2}$$

is a Lagrangian submanifold of $(TT^*Q, \Omega_{TT^*Q} = \mathbf{d}\chi)$ with $\frac{1}{2}\dim TT^*Q$, where the submanifold

$$P = \tau_{T^*Q}(N) \subset T^*Q$$

is the *constraint momentum space* called a *Hamiltonian constraint*. Similarly, the Hamiltonian H is a generating function of N, because $N \subset TT^*Q$ is the graph of $(\Omega^\flat)^{-1}(\mathbf{d}H)$.

2.4. Symplectomorphism and the momentum function

Consider the identity map, which we can regard as a symplectomorphism $\varphi : (P_1 = TT^*Q, \Omega_{P_1} = -\mathbf{d}\lambda) \to (P_2 = TT^*Q, \Omega_{P_2} = \mathbf{d}\chi)$, and it follows

$$\varphi^*\,\Omega_{P_2} = \Omega_{P_1},$$

since $P_1 = P_2$, and, as we have seen in Part I, $\Omega_{P_2} = \Omega_{P_1}$. The graph of the symplectomorphism φ is a submanifold of $P_1 \times P_2$, which is denoted by

$$\Gamma(\varphi) \subset P_1 \times P_2.$$

Let $i_\varphi : \Gamma(\varphi) \to P_1 \times P_2$ be the inclusion and let $\pi_i : P_1 \times P_2 \to P_i$ be the canonical projection. Define

$$\begin{aligned} \omega &= \pi_1^*\,\Omega_{P_1} - \pi_2^*\,\Omega_{P_2} \\ &= \pi_1^*(-\mathbf{d}\lambda) - \pi_2^*\mathbf{d}\chi. \end{aligned}$$

Since φ is symplectic, it follows that

$$\begin{aligned} i_\varphi^*\omega &= (\pi_1|\Gamma(\varphi))^*(\Omega_{P_1} - \varphi^*\Omega_{P_2}) \\ &= (\pi_1|\Gamma(\varphi))^*(-\mathbf{d}\lambda - \varphi^*\mathbf{d}\chi) \\ &= 0. \end{aligned}$$

In the above, $\pi_1 \circ i_\varphi$ is the projection restricted to $\Gamma(\varphi)$ and $\pi_2 \circ i_\varphi = \varphi \circ \pi_1$ on $\Gamma(\varphi)$. So, we can write $\omega = -\mathbf{d}\theta$ where $\theta = \lambda \oplus \chi = \pi_1^*\lambda + \pi_2^*\chi$. Clearly, $\Gamma(\varphi)$ is a maximally isotropic submanifold with half of the dimension of $P_1 \times P_2 = TT^*Q \times TT^*Q$.

Letting $\Psi : TT^*Q \to TT^*Q \times TT^*Q$ be the diagonal map, we have the one-form $\Psi^*\theta$ on TT^*Q, which is represented, by using local coordinates $(q, \delta q)$, (q, p) and $(q, p, \delta q, \delta p)$ for TQ, T^*Q and TT^*Q, as

$$\begin{aligned}\Psi^*\theta &= \Psi^*(\lambda \oplus \chi) = \lambda + \varphi^*\chi \\ &= (\delta p\, \mathrm{d}q + p\, \mathrm{d}\delta q) + (-\delta p\, \mathrm{d}q + \delta q\, \mathrm{d}p) \\ &= p\, \mathrm{d}\delta q + \delta q\, \mathrm{d}p \\ &= \mathbf{d}(p \cdot \delta q) \\ &= \mathbf{d}(G \circ \rho_{TT^*Q}),\end{aligned}$$

where we recall from Part I that the map $\rho_{TT^*Q} : TT^*Q \to TQ \oplus T^*Q$ is given in coordinates by $(q, p, \delta q, \delta p) \mapsto (q, \delta q, p)$. We shall also need the function G defined on the Pontryagin bundle $TQ \oplus T^*Q$ that simply pairs an element of T_qQ with that of T_q^*Q; it is given in local coordinates by

$$G(q, \delta q, p) = p \cdot \delta q,$$

which we shall call the *momentum function*.

2.5. The generalized Legendre transform

There are two different ways of realizing the submanifold N in TT^*Q, as shown in Eqs. (2.1) and (2.2), as graphs of one-forms on TQ and T^*Q and the passage between them implies the Legendre transformation. This procedure is called the *generalized Legendre transform*, which enables us to treat the case in which a given Lagrangian is degenerate.

The generalized Legendre transformation is the procedure to obtain the submanifold $\mathcal{K}$ of the Pontryagin bundle $TQ \oplus T^*Q$ from a submanifold N of TT^*Q associated with $(TT^*Q, TQ, T\pi_Q, \lambda)$ and with a Lagrangian L (possibly degenerate) on $\Delta_Q \subset TQ$ as in Eq. (2.1). This can be understood by the passage of the identity symplectomorphism

$$\varphi : (TT^*Q, \Omega_{TT^*Q} = -\mathbf{d}\lambda) \to (TT^*Q, \Omega_{TT^*Q} = \mathbf{d}\chi).$$

Let $T\pi_Q \times \tau_{T^*Q} : TT^*Q \times TT^*Q \to TQ \times T^*Q$ and define the map

$$(T\pi_Q \times \tau_{T^*Q}) \circ \Psi : TT^*Q \to TQ \times T^*Q.$$

We can define a submanifold $\mathcal{K}$ by the image of N obtained by the map $(T\pi_Q \times \tau_{T^*Q}) \circ \Psi$ such that

$$\mathcal{K} = (T\pi_Q \times \tau_{T^*Q}) \circ \Psi(N) \subset TQ \times T^*Q,$$

which is to be the graph of the Legendre transform $\mathbb{F}L : TQ \to T^*Q$ with respect to a constraint distribution $\Delta_Q = T\pi_Q(N) \subset TQ$. Define the *generalized energy* E on $TQ \oplus T^*Q$, using local coordinates (q, v), (q, p) and (q, v, p) for TQ, T^*Q and $TQ \oplus T^*Q$ and $\mathrm{pr}_{TQ} : TQ \oplus T^*Q \to TQ$, by

$$\begin{aligned}E(q, v, p) &= G(q, v, p) - L(\mathrm{pr}_{TQ}(q, v, p)) \\ &= p \cdot v - L(q, v).\end{aligned}$$

In fact, the submanifold $\mathcal{K}$ may be given by

$$\mathcal{K} = \{(q, v, p) \in TQ \oplus T^*Q \mid (q, v) \in \Delta_Q \subset TQ \text{ is a stationary point of } E(q, v, p) \text{ for each } (q, p) \in T^*Q\},$$

which is eventually represented by

$$\mathcal{K} = \left\{ (q, v, p) \in TQ \oplus T^*Q \,\middle|\, (q, v) \in \Delta_Q,\ p = \frac{\partial L}{\partial v} \right\}.$$

3. The variational framework

In this section, we illustrate the variational framework of implicit Lagrangian systems. First, we show that *implicit Euler–Lagrange equations* can be formulated by using an extended variational principle of Hamilton called the *Hamilton–Pontryagin principle*, which we develop by inspiration from Pontryagin's maximum principle (see [29]).

This variational principle naturally includes the generalized Legendre transformation. Second, we investigate an implicit Lagrangian system (L, Δ_Q, X) by developing an extended Lagrange–d'Alembert principle called the *Lagrange–d'Alembert–Pontryagin principle* together with its intrinsic expression. Then, we show the variational link between an implicit Lagrangian system and the induced Dirac structure on T^*Q. Third, in the case where a given Lagrangian L is hyperregular, a Hamiltonian H is well defined by the usual Legendre transformation. Then, we also show how an implicit Hamiltonian system (H, Δ_Q, X) that is defined by an induced Dirac structure on T^*Q can be naturally associated with an extension of Hamilton's phase space principle that we call the *Hamilton–d'Alembert principle in phase space*.

3.1. Variational principle of Hamilton

Before going into an extended variational principle of Hamilton, we shall review the variational principle of Hamilton.

Let L be a Lagrangian on TQ and $q(t),\ t_1 \leq t \leq t_2$, be a curve in the manifold Q. Define the path space from q_1 to q_2 by

$$C(q_1, q_2, [t_1, t_2]) = \{q : [t_1, t_2] \to Q \mid q(t_1) = q_1,\ q(t_2) = q_2\}$$

and the map called the action functional $\mathfrak{S} : C(q_1, q_2, [t_1, t_2]) \to \mathbb{R}$ by

$$\mathfrak{S}(q(t)) = \int_{t_1}^{t_2} L(q(t),\ \dot{q}(t))\ \mathrm{d}t.$$

The variation of the action functional $\mathfrak{S}(q(t))$ at $q(t)$ in direction of $\delta q(t)$ is

$$\begin{aligned}
\mathbf{d}\mathfrak{S}(q(t)) \cdot \delta q(t) &= \delta \int_{t_1}^{t_2} L(q(t), \dot{q}(t))\ \mathrm{d}t \\
&= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q}\,\delta q + \frac{\partial L}{\partial \dot{q}}\,\delta \dot{q} \right) \mathrm{d}t \\
&= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} \right) \cdot \delta q + \left. \frac{\partial L}{\partial \dot{q}} \cdot \delta q \right|_{t_1}^{t_2}.
\end{aligned}$$

In the above, $\dot{q}$ denotes $\mathrm{d}q/\mathrm{d}t$ and we employ $\delta\dot{q} = \mathrm{d}(\delta q)/\mathrm{d}t$. When $q(t)$ is a critical point of the action functional $\mathfrak{S}$, that is, $\mathbf{d}\mathfrak{S}(q(t)) \cdot \delta q(t) = 0$ for all $\delta q(t) \in T_{q(t)}C(q_1, q_2, [t_1, t_2])$, the curve $q(t)$ satisfies, keeping the endpoints fixed, the Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}.$$

3.2. The Hamilton–Pontryagin principle

We shall illustrate that implicit Euler–Lagrange equations can be formulated using an extended variational principle of Hamilton, which we call the *Hamilton–Pontryagin principle*, that incorporates the second-order condition $v = \dot{q}$ into the action functional for a Lagrangian as a kinematical constraint. This variational principle naturally includes the generalized Legendre transform.

Proposition 3.1. *Let q, (q, v) and (q, p) be local coordinates respectively for Q, TQ and T^*Q. Let (q, v, p) be local coordinates for the Pontryagin bundle $TQ \oplus T^*Q$. Let $L : TQ \to \mathbb{R}$ be a Lagrangian (possibly degenerate). Consider the action functional defined by*

$$\int_{t_1}^{t_2} \{L(q(t), v(t)) + p(t)\,(\dot{q}(t) - v(t))\}\ \mathrm{d}t = \int_{t_1}^{t_2} \{p(t) \cdot \dot{q}(t) - E(q(t), v(t), p(t))\}\ \mathrm{d}t.$$

In the above, as previously illustrated, E is the generalized energy on $TQ \oplus T^*Q$ given by

$$\begin{aligned}
E(q, v, p) &= G(q, v, p) - L(q, v) \\
&= p \cdot v - L(q, v),
\end{aligned}$$

where $G(q, v, p) = p \cdot v$ is the momentum function. Keeping the endpoints of $q(t)$ fixed whereas the endpoints of $v(t)$ and $p(t)$ are allowed to be free, the stationary condition for the action functional implies

$$\dot{q} = v, \quad \dot{p} = \frac{\partial L}{\partial q}, \quad p = \frac{\partial L}{\partial v}, \tag{3.1}$$

which we shall call the implicit Euler–Lagrange equations.

Proof. The variation of the action functional is given by

$$\begin{aligned}
&\delta \int_{t_1}^{t_2} \{L(q(t), v(t)) + p(t)\,(\dot{q}(t) - v(t))\}\ \mathrm{d}t \\
&\quad = \delta \int_{t_1}^{t_2} \{p(t) \cdot \dot{q}(t) - E(q(t), v(t), p(t))\}\,\mathrm{d}t \\
&\quad = \int_{t_1}^{t_2} \{\delta p\,\dot{q} + p\,\delta\dot{q} - \delta E(q, v, p)\}\,\mathrm{d}t \\
&\quad = \int_{t_1}^{t_2} \left(-\dot{p}\,\delta q + \dot{q}\,\delta p - \frac{\partial E}{\partial q}\delta q - \frac{\partial E}{\partial v}\delta v - \frac{\partial E}{\partial p}\delta p\right)\mathrm{d}t + p\,\delta q\Big|_{t_1}^{t_2} \\
&\quad = \int_{t_1}^{t_2} \left\{\left(\dot{q} - \frac{\partial E}{\partial p}\right)\delta p + \left(-\dot{p} - \frac{\partial E}{\partial q}\right)\delta q - \frac{\partial E}{\partial v}\delta v\right\}\mathrm{d}t + p\,\delta q\Big|_{t_1}^{t_2} \\
&\quad = \int_{t_1}^{t_2} \left\{(\dot{q} - v)\,\delta p + \left(-\dot{p} + \frac{\partial L}{\partial q}\right)\delta q + \left(-p + \frac{\partial L}{\partial v}\right)\delta v\right\}\mathrm{d}t + p\,\delta q\Big|_{t_1}^{t_2},
\end{aligned} \tag{3.2}$$

where $(\delta q, \delta v, \delta p) \in T_{(q,v,p)}(TQ \oplus T^*Q)$. Keeping the endpoints of $q(t)$ fixed, that is, $q(t_1) = q_1$ and $q(t_2) = q_2$, the stationary condition for the action functional for all $(\delta q, \delta v, \delta p)$ provides Eq. (3.1). □

Notice that the Hamilton–Pontryagin principle naturally includes the Legendre transform and also that Eq. (3.1) is nothing but the local expression of the implicit Lagrangian system (L, Δ_Q, X) for the case in which $\Delta_Q = TQ$, as shown in Part I.

3.3. The intrinsic form of the implicit Euler–Lagrange equations

We shall develop the intrinsic form of the implicit Euler–Lagrange equations in the context of the Hamilton–Pontryagin principle.

Let $\rho_{TT^*Q} : TT^*Q \to TQ \oplus T^*Q$. Let $\mathrm{pr}_{TQ} : TQ \oplus T^*Q \to TQ$, $\mathrm{pr}_{T^*Q} : TQ \oplus T^*Q \to T^*Q$ and $\tau_{TQ\oplus T^*Q} : T(TQ \oplus T^*Q) \to TQ \oplus T^*Q$. Define the path space of curves $x(t) = (q(t), v(t), p(t))$, $t_1 \le t \le t_2$, in $TQ \oplus T^*Q$ by

$$\begin{aligned}
\widetilde{C}(q_1, q_2, [t_1, t_2]) = \{&(q, v, p) : [t_1, t_2] \to TQ \oplus T^*Q \mid \\
&\mathrm{pr}_Q(q(t_1), v(t_1), p(t_1)) = q_1,\ \mathrm{pr}_Q(q(t_2), v(t_2), p(t_2)) = q_2\},
\end{aligned}$$

where $\mathrm{pr}_Q : TQ \oplus T^*Q \to Q$.

The action functional on $\widetilde{C}(q_1, q_2, [t_1, t_2])$ of curves $x(t) = (q(t), v(t), p(t))$, $t_1 \le t \le t_2$, in $TQ \oplus T^*Q$ is represented by

$$\begin{aligned}
&\int_{t_1}^{t_2} \{L(q(t), v(t)) + p(t) \cdot (\dot{q}(t) - v(t))\}\ \mathrm{d}t \\
&\quad = \int_{t_1}^{t_2} \{p(t) \cdot \dot{q}(t) - E(q(t), v(t), p(t))\}\ \mathrm{d}t \\
&\quad = \int_{t_1}^{t_2} \left\{G(\rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q}(x(t), \dot{x}(t))) - E(\tau_{TQ\oplus T^*Q}(x(t), \dot{x}(t)))\right\}\ \mathrm{d}t,
\end{aligned} \tag{3.3}$$

where $\dot{x}(t)$ denotes the time derivative of $x(t)$, $T\mathrm{pr}_{T^*Q} : T(TQ \oplus T^*Q) \to TT^*Q$ is the tangent map of pr_{T^*Q} and then $\rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q} : T(TQ \oplus T^*Q) \to TQ \oplus T^*Q$.

Let us call Eq. (3.3) the *Hamilton–Pontryagin integral*.

3.4. Tangent bundle of the Pontryagin bundle

Let $\tau_Q : TQ \to Q$ be the tangent projection and $\pi_Q : T^*Q \to Q$ be the cotangent projection. The tangent space of $TQ \oplus T^*Q$ at a point (q, v, p), that is, $T_{(q,v,p)}(TQ \oplus T^*Q)$, is the subset of $T_{v_q}TQ \oplus T_{p_q}T^*Q$ consisting of vectors that project to the same point of TQ; that is,

$$T\tau_Q(q, v, \dot{q}, \dot{v}) = T\pi_Q(q, p, \dot{q}, \dot{p}),$$

where the tangent map $T\tau_Q : TTQ \to TQ$ is given in coordinates by $(q, v, \dot{q}, \dot{v}) \mapsto (q, \dot{q})$ and the tangent map $T\pi_Q : TT^*Q \to TQ$ is given in coordinates by $(q, p, \dot{q}, \dot{p}) \mapsto (q, \dot{q})$. Thus, tangent vectors of the Pontryagin bundle in coordinates have a base point (q, v, p) and a vector part $(\dot{q}, \dot{v}, \dot{p})$. That is, the $\dot{q}$ piece for the two tangent vectors of TQ and T^*Q agree.

There are two different maps from $T(TQ \oplus T^*Q)$ to $TQ \oplus T^*Q$; namely,

$$\begin{aligned}
&\tau_{TQ\oplus T^*Q} : T(TQ \oplus T^*Q) \to TQ \oplus T^*Q; \quad (q, v, p, \dot{q}, \dot{v}, \dot{p}) \mapsto (q, v, p),\\
&\rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q} : T(TQ \oplus T^*Q) \to TQ \oplus T^*Q; \quad (q, v, p, \dot{q}, \dot{v}, \dot{p}) \mapsto (q, \dot{q}, p),
\end{aligned}$$

where we recall that the map $\rho_{TT^*Q} : TT^*Q \to TQ \oplus T^*Q$ is given by $(q, p, \dot{q}, \dot{p}) \mapsto (q, \dot{q}, p)$. The map $T\mathrm{pr}_{T^*Q}$ is the tangent of the projection map $\mathrm{pr}_{T^*Q} : TQ \oplus T^*Q \to T^*Q$ and is given by $(q, v, p, \dot{q}, \dot{v}, \dot{p}) \mapsto (q, p, \dot{q}, \dot{p})$.

Proposition 3.2. *Let $\Omega^\flat : TT^*Q \to T^*T^*Q$ be the bundle map associated with the canonical symplectic structure Ω on T^*Q and Θ_{T^*Q} be the canonical one-form on T^*Q. Let $\chi = (\Omega^\flat)^* \, \Theta_{T^*T^*Q}$ be the induced one-form on TT^*Q, where $\Theta_{T^*T^*Q}$ is the canonical one-form on T^*T^*Q. Then, the variation of the Hamilton–Pontryagin integral in Eq. (3.3) is represented by*

$$\begin{aligned}
&\delta \int_{t_1}^{t_2} \{L(q(t), v(t)) + p(t)\cdot(\dot{q}(t) - v(t))\}\, \mathrm{d}t\\
&\quad = \delta \int_{t_1}^{t_2} \{p(t)\cdot \dot{q}(t) - E(q(t), v(t), p(t))\}\,\mathrm{d}t\\
&\quad = \delta \int_{t_1}^{t_2} \{G(\rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q}(x, \dot{x})) - E(\tau_{TQ\oplus T^*Q}(x, \dot{x}))\}\,\mathrm{d}t\\
&\quad = \int_{t_1}^{t_2} \Big\{\Theta_{T^*T^*Q}\left(\Omega^\flat \cdot T\mathrm{pr}_{T^*Q}(x, \dot{x})\right) \cdot T\Omega^\flat\left(T_{(x,\dot{x})}(T\mathrm{pr}_{T^*Q})(w)\right)\\
&\qquad - \mathbf{d}E\left(\tau_{TQ\oplus T^*Q}(x, \dot{x})\right)\cdot T_{(x,\dot{x})}\tau_{TQ\oplus T^*Q}(w)\Big\}\,\mathrm{d}t\\
&\qquad + \Theta_{T^*Q}\left(\mathrm{pr}_{T^*Q} \circ (\tau_{TQ\oplus T^*Q})(x, \dot{x})\right)\cdot\left(T\tau_{T^*Q} \circ T\left(T\mathrm{pr}_{T^*Q}\right)(w)\right)\Big|_{t_1}^{t_2}\\
&\quad = \int_{t_1}^{t_2} \{(T\mathrm{pr}_{T^*Q})^*\chi(x, \dot{x}) - (\tau_{TQ\oplus T^*Q})^*\mathbf{d}E(x, \dot{x})\}\cdot w\,\mathrm{d}t\\
&\qquad + \Theta_{T^*Q}\left(\mathrm{pr}_{T^*Q} \circ (\tau_{TQ\oplus T^*Q})(x, \dot{x})\right)\cdot\left(T\tau_{T^*Q} \circ T\left(T\mathrm{pr}_{T^*Q}\right)(w)\right)\Big|_{t_1}^{t_2}.
\end{aligned} \tag{3.4}$$

In the above, $(x, \dot{x}) \in T(TQ \oplus T^*Q)$, $w \in T_{(x,\dot{x})}T(TQ \oplus T^*Q)$, and

$$T(T\mathrm{pr}_{T^*Q}) : TT(TQ \oplus T^*Q) \to T(TT^*Q),$$

where $T_{(x,\dot{x})}(T\mathrm{pr}_{T^*Q})(w) \in T_{T\mathrm{pr}_{T^*Q}(x,\dot{x})}(TT^*Q)$.

Proof. Let us check, by using local coordinates, that Eq. (3.4) is the intrinsic representation of Eq. (3.2). Let (q, v), (q, p) and (q, v, p) be local coordinates for TQ, T^*Q and $TQ \oplus T^*Q$.

First, by using the two different maps $\tau_{TQ\oplus T^*Q} : T(TQ \oplus T^*Q) \to TQ \oplus T^*Q$ and $\rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q} : T(TQ \oplus T^*Q) \to TQ \oplus T^*Q$, for each $x = (q, v, p) \in TQ \oplus T^*Q$, one obtains

$$\begin{aligned}\rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q}(x, \dot{x}) &= \rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q}(q, v, p, \dot{q}, \dot{v}, \dot{p})\\ &= \rho_{TT^*Q} \circ (q, p, \dot{q}, \dot{p})\\ &= (q, \dot{q}, p) \in TQ \oplus T^*Q,\end{aligned}$$

while

$$\begin{aligned}\tau_{TQ\oplus T^*Q}(x, \dot{x}) &= \tau_{TQ\oplus T^*Q}(q, v, p, \dot{q}, \dot{v}, \dot{p})\\ &= (q, v, p) \in TQ \oplus T^*Q,\end{aligned}$$

where $(x, \dot{x}) = (q, v, p, \dot{q}, \dot{v}, \dot{p})$. Since the momentum function G on $TQ \oplus T^*Q$ is locally given by

$$\begin{aligned}G(\rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q}(x, \dot{x})) &= G(q, \dot{q}, p)\\ &= p \cdot \dot{q},\end{aligned}$$

it reads, using local coordinates,

$$\begin{aligned}\delta \int_{t_1}^{t_2} G(\rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q}(x, \dot{x}))\,\mathrm{d}t &= \delta \int_{t_1}^{t_2} G(q, \dot{q}, p)\,\mathrm{d}t\\ &= \delta \int_{t_1}^{t_2} (p \cdot \dot{q})\ \mathrm{d}t\\ &= \int_{t_1}^{t_2} (p\,\delta\dot{q} + \delta p\,\dot{q})\ \mathrm{d}t\\ &= \int_{t_1}^{t_2} (-\dot{p}\,\delta q + \delta p\,\dot{q})\ \mathrm{d}t + p\,\delta q\Big|_{t_1}^{t_2}.\end{aligned} \tag{3.5}$$

Second, let us check the terms concerning the canonical one-forms $\Theta_{T^*T^*Q}$ in Eq. (3.4) by using local coordinates. From the map $\Omega^\flat \circ T\mathrm{pr}_{T^*Q} : T(TQ \oplus T^*Q) \to T^*T^*Q$, we can easily see that

$$\begin{aligned}\Theta_{T^*T^*Q}\left(\Omega^\flat \cdot T\mathrm{pr}_{T^*Q}(x, \dot{x})\right) &= \Theta_{T^*T^*Q}(q, p, -\dot{p}, \dot{q})\\ &= -\dot{p}\,\mathrm{d}q + \dot{q}\,\mathrm{d}p,\end{aligned}$$

where $T\mathrm{pr}_{T^*Q}(x, \dot{x}) = (q, p, \dot{q}, \dot{p})$ and then $\Omega^\flat \circ T\mathrm{pr}_{T^*Q}(x, \dot{x}) = (q, p, -\dot{p}, \dot{q})$. Furthermore, we can write $w \in T_{(x,\dot{x})}T(TQ \oplus T^*Q)$ in coordinates as

$$w = (q, v, p, \dot{q}, \dot{v}, \dot{p},\ \delta q, \delta v, \delta p, \delta\dot{q}, \delta\dot{v}, \delta\dot{p}),$$

and it follows that

$$\begin{aligned}T_{(x,\dot{x})}\left(T\mathrm{pr}_{T^*Q}\right)(w) &= T_{(x,\dot{x})}\left(T\mathrm{pr}_{T^*Q}\right)(q, v, p, \dot{q}, \dot{v}, \dot{p},\ \delta q, \delta v, \delta p, \delta\dot{q}, \delta\dot{v}, \delta\dot{p})\\ &= (q, p, \dot{q}, \dot{p},\ \delta q, \delta p, \delta\dot{q}, \delta\dot{p})\end{aligned}$$

and then

$$\begin{aligned}\left(T\Omega^\flat \circ T_{(x,\dot{x})}(T\mathrm{pr}_{T^*Q})\right)(w) &= T\Omega^\flat(q, p, \dot{q}, \dot{p},\ \delta q, \delta p, \delta\dot{q}, \delta\dot{p})\\ &= (q, p, -\dot{p}, \dot{q},\ \delta q, \delta p, -\delta\dot{p}, \delta\dot{q}),\end{aligned}$$

where $T(T\mathrm{pr}_{T^*Q}) : TT(TQ \oplus T^*Q) \to T(TT^*Q)$ and $T\Omega^\flat : T(TT^*Q) \to T(T^*T^*Q)$.

Therefore, noting $\chi = (\Omega^\flat)^* \, \Theta_{T^*T^*Q}$, we have

$$\begin{aligned} \Theta_{T^*T^*Q}\left(\Omega^\flat \circ T\mathrm{pr}_{T^*Q}(x,\dot{x})\right) \cdot \left(T\Omega^\flat \circ T_{(x,\dot{x})}(T\mathrm{pr}_{T^*Q})(w)\right) &= (T\mathrm{pr}_{T^*Q})^* \chi(x,\dot{x}) \cdot w \\ &= -\dot{p}\,\delta q + \dot{q}\,\delta p. \end{aligned} \tag{3.6}$$

As to the term associated with the canonical one-form Θ_{T^*Q} in Eq. (3.4), by employing the map $\mathrm{pr}_{T^*Q} \circ \left(\tau_{TQ\oplus T^*Q}\right) : T(TQ \oplus T^*Q) \to T^*Q$, we can obtain

$$\begin{aligned} \mathrm{pr}_{T^*Q} \circ \left(\tau_{TQ\oplus T^*Q}\right)(x,\dot{x}) &= \mathrm{pr}_{T^*Q} \circ (\tau_{TQ\oplus T^*Q})(q,v,p,\dot{q},\dot{v},\dot{p}) \\ &= \mathrm{pr}_{T^*Q}(q,v,p) \\ &= (q,p), \end{aligned}$$

and then

$$\begin{aligned} \Theta_{T^*Q}\left(\mathrm{pr}_{T^*Q} \circ (\tau_{TQ\oplus T^*Q})(x,\dot{x})\right) &= \Theta_{T^*Q}\,(q,p) \\ &= p\,\mathrm{d}q. \end{aligned}$$

Recall that $\tau_{T^*Q} : TT^*Q \to T^*Q$ and $T\tau_{T^*Q} : T(TT^*Q) \to TT^*Q$, and it follows that

$$\Theta_{T^*Q}\left(\mathrm{pr}_{T^*Q} \circ (\tau_{TQ\oplus T^*Q})(x,\dot{x})\right) \cdot \left(T\tau_{T^*Q} \circ T\left(T\mathrm{pr}_{T^*Q}\right)(w)\right) = p\,\delta q, \tag{3.7}$$

where the map $T\tau_{T^*Q} \circ T\left(T\mathrm{pr}_{T^*Q}\right) : TT(TQ \oplus T^*Q) \to TT^*Q$ is locally indicated by

$$\begin{aligned} T\tau_{T^*Q} \circ T\left(T\mathrm{pr}_{T^*Q}\right)(w) &= T\tau_{T^*Q} \circ T\left(T\mathrm{pr}_{T^*Q}\right)(q,v,p,\dot{q},\dot{v},\dot{p},\ \delta q,\delta v,\delta p,\delta\dot{q},\delta\dot{v},\delta\dot{p}) \\ &= T\tau_{T^*Q}(q,p,\dot{q},\dot{p},\ \delta q,\delta p,\delta\dot{q},\delta\dot{p}) \\ &= (q,p,\delta q,\delta p). \end{aligned}$$

From Eqs. (3.5)–(3.7), we can easily check that the following relation holds:

$$\begin{aligned} &\delta \int_{t_1}^{t_2} G\ (\rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q}(x,\dot{x}))\,\mathrm{d}t \\ &\quad = \int_{t_1}^{t_2} \Theta_{T^*T^*Q}\left(\Omega^\flat \cdot T\mathrm{pr}_{T^*Q}(x,\dot{x})\right) \cdot \left(T\Omega^\flat \circ T_{(x,\dot{x})}(T\mathrm{pr}_{T^*Q})(w)\right)\,\mathrm{d}t \\ &\qquad + \left.\Theta_{T^*Q}\left(\mathrm{pr}_{T^*Q} \circ (\tau_{TQ\oplus T^*Q})(x,\dot{x})\right) \cdot \left(T\tau_{T^*Q} \circ T\left(T\mathrm{pr}_{T^*Q}\right)(w)\right)\right|_{t_1}^{t_2} \\ &\quad = \int_{t_1}^{t_2} (T\mathrm{pr}_{T^*Q})^* \chi(x,\dot{x}) \cdot w\,\mathrm{d}t \\ &\qquad + \left.\Theta_{T^*Q}\left(\mathrm{pr}_{T^*Q} \circ (\tau_{TQ\oplus T^*Q})(x,\dot{x})\right) \cdot \left(T\tau_{T^*Q} \circ T\left(T\mathrm{pr}_{T^*Q}\right)(w)\right)\right|_{t_1}^{t_2}. \end{aligned} \tag{3.8}$$

Third, let us check the terms relating to E on $TQ \oplus T^*Q$ in Eq. (3.4). Since the local representation of E is given by

$$\begin{aligned} E\left(\tau_{TQ\oplus T^*Q}(x,\dot{x})\right) &= E(q,v,p) \\ &= p \cdot v - L(q,v), \end{aligned}$$

one can directly compute the differential of E in local coordinates such that

$$\begin{aligned} \mathbf{d}E\left(\tau_{TQ\oplus T^*Q}(x,\dot{x})\right) &= \mathbf{d}E(q,v,p) \\ &= \frac{\partial E}{\partial q}\mathrm{d}q + \frac{\partial E}{\partial v}\mathrm{d}v + \frac{\partial E}{\partial p}\mathrm{d}p \\ &= -\frac{\partial L}{\partial q}\mathrm{d}q + \left(p - \frac{\partial L}{\partial v}\right)\mathrm{d}v + v\,\mathrm{d}p. \end{aligned}$$

Recall that the map $T\tau_{TQ\oplus T^*Q} : TT(TQ \oplus T^*Q) \to T(TQ \oplus T^*Q)$ is locally represented by

$$\begin{aligned} T_{(x,\dot{x})}\tau_{TQ\oplus T^*Q}(w) &= T_{(x,\dot{x})}\tau_{TQ\oplus T^*Q}(q, v, p, \dot{q}, \dot{v}, \dot{p}, \delta q, \delta v, \delta p, \delta\dot{q}, \delta\dot{v}, \delta\dot{p}) \\ &= (q, v, p, \delta q, \delta v, \delta p), \end{aligned}$$

and hence it follows that

$$\begin{aligned} \mathbf{d}E(\tau_{TQ\oplus T^*Q}(x, \dot{x})) \cdot T_{(x,\dot{x})}\tau_{TQ\oplus T^*Q}(w) &= (\tau_{TQ\oplus T^*Q})^*\mathbf{d}E(x, \dot{x}) \cdot w \\ &= \frac{\partial E}{\partial q}\delta q + \frac{\partial E}{\partial v}\delta v + \frac{\partial E}{\partial p}\delta p \\ &= -\frac{\partial L}{\partial q}\delta q + \left(p - \frac{\partial L}{\partial v}\right)\delta v + v\,\delta p. \end{aligned} \tag{3.9}$$

Thus, by Eqs. (3.8) and (3.9), it reads that Eq. (3.4) is the intrinsic representation of Eq. (3.2). □

Proposition 3.3. *A curve $x(t) = (q(t), v(t), p(t))$, $t_1 \leq t \leq t_2$, in $TQ \oplus T^*Q$ joining $\mathrm{pr}_Q(x(t_1)) = q_1$ and $\mathrm{pr}_Q(x(t_2)) = q_2$ satisfies*

$$(T\mathrm{pr}_{T^*Q})^*\chi(x(t), \dot{x}(t)) = (\tau_{TQ\oplus T^*Q})^*\mathbf{d}E(x(t), \dot{x}(t)), \tag{3.10}$$

if and only if $x(t)$ is a stationary point of the Hamilton–Pontryagin integral in Eq. (3.3).

Proof. The stationary condition of the Hamilton–Pontryagin integral is given by

$$\begin{aligned} &\delta \int_{t_1}^{t_2} \{L(q(t), v(t)) + p(t) \cdot (\dot{q}(t) - v(t))\}\ \mathrm{d}t \\ &= \delta \int_{t_1}^{t_2} \{p(t) \cdot \dot{q}(t) - E(q(t), v(t), p(t))\}\ \mathrm{d}t \\ &= \delta \int_{t_1}^{t_2} \{G(\rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q}(x, \dot{x})) - E(\tau_{TQ\oplus T^*Q}(x, \dot{x}))\}\,\mathrm{d}t \\ &= \int_{t_1}^{t_2} \left\{(T\mathrm{pr}_{T^*Q})^*\chi(x, \dot{x}) - (\tau_{TQ\oplus T^*Q})^*\mathbf{d}E(x, \dot{x})\right\} \cdot w\ \mathrm{d}t \\ &\quad + \Theta_{T^*Q}\left(\mathrm{pr}_{T^*Q} \circ (\tau_{TQ\oplus T^*Q})(x, \dot{x})\right) \cdot \left(T\tau_{T^*Q} \circ T\left(T\mathrm{pr}_{T^*Q}\right)(w)\right)\Big|_{t_1}^{t_2} \\ &= 0, \end{aligned}$$

which satisfies for all $w \in T_{(x,\dot{x})}T(TQ \oplus T^*Q)$. Since the endpoints of $q(t)$ are fixed, one has

$$\begin{aligned} \Theta_{T^*Q}\left(\mathrm{pr}_{T^*Q} \circ (\tau_{TQ\oplus T^*Q})(x, \dot{x})\right) \cdot \left(T\tau_{T^*Q} \circ T\left(T\mathrm{pr}_{T^*Q}\right)(w)\right)\Big|_{t_1}^{t_2} &= p\,\delta q\Big|_{t_1}^{t_2} \\ &= 0. \end{aligned}$$

Thus, we obtain Eq. (3.10). □

Needless to say, Eq. (3.10) is the intrinsic expression of Eq. (3.1), and so we shall call Eq. (3.10) the *intrinsic implicit Euler–Lagrange equations*, which are the implicit Lagrangian system (L, Δ_Q, X) for the case in which $\Delta_Q = TQ$.

3.5. The Lagrange–d'Alembert–Pontryagin principle

Next, we shall investigate an implicit Lagrangian system for the case in which a regular constraint distribution is given. To do this, we introduce an extended Lagrange–d'Alembert principle called the Lagrange–d'Alembert–Pontryagin principle.

Let L be a Lagrangian on TQ and $\Delta_Q \subset TQ$ be a constraint distribution on Q. Define a generalized energy $E(q, v, p) = p\cdot v - L(q, v)$ on $TQ \oplus T^*Q$. Keeping the endpoints of $q(t)$ fixed, the *Lagrange–d'Alembert–Pontryagin principle* is expressed by

$$\begin{aligned}\delta \int_{t_1}^{t_2} \{L(q(t), v(t)) + p(t) \cdot (\dot{q}(t) - v(t))\} \, \mathrm{d}t &= \delta \int_{t_1}^{t_2} \{p(t) \cdot \dot{q}(t) - E(q(t), v(t), p(t))\} \, \mathrm{d}t \\ &= \int_{t_1}^{t_2} \left\{ \left(\frac{\partial L}{\partial q} - \dot{p} \right) \delta q + \left(\frac{\partial L}{\partial v} - p \right) \delta v + (\dot{q} - v) \delta p \right\} \mathrm{d}t \\ &= 0, \end{aligned} \tag{3.11}$$

where we choose a variation $(\delta q(t), \delta v(t), \delta p(t))$ of the curve $(q(t), v(t), p(t))$, $t_1 \leq t \leq t_2$, in $TQ \oplus T^*Q$ such that $\delta q(t) \in \Delta_Q(q(t))$ and with the constraint $v(t) \in \Delta_Q(q(t))$. Hence, the Lagrange–d'Alembert–Pontryagin principle is represented by

$$\int_{t_1}^{t_2} \left\{ \left(\frac{\partial L}{\partial q} - \dot{p} \right) \delta q + \left(\frac{\partial L}{\partial v} - p \right) \delta v + (\dot{q} - v) \delta p \right\} \, \mathrm{d}t = 0,$$

which is equivalent to the equation

$$\left(\frac{\partial L}{\partial q} - \dot{p} \right) \delta q + \left(\frac{\partial L}{\partial v} - p \right) \delta v + (\dot{q} - v) \, \delta p = 0 \tag{3.12}$$

for all variations $\delta q(t) \in \Delta_Q(q(t))$, for arbitrary $\delta p(t)$ and $\delta v(t)$, and with the constraint $v(t) \in \Delta_Q(q(t))$.

Proposition 3.4. *Let a distribution Δ_Q be locally denoted by $\Delta(q) \subset \mathbb{R}^n$ at each $q \in U \subset \mathbb{R}^n$. The Lagrange–d'Alembert–Pontryagin principle for a curve $(q(t), v(t), p(t))$ provides equations of motion, in coordinates, such that*

$$\dot{q} = v, \quad \dot{p} - \frac{\partial L}{\partial q} \in \Delta^\circ(q), \quad p = \frac{\partial L}{\partial v}, \quad v \in \Delta(q). \tag{3.13}$$

Proof. From Eq. (3.12), we obtain the second-order condition $\dot{q} = v$, the Legendre transform $p = \partial L / \partial v$, the equations of motion $\dot{p} - \partial L / \partial q \in \Delta^\circ$ and with the constraint $v \in \Delta(q)$. Thus, we obtain Eq. (3.13). □

Notice that, as shown in Part I, Eq. (3.13) is the local expression of an *implicit Lagrangian system*.

3.6. Constraint distributions

We shall define constraint distributions for the intrinsic expression of the Lagrange–d'Alembert–Pontryagin principle.

Consider a regular constraint distribution Δ_Q on Q. Let $\mathrm{pr}_Q : TQ \oplus T^*Q \to Q$ and $T\mathrm{pr}_Q : T(TQ \oplus T^*Q) \to TQ$. Let $\tau_{TQ \oplus T^*Q} : T(TQ \oplus T^*Q) \to TQ \oplus T^*Q$ and $T\tau_{TQ \oplus T^*Q} : TT(TQ \oplus T^*Q) \to T(TQ \oplus T^*Q)$.

Define the submanifold $\mathcal{K} \subset TQ \oplus T^*Q$ by

$$\mathcal{K} = \left\{ (q, v, p) \in TQ \oplus T^*Q | (q, v) \in \Delta_Q, \quad p = \frac{\partial L}{\partial v} \right\}, \tag{3.14}$$

and also define the distribution $\mathcal{B}$ on $TQ \oplus T^*Q$ by

$$\mathcal{B} = (T\mathrm{pr}_Q)^{-1}(\Delta_Q) \subset T(TQ \oplus T^*Q).$$

Let $\mathcal{C}$ be the restriction of $\mathcal{B}$ to $\mathcal{K}$; that is,

$$\mathcal{C} = \mathcal{B} \cap T\mathcal{K} \subset T(TQ \oplus T^*Q).$$

In the above, we assume that $\mathcal{C}$ is a regular distribution on $\mathcal{K}$. Furthermore, note that $T\mathrm{pr}_Q \circ T\tau_{TQ \oplus T^*Q} : TT(TQ \oplus T^*Q) \to TQ$ and define the distribution $\mathcal{F}$ on $T(TQ \oplus T^*Q)$ by

$$\mathcal{F} = \left(T\mathrm{pr}_Q \circ T\tau_{TQ \oplus T^*Q} \right)^{-1} (\Delta_Q) \subset TT(TQ \oplus T^*Q).$$

Let $\mathcal{G}$ be defined by the restriction of $\mathcal{F}$ to $\mathcal{C}$ as

$$\mathcal{G} = \mathcal{F} \cap T\mathcal{C} \subset TT(TQ \oplus T^*Q), \tag{3.15}$$

where we assume $\mathcal{G}$ is a regular distribution on $\mathcal{C}$.

3.7. The intrinsic Lagrange–d'Alembert–Pontryagin equations

We can intrinsically represent the Lagrange–d'Alembert–Pontryagin principle for a curve $x(t) = (q(t), v(t), p(t))$, $t_1 \leq t \leq t_2$, in $TQ \oplus T^*Q$, with the endpoints of $q(t)$ fixed, by

$$\begin{aligned}\delta \int_{t_1}^{t_2} \{L(q(t), v(t)) + p(t) \cdot (\dot{q}(t) - v(t))\} \, \mathrm{d}t &= \delta \int_{t_1}^{t_2} \{p(t) \cdot \dot{q}(t) - E(q(t), v(t), p(t))\} \, \mathrm{d}t \\ &= \delta \int_{t_1}^{t_2} \{G(\rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q}(x, \dot{x})) - E(\tau_{TQ\oplus T^*Q}(x, \dot{x}))\} \, \mathrm{d}t \\ &= \int_{t_1}^{t_2} \{(T\mathrm{pr}_{T^*Q})^* \chi(x, \dot{x}) - (\tau_{TQ\oplus T^*Q})^* \mathbf{d}E(x, \dot{x})\} \cdot w \, \mathrm{d}t \\ &= 0, \end{aligned} \tag{3.16}$$

which holds for all $w = (q, v, p, \dot{q}, \dot{v}, \dot{p}, \delta q, \delta v, \delta p, \delta\dot{q}, \delta\dot{v}, \delta\dot{p}) \in \mathcal{G}(x, \dot{x}) \subset T_{(x,\dot{x})}T(TQ \oplus T^*Q)$.

Proposition 3.5. *The Lagrange–d'Alembert–Pontryagin principle is equivalent to the equation*

$$(T\mathrm{pr}_{T^*Q})^* \chi(x(t), \dot{x}(t)) \cdot w(t) = (\tau_{TQ\oplus T^*Q})^* \mathbf{d}E(x(t), \dot{x}(t)) \cdot w(t) \tag{3.17}$$

for all $w \in \mathcal{G}(x(t), \dot{x}(t))$.

Proof. Note that Eq. (3.16) is the intrinsic expression of Eq. (3.11). From Eq. (3.16), it is obvious that we obtain Eq. (3.17), which is the intrinsic expression of Eq. (3.13). □

Notice that we have derived Eq. (3.17) from the variational viewpoint, and that the result is consistent with the geometry of the generalized Legendre transform that we studied in Section 2.

3.8. The intrinsic form of implicit Lagrangian systems

Let X be a vector field on T^*Q, defined at points of $P = \mathbb{F}L(\Delta_Q)$, and let $\widetilde{X}$ be a choice of vector field on $TQ \oplus T^*Q$, defined at points (v_q, p_q) of $\mathcal{K}$ (defined by Eq. (3.14)) in the following way. Let $X(p_q)$ be tangent to a curve $p_{q(t)}(t) \in T^*_{q(t)}Q$. Then consider a curve in $TQ \oplus T^*Q$ having the form $v_{q(t)}(t)$ in the first component, where $v_{q(0)}(0) = v_q$ but the curve is otherwise arbitrary, and the given curve $p_{q(t)}(t) \in T^*_{q(t)}Q$ in the second component. The tangent to this curve defines the value of $\widetilde{X}(v_q, p_q)$. Of course this vector field is not unique. However, the vector field $\widetilde{X}$ has the property that, for each $x = (q, v, p)$ in $\mathcal{K} \subset TQ \oplus T^*Q$,

$$T\mathrm{pr}_{T^*Q}(\widetilde{X}(x)) = X(\mathrm{pr}_{T^*Q}(x)), \tag{3.18}$$

where $\mathrm{pr}_{T^*Q} : TQ \oplus T^*Q \to T^*Q$. If the curve $x(t) = (q(t), v(t), p(t))$ is an integral curve of $\widetilde{X}$, then, it follows that

$$\begin{aligned}\widetilde{X}(x(t)) &= \frac{\mathrm{d}x(t)}{\mathrm{d}t} \\ &= (q(t), v(t), p(t), \dot{q}(t), \dot{v}(t), \dot{p}(t)).\end{aligned}$$

Proposition 3.6. *Let $x(t)$, $t_1 \leq t \leq t_2$, be an integral curve of the vector field $\widetilde{X}$ on $TQ \oplus T^*Q$ that is naturally induced from a vector field X on T^*Q as in Eq. (3.18). If $x(t)$ is a solution curve of the Lagrange–d'Alembert–Pontryagin principle, then it satisfies*

$$(T\mathrm{pr}_{T^*Q})^* \chi(\widetilde{X}(x(t))) \cdot w(t) = (\tau_{TQ\oplus T^*Q})^* \mathbf{d}E(\widetilde{X}(x(t))) \cdot w(t) \tag{3.19}$$

for a given variation $w(t) \in \mathcal{G}(\widetilde{X}(x(t)))$, where $\mathcal{G}$ is the regular distribution defined by Eq. (3.15).

Proof. Since $x(t)$ is the integral curve of the induced vector field $\widetilde{X}$, we have $\dot{x}(t) = \widetilde{X}(x(t))$. Thus, we obtain Eq. (3.19) by substituting $\dot{x}(t) = \widetilde{X}(x(t))$ into Eq. (3.17). □

We shall call Eq. (3.19) the *intrinsic implicit Lagrangian system*.

3.9. Variational link with Dirac structures

Next, we shall see the variational link of an implicit Lagrangian systems (L, Δ_Q, X) with the induced Dirac structure D_{Δ_Q} on T^*Q.

Recall from Part I that D_{Δ_Q} is defined, for each $z \in T^*Q$, by

$$\begin{aligned} D_{\Delta_Q}(z) = \{(u_z, \alpha_z) \in T_zT^*Q \times T_z^*T^*Q \mid u_z \in \Delta_{T^*Q}(z), \text{ and} \\ \alpha_z(w_z) = \Omega_{\Delta_Q}(u_z, w_z) \text{ for all } w_z \in \Delta_{T^*Q}(z)\}, \end{aligned} \tag{3.20}$$

where $\Delta_{T^*Q} = (T\pi_Q)^{-1}(\Delta_Q)$ and Ω_{Δ_Q} is the restriction of the canonical symplectic form Ω on T^*Q to Δ_{T^*Q}.

As to the variational link of implicit Lagrangian systems, we have the following proposition.

Proposition 3.7. *If a curve $x(t) = (q(t), v(t), p(t))$ is an integral curve of the induced vector field $\widetilde{X}$ that is associated with Eq. (3.19), then the curve $x(t) = (q(t), v(t), p(t))$ is a solution curve of an implicit Lagrangian system (L, Δ_Q, X), which satisfies, for each $v(t) \in \Delta_Q(q(t))$,*

$$(X(q(t), p(t)), \mathfrak{D}L(q(t), v(t))) \in D_{\Delta_Q}(q(t), p(t)),$$

where $(q(t), p(t)) = \mathbb{F}L(q(t), v(t))$ is an integral curve of X.

Proof. It is logically obvious that the above proposition holds; however, we shall prove this by direct computations.

Let us rewrite the left-hand side of Eq. (3.19). Recall that the one-form χ on TT^*Q is defined by $\chi = (\Omega^\flat)^*\Theta_{T^*T^*Q}$, and we have

$$\begin{aligned} (T\mathrm{pr}_{T^*Q})^*\chi(\widetilde{X}(x)) \cdot w &= \chi(T\mathrm{pr}_{T^*Q}(\widetilde{X}(x))) \cdot T_{\widetilde{X}(x)}T\mathrm{pr}_{T^*Q}(w) \\ &= \Theta_{T^*T^*Q}\left(\Omega^\flat \circ T\mathrm{pr}_{T^*Q}(\widetilde{X}(x))\right) \cdot T\Omega^\flat\left(T_{\widetilde{X}(x)}(T\mathrm{pr}_{T^*Q})(w)\right). \end{aligned} \tag{3.21}$$

Recall also that the canonical one-form on T^*T^*Q is defined by

$$\Theta_{T^*T^*Q}(\alpha) \cdot V = \left\langle \alpha, T\pi_{T^*Q}(V) \right\rangle,$$

where $\alpha \in T^*T^*Q$, $V \in T_\alpha(T^*T^*Q)$, $\pi_{T^*Q} : T^*T^*Q \to T^*Q$ is the canonical projection and $T\pi_{T^*Q} : TT^*T^*Q \to TT^*Q$. Using local coordinates $x = (q, v, p)$, Eq. (3.21) further reads that

$$\begin{aligned} &\Theta_{T^*T^*Q}\left(\Omega^\flat \circ T\mathrm{pr}_{T^*Q}(\widetilde{X}(x))\right) \cdot \left(T\Omega^\flat \circ T_{\widetilde{X}(x)}(T\mathrm{pr}_{T^*Q})(w)\right) \\ &= \Omega^\flat \circ T\mathrm{pr}_{T^*Q}(\widetilde{X}(x)) \cdot T\pi_{T^*Q}\left(T\Omega^\flat \circ T_{\widetilde{X}(x)}\left(T\mathrm{pr}_{T^*Q}\right)(w)\right) \\ &= -\dot{p}\,\delta q + \dot{q}\,\delta p, \end{aligned} \tag{3.22}$$

where, using local coordinates, $\widetilde{X}(x) = (q, v, p, \dot{q}, \dot{v}, \dot{p})$ and hence $\Omega^\flat \circ T\mathrm{pr}_{T^*Q}(\widetilde{X}(x)) = \Omega^\flat(q, p, \dot{q}, \dot{p}) = (q, p, -\dot{p}, \dot{q})$, and further, noting that $w = (q, v, p, \dot{q}, \dot{v}, \dot{p}, \delta q, \delta v, \delta p, \delta\dot{q}, \delta\dot{v}, \delta\dot{p})$,

$$\begin{aligned} T\pi_{T^*Q}\left(T\Omega^\flat \circ T_{\widetilde{X}(x)}\left(T\mathrm{pr}_{T^*Q}\right)(w)\right) &= T\pi_{T^*Q}(q, p, -\dot{p}, \dot{q}, \delta q, \delta p, -\delta\dot{p}, \delta\dot{q}) \\ &= (q, p, \delta q, \delta p). \end{aligned}$$

On the other hand, the right-hand side of Eq. (3.19) is locally expressed by

$$\begin{aligned} \left(\tau_{TQ\oplus T^*Q}\right)^*\mathbf{d}E(\widetilde{X}(x)) \cdot w &= \mathbf{d}E\left(\tau_{TQ\oplus T^*Q}(\widetilde{X}(x))\right) \cdot T_{\widetilde{X}(x)}\tau_{TQ\oplus T^*Q}(w) \\ &= \left(\frac{\partial E}{\partial q}\mathrm{d}q + \frac{\partial E}{\partial v}\mathrm{d}v + \frac{\partial E}{\partial p}\mathrm{d}p\right) \cdot \left(\delta q\frac{\partial}{\partial q} + \delta v\frac{\partial}{\partial v} + \delta p\frac{\partial}{\partial p}\right) \\ &= \left(-\frac{\partial L}{\partial q}\right)\delta q + \left(p - \frac{\partial L}{\partial v}\right)\delta v + v\,\delta p \\ &= \left(-\frac{\partial L}{\partial q}\right)\delta q + v\,\delta p. \end{aligned}$$

In the above, note that $p = \partial L/\partial v$ on $\mathcal{K}$ and also that $\tau_{TQ\oplus T^*Q}(\widetilde{X}(x)) = x$ and

$$T_{\widetilde{X}(x)}\tau_{TQ\oplus T^*Q}(w) = (q, v, p, \delta q, \delta v, \delta p) \in \mathcal{C}(x).$$

Using the projection $\mathrm{pr}_{T^*Q} : TQ \oplus T^*Q \to T^*Q$ and its differential map $T\mathrm{pr}_{T^*Q} : T(TQ \oplus T^*Q) \to T(T^*Q)$, it follows that

$$\begin{aligned}(\mathrm{pr}_{T^*Q})^*(\tau_{TQ\oplus T^*Q})^*\mathbf{d}E(\widetilde{X}(x)) \cdot w &= (\mathrm{pr}_{T^*Q})^*\mathbf{d}E\,(x) \cdot \left(T_{\widetilde{X}(x)}\tau_{TQ\oplus T^*Q}(w)\right)\\ &= \mathbf{d}E\left(\mathrm{pr}_{T^*Q}(x)\right) \cdot T\mathrm{pr}_{T^*Q}\left(T_{\widetilde{X}(x)}\tau_{TQ\oplus T^*Q}(w)\right)\\ &= \frac{\partial E}{\partial q}\delta q + \frac{\partial E}{\partial p}\delta p\\ &= \left(-\frac{\partial L}{\partial q}\right)\delta q + v\,\delta p,\end{aligned} \tag{3.23}$$

while the Dirac differential of a Lagrangian, that is, $\mathfrak{D}L : TT^*Q \to T^*T^*Q$, is defined by using the diffeomorphism $\gamma_Q = \Omega^\flat \circ (\kappa_Q)^{-1} : T^*TQ \to T^*T^*Q$, as shown in Part I, such that

$$\begin{aligned}\mathfrak{D}L &= \gamma_Q \circ \mathbf{d}L\\ &= \left(q, \frac{\partial L}{\partial v}, -\frac{\partial L}{\partial q}, v\right).\end{aligned}$$

Recall that $\mathrm{pr}_{TQ} : TQ \oplus T^*Q \to TQ$ is given in coordinates by $\mathrm{pr}_{TQ}(q, v, p) = (q, v)$, and we obtain

$$\begin{aligned}\mathbf{d}E\left(\mathrm{pr}_{T^*Q}(x)\right) &= \mathfrak{D}L\left(\mathrm{pr}_{TQ}(x)\right)\\ &= \left(-\frac{\partial L}{\partial q}\right)\mathrm{d}q + v\,\mathrm{d}p,\end{aligned} \tag{3.24}$$

where $p = \partial L/\partial v$. From Eqs. (3.19) and (3.21)–(3.24), we have

$$\begin{aligned}&\Omega^\flat\left(T\mathrm{pr}_{T^*Q}(\widetilde{X}(x))\right) \cdot T\pi_{T^*Q}\left(T\Omega^\flat \circ T_{\widetilde{X}(x)}\left(T\mathrm{pr}_{T^*Q}\right)(w)\right)\\ &\quad= \mathbf{d}E\left(\mathrm{pr}_{T^*Q}(x)\right) \cdot T\mathrm{pr}_{T^*Q}\left(T_{\widetilde{X}(x)}\tau_{TQ\oplus T^*Q}(w)\right)\\ &\quad= \mathfrak{D}L\left(\mathrm{pr}_{TQ}(x)\right) \cdot T\mathrm{pr}_{T^*Q}\left(T_{\widetilde{X}(x)}\tau_{TQ\oplus T^*Q}(w)\right).\end{aligned} \tag{3.25}$$

Notice that there exists the identity

$$T\pi_{T^*Q}\left(T\Omega^\flat \circ T\left(T\mathrm{pr}_{T^*Q}\right)\right) = T\mathrm{pr}_{T^*Q} \circ T\tau_{TQ\oplus T^*Q},$$

and we set

$$\begin{aligned}\delta z &= T\pi_{T^*Q}\left(T\Omega^\flat \circ T_{\widetilde{X}(x)}\left(T\mathrm{pr}_{T^*Q}\right)(w)\right)\\ &= T\mathrm{pr}_{T^*Q}\left(T_{\widetilde{X}(x)}\tau_{TQ\oplus T^*Q}(w)\right)\\ &= (q, p, \delta q, \delta p),\end{aligned} \tag{3.26}$$

where $z = \mathrm{pr}_{T^*Q}(x)$ and hence $z = (q, p)$.

From Eqs. (3.25) and (3.26), it follows that Eq. (3.19) can be restated as

$$\Omega^\flat\left(X(z(t))\right) \cdot \delta z(t) = \mathfrak{D}L(q(t), v(t)) \cdot \delta z(t)$$

for all $\delta z(t) = (\delta q(t), \delta p(t)) \in \Delta_{T^*Q}(z(t))$, where $z(t) = \mathbb{F}L(q(t), v(t))$ is an integral curve of a vector field X on T^*Q and $\Delta_{T^*Q} = (T\pi_Q)^{-1}(\Delta_Q) \subset TT^*Q$. In other words, a curve $(q(t), v(t)) \in \Delta_Q(q(t))$ satisfies

$$\Omega_{\Delta_Q}\left(X(z(t)),\ \delta z(t)\right) = \mathfrak{D}L(q(t), v(t)) \cdot \delta z(t), \tag{3.27}$$

for all $\delta z(t) \in \Delta_{T^*Q}(z(t))$, where $z(t) = \mathbb{F}L(q(t), v(t))$ is the integral curve of X, defined at points in $P = \mathbb{F}L(\Delta_Q)$.

As shown in Eq. (3.20), the set of Δ_{T^*Q} and the skew-symmetric bilinear form Ω_{Δ_Q} defines an *induced Dirac structure* D_{Δ_Q} on T^*Q. Then, needless to say, Eq. (3.27) represents the condition of an implicit Lagrangian system (L, Δ_Q, X) associated with the induced Dirac structure D_{Δ_Q} on T^*Q; that is,

$$(X, \mathfrak{D}L) \in D_{\Delta_Q}$$

together with the Legendre transform $P = \mathbb{F}L(\Delta_Q)$. $\square$

Remarks. From Eq. (3.24), the following relation is satisfied for each $(q, v) \in \Delta_Q$:

$$\left(X(q, p), \mathbf{d}E(q, v, p)|_{TT^*Q}\right) \in D_{\Delta_Q}(q, p),$$

where $(q, p) = \mathbb{F}L(q, v)$ and the restriction $\mathbf{d}E(q, v, p)|_{TT^*Q}$ is understood in the sense that TT^*Q is naturally included in $T(TQ \oplus T^*Q)$. Hence, the condition for an implicit Lagrangian system (L, Δ_Q, X), namely, $(X, \mathfrak{D}L) \in D_{\Delta_Q}$, can be restated as

$$\left(X, \mathbf{d}E|_{TT^*Q}\right) \in D_{\Delta_Q}$$

together with the Legendre transform $P = \mathbb{F}L(\Delta_Q)$.

We can summarize the results so far obtained in the following theorem.

Theorem 3.8. *Let L be a Lagrangian on TQ (possibly degenerate) and Δ_Q be a constraint distribution on Q. Let X be a vector field on T^*Q, defined at points of $P = \mathbb{F}L(\Delta_Q)$, such that (L, Δ_Q, X) is an implicit Lagrangian system. Denote by $x(t) = (q(t), v(t), p(t))$, $t_1 \le t \le t_2$, a curve in $TQ \oplus T^*Q$. The following statements are equivalent:*

(a) *$x(t)$ is a solution curve of the implicit Lagrangian system (L, Δ_Q, X);*
(b) *$x(t)$ satisfies the Lagrange–d'Alembert–Pontryagin principle in Eq. (3.16);*
(c) *$x(t)$ is the integral curve of a choice of vector field $\tilde{X}$ on $TQ \oplus T^*Q$ that is naturally induced from X.*

3.10. Hamilton's phase space principle

If a given Lagrangian is hyperregular, then, a hyperregular Hamiltonian is well defined on the cotangent bundle via the Legendre transform. Hence, we can also develop an *implicit Hamiltonian system* for the regular case in terms of the induced Dirac structure on the cotangent bundle. Before going into the construction of implicit Hamiltonian systems, we first show how *intrinsic Hamilton's equations* can be developed in the context of Hamilton's phase space principle.

Let L be a hyperregular Lagrangian on TQ. Define the energy E on TQ, by employing local coordinates (q, v) for TQ, such that

$$E(q, v) = \frac{\partial L}{\partial v} \cdot v - L(q, v).$$

Since the Legendre transform $\mathbb{F}L : TQ \to T^*Q$ is diffeomorphism, a hyperregular Hamiltonian H can be defined on T^*Q such that

$$H = E \circ (\mathbb{F}L)^{-1}.$$

Then, define the path space of curves $z(t) = (q(t), p(t))$, $t_1 \le t \le t_2$, in T^*Q as

$$S(q_1, q_2, [t_1, t_2]) = \{z = (q, p) : [t_1, t_2] \to T^*Q \mid \pi_Q(z(t_1)) = q_1, \pi_Q(z(t_2)) = q_2\},$$

where $\pi_Q : T^*Q \to Q$ and also define the action functional on the path space $S(q_1, q_2, [t_1, t_2])$ of curves $z(t) = (q(t), p(t))$, $t_1 \le t \le t_2$, by

$$\int_{t_1}^{t_2} \{p(t) \cdot \dot{q}(t) - H(q(t), p(t))\}\, \mathrm{d}t, \tag{3.28}$$

which is called the *Poincaré–Cartan integral*.

Proposition 3.9. *Keeping the endpoints $q(t_1)$ and $q(t_2)$ of $q(t)$ fixed whereas $p(t_1)$ and $p(t_2)$ of $p(t)$ are allowed to be free, the stationary condition for the Poincaré–Cartan integral in Eq. (3.28) gives Hamilton's equations*

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}. \tag{3.29}$$

Proof. The variation of the Poincaré–Cartan integral is locally represented by

$$\begin{aligned} \delta \int_{t_1}^{t_2} \{p(t)\cdot\dot{q}(t) - H(q(t), p(t))\}\, \mathrm{d}t &= \int_{t_1}^{t_2} \left(\delta p\, \dot{q} + p\, \delta\dot{q} - \frac{\partial H}{\partial q}\delta q - \frac{\partial H}{\partial p}\delta p\right) \mathrm{d}t \\ &= \int_{t_1}^{t_2} \left\{\left(\dot{q} - \frac{\partial H}{\partial p}\right)\delta p + \left(-\dot{p} - \frac{\partial H}{\partial q}\right)\delta q\right\} \mathrm{d}t + p\,\delta q\Big|_{t_1}^{t_2}. \end{aligned} \tag{3.30}$$

Keeping the endpoints $q(t_1)$ and $q(t_2)$ of $q(t)$ fixed, the stationary condition for the Poincaré–Cartan integral gives Hamilton's equations in Eq. (3.29). □

3.11. The intrinsic form of Hamilton's equations

Let us demonstrate the intrinsic expression for Hamilton's phase space principle in the following.

Proposition 3.10. *Let Θ_{T^*Q} be the canonical one-form on T^*Q and $\Omega^\flat : TT^*Q \to T^*T^*Q$ be the bundle map associated with the canonical symplectic form Ω. Let $\Theta_{T^*T^*Q}$ be the canonical one-form on T^*T^*Q. Let $\tau_{T^*Q} : TT^*Q \to T^*Q$ be the tangent projection and $\rho_{TT^*Q} : TT^*Q \to TQ \oplus T^*Q$ be the projection. Denote by H a hyperregular Hamiltonian on T^*Q. Let $G(q, v, p) = p \cdot v$ be the momentum function on $TQ \oplus T^*Q$, where $(q, v, p) \in TQ \oplus T^*Q$. The variation of the Poincaré–Cartan integral in Eq. (3.28) is represented by*

$$\begin{aligned} &\delta \int_{t_1}^{t_2} \{p(t)\cdot\dot{q}(t) - H(q(t), p(t))\}\, \mathrm{d}t \\ &= \delta \int_{t_1}^{t_2} \left\{G(\rho_{TT^*Q}(z,\dot{z})) - H(\tau_{T^*Q}(z,\dot{z}))\right\} \mathrm{d}t \\ &= \int_{t_1}^{t_2} \left\{\Theta_{T^*T^*Q}(\Omega^\flat_z(\dot{z})) \cdot T\Omega^\flat(w) - \mathbf{d}H(\tau_{T^*Q}(z,\dot{z})) \cdot T\tau_{T^*Q}(w)\right\} \mathrm{d}t + \Theta_{T^*Q}(\tau_{T^*Q}(z,\dot{z})) \cdot T\tau_{T^*Q}(w)\Big|_{t_1}^{t_2} \\ &= \int_{t_1}^{t_2} \left\{\chi(z,\dot{z}) - (\tau_{T^*Q})^*\mathbf{d}H(z,\dot{z})\right\} \cdot w\, \mathrm{d}t + \Theta_{T^*Q}(\tau_{T^*Q}(z,\dot{z})) \cdot T\tau_{T^*Q}(w)\Big|_{t_1}^{t_2}, \end{aligned} \tag{3.31}$$

where $z = (q, p) \in T^*Q$, $\dot{z} = \mathrm{d}z/\mathrm{d}t \in T_zT^*Q$ and $w \in T_{(z,\dot{z})}(TT^*Q)$.

Proof. Let us check that Eq. (3.31) is the intrinsic representation of Eq. (3.30). Recall that the one-form χ on TT^*Q is defined by the canonical one-form $\Theta_{T^*T^*Q}$ on T^*T^*Q such that

$$\chi = (\Omega^\flat)^*\Theta_{T^*T^*Q}.$$

Since $(z, \dot{z}) \in TT^*Q$ is locally denoted by $(q, p, \dot{q}, \dot{p})$, we have

$$\chi(z,\dot{z}) = -\dot{p}\,\mathrm{d}q + \dot{q}\,\mathrm{d}p \in T^*_{(z,\dot{z})}(TT^*Q).$$

Then, noting $T\Omega^\flat : T(TT^*Q) \to T(T^*T^*Q)$, the following relation holds:

$$\begin{aligned} \Theta_{T^*T^*Q}(\Omega^\flat_z(\dot{z})) \cdot T\Omega^\flat(w) &= \chi(z,\dot{z}) \cdot w \\ &= -\dot{p}\,\delta q + \dot{q}\,\delta p \end{aligned}$$

for all $\dot{z} \in T_zT^*Q$ and $w = (q, p, \dot{q}, \dot{p}, \delta q, \delta p, \delta\dot{q}, \delta\dot{p}) \in T_{(z,\dot{z})}(TT^*Q)$.

On the other hand, recall that the canonical one-form $\Theta_{T^*T^*Q}$ is defined such that

$$\Theta_{T^*T^*Q}(\Omega^\flat_z(\dot{z})) \cdot T\Omega^\flat(w) = \Omega^\flat_z(\dot{z}) \cdot T\pi_{T^*Q}(T\Omega^\flat(w)), \tag{3.32}$$

where $\pi_{T^*Q} : T^*T^*Q \to T^*Q$ is the cotangent projection, $\Omega_z^\flat(\dot{z}) = (q, p, -\dot{p}, \dot{q}) \in T^*T^*Q$ and $T\pi_{T^*Q}(T\Omega^\flat(w)) = (q, p, \delta q, \delta p) \in TT^*Q$.

The differential of the Hamiltonian $\mathbf{d}H : T^*Q \to T^*T^*Q$ is locally denoted by

$$\mathbf{d}H = \left(q, p, \frac{\partial H}{\partial q}, \frac{\partial H}{\partial p}\right).$$

By the projection $T\tau_{T^*Q} : T(TT^*Q) \to TT^*Q$, we have

$$\begin{aligned} T\tau_{T^*Q}(w) &= T\tau_{T^*Q}(q, p, \dot{q}, \dot{p}, \delta q, \delta p, \delta\dot{q}, \delta\dot{p}) \\ &= (q, p, \delta q, \delta p). \end{aligned}$$

Hence, it follows that, for all $z \in T^*Q$ and for all $w \in T_{(z,\dot{z})}TT^*Q$,

$$\begin{aligned} \mathbf{d}H(\tau_{T^*Q}(z,\dot{z})) \cdot T\tau_{T^*Q}(w) &= (\tau_{T^*Q})^*\mathbf{d}H(z,\dot{z}) \cdot w \\ &= \frac{\partial H}{\partial q}\,\delta q + \frac{\partial H}{\partial p}\delta p. \end{aligned} \tag{3.33}$$

Furthermore, we have the local expression

$$\Theta_{T^*Q}(\tau_{T^*Q}(z,\dot{z})) \cdot T\tau_{T^*Q}(w) = p\,\delta q. \tag{3.34}$$

From Eqs. (3.32) to (3.34), it immediately reads that Eq. (3.31) is the intrinsic representation of Eq. (3.30). □

Remarks. Recall that $\theta = \lambda \oplus \chi$ is the one-form on $TT^*Q \times TT^*Q$ and also that $\Psi^*\theta = \mathbf{d}(G \circ \rho_{TT^*Q})$ using the diagonal map $\Psi : TT^*Q \to TT^*Q \times TT^*Q$. Then, the following relation is satisfied, for each $z \in T^*Q$,

$$\begin{aligned} \delta \int_{t_1}^{t_2} G(\rho_{TT^*Q}(z,\dot{z}))\,\mathrm{d}t &= \int_{t_1}^{t_2} \mathbf{d}G(\rho_{TT^*Q}(z,\dot{z})) \cdot T\rho_{TT^*Q}(w)\,\mathrm{d}t \\ &= \int_{t_1}^{t_2} \Psi^*\theta(z,\dot{z}) \cdot w\,\mathrm{d}t \\ &= \int_{t_1}^{t_2} \{\Theta_{T^*T^*Q}(\Omega_z^\flat(\dot{z})) \cdot T\Omega^\flat(w) + \Theta_{T^*TQ}(\kappa_Q(z,\dot{z})) \cdot T\kappa_Q(w)\}\,\mathrm{d}t \\ &= \int_{t_1}^{t_2} \Theta_{T^*T^*Q}(\Omega_z^\flat(\dot{z})) \cdot T\Omega^\flat(w)\,\mathrm{d}t + \frac{\mathrm{d}}{\mathrm{d}t}\int_{t_1}^{t_2} \Theta_{T^*Q}(\tau_{T^*Q}(z,\dot{z})) \cdot T\tau_{T^*Q}(w)\,\mathrm{d}t \\ &= \int_{t_1}^{t_2} \Theta_{T^*T^*Q}(\Omega_z^\flat(\dot{z})) \cdot T\Omega^\flat(w)\,\mathrm{d}t + \Theta_{T^*Q}(\tau_{T^*Q}(z,\dot{z})) \cdot T\tau_{T^*Q}(w)\Big|_{t_1}^{t_2}. \end{aligned}$$

Using local coordinates $z = (q, p) \in T^*Q$, $\dot{z} = (\dot{q}, \dot{p}) \in T_zT^*Q$, and $w = (\delta q, \delta p, \delta\dot{q}, \delta\dot{p}) \in T_{(z,\dot{z})}(TT^*Q)$, it is easy to check the above relation:

$$\begin{aligned} \delta \int_{t_1}^{t_2} G(q, \dot{q}, p)\,\mathrm{d}t &= \delta \int_{t_1}^{t_2} (p \cdot \dot{q})\;\mathrm{d}t \\ &= \int_{t_1}^{t_2} (p\,\delta\dot{q} + \dot{q}\,\delta p)\;\mathrm{d}t \\ &= \int_{t_1}^{t_2} \{(-\dot{p}\,\delta q + \dot{q}\,\delta p) + (\dot{p}\,\delta q + p\,\delta\dot{q})\}\,\mathrm{d}t \\ &= \int_{t_1}^{t_2} (-\dot{p}\,\delta q + \dot{q}\,\delta p)\;\mathrm{d}t + \frac{\mathrm{d}}{\mathrm{d}t}\int_{t_1}^{t_2} p\,\delta q\,\mathrm{d}t \\ &= \int_{t_1}^{t_2} (-\dot{p}\,\delta q + \dot{q}\,\delta p)\;\mathrm{d}t + p\,\delta q\Big|_{t_1}^{t_2}. \end{aligned}$$

Proposition 3.11. *Keeping the endpoints $\pi_Q(z(t_1)) = q(t_1)$ and $\pi_Q(z(t_2)) = q(t_2)$ of the curve $\pi_Q(z(t)) = q(t)$ fixed, the stationary condition for the Poincaré–Cartan integral in Eq. (3.31) provides intrinsic Hamilton's equations such that, for each $z = (q, p) \in T^*Q$,*

$$\chi(z, \dot{z}) = (\tau_{T^*Q})^* \mathbf{d}H(z, \dot{z}). \tag{3.35}$$

Proof. The stationary condition of the Poincaré–Cartan integral is given by

$$\begin{aligned}
&\delta \int_{t_1}^{t_2} \{p(t) \cdot \dot{q}(t) - H(q(t), p(t))\} \, \mathrm{d}t \\
&\quad = \delta \int_{t_1}^{t_2} \left\{G(\rho_{TT^*Q}(z, \dot{z})) - H(\tau_{T^*Q}(z, \dot{z}))\right\} \, \mathrm{d}t \\
&\quad = \int_{t_1}^{t_2} \left\{\chi(z, \dot{z}) - (\tau_{T^*Q})^* \mathbf{d}H(z, \dot{z})\right\} \cdot w \, \mathrm{d}t + \Theta_{T^*Q}(\tau_{T^*Q}(z, \dot{z})) \cdot T\tau_{T^*Q}(w) \Big|_{t_1}^{t_2}
\end{aligned}$$

for all $w \in T_{(z,\dot{z})}(TT^*Q)$. Keeping the endpoints of $q(t)$ fixed, we can obtain the intrinsic Hamilton equations in Eq. (3.35). □

3.12. Implicit Hamiltonian systems

Let us illustrate an *implicit Hamiltonian system* for the case in which a hyperregular Hamiltonian is given on the cotangent bundle and with a constraint distribution on a configuration manifold.

Definition 3.12. Let H be a hyperregular Hamiltonian on T^*Q and $\Delta_Q \subset TQ$ be a constraint distribution on Q. Let X be a vector field on T^*Q and Ω be the canonical symplectic form on T^*Q. Let D_{Δ_Q} be the induced Dirac structure on T^*Q defined by Eq. (3.20).

Then, an *implicit Hamiltonian system* is the triple (H, Δ_Q, X) that satisfies, for each point $z \in T^*Q$,

$$(X(z), \mathbf{d}H(z)) \in D_{\Delta_Q}(z),$$

that is,

$$(X, \mathbf{d}H) \in D_{\Delta_Q}.$$

Definition 3.13. A *solution curve* of an implicit Hamiltonian system (H, Δ_Q, X) is a curve $(q(t), p(t))$, $t_1 \leq t \leq t_2$, in T^*Q such that $(q(t), p(t))$ is an integral curve of X.

Proposition 3.14. *Using local coordinates (q, p) for T^*Q, it follows from the condition $(X, \mathbf{d}H) \in D_{\Delta_Q}$ that the local expression for an implicit Hamiltonian system is given by*

$$\dot{q} = \frac{\partial H}{\partial p} \in \Delta(q), \quad \dot{p} + \frac{\partial H}{\partial q} \in \Delta^\circ(q), \tag{3.36}$$

where the distribution Δ_Q is locally denoted by $\Delta(q) \subset \mathbb{R}^n$ at each $q \in U \subset \mathbb{R}^n$.

Proof. Recall that the local expression for the canonical symplectic form is given by

$$\Omega\left((q, p, u_1, \alpha_1), (q, p, u_2, \alpha_2)\right) = \langle \alpha_2, u_1 \rangle - \langle \alpha_1, u_2 \rangle,$$

and the condition for an implicit Hamiltonian system (H, Δ_Q, X) is given by, for each $(q, p) \in T^*Q$,

$$(X(q, p), \mathbf{d}H(q, p)) \in D_{\Delta_Q}(q, p).$$

Using local expressions $X(q, p) = (\dot{q}, \dot{p})$ and $\mathbf{d}H(q, p) = (\partial H/\partial q, \partial H/\partial p)$, it follows that

$$\left\langle \frac{\partial H}{\partial q}, \delta q \right\rangle + \left\langle \frac{\partial H}{\partial p}, \delta p \right\rangle = \langle \delta p, \dot{q} \rangle - \langle \dot{p}, \delta q \rangle$$

for all $\delta q \in \Delta(q)$, for all δp, and with $\dot{q} \in \Delta(q)$. Thus, we obtain Eq. (3.36). □

Notice that Eq. (3.36) is the local expression for an *implicit Hamiltonian system*.

3.13. The Hamilton–d'Alembert principle in phase space

We now show how to obtain an *implicit Hamiltonian system* in the context of a generalization of Hamilton's phase space principle that we refer to as the *Hamilton–d'Alembert principle in phase space*.

The *Hamilton–d'Alembert principle in phase space* for a curve $(q(t), p(t)),\ t_1 \leq t \leq t_2$, in T^*Q is given by

$$\delta \int_{t_1}^{t_2} \{p(t) \cdot \dot{q}(t) - H(q(t), p(t))\}\ \mathrm{d}t = 0, \tag{3.37}$$

and with $\dot{q}(t) \in \Delta(q(t))$. The variation of the left-hand side in Eq. (3.37) is locally given, keeping the endpoints of $q(t)$ fixed, by

$$\delta \int_{t_1}^{t_2} \{p(t) \cdot \dot{q}(t) - H(q(t), p(t))\}\ \mathrm{d}t = \int_{t_1}^{t_2} \left\{ \left(\dot{q} - \frac{\partial H}{\partial p}\right) \delta p + \left(-\dot{p} - \frac{\partial H}{\partial q}\right) \delta q \right\} \mathrm{d}t, \tag{3.38}$$

where we choose a variation δq of curves $q(t)$ such that $\delta q \in \Delta(q)$.

Note that in the case of regular Lagrangians, if one starts with the Lagrange–d'Alembert–Pontryagin principle, and optimizes first over v, then one arrives at the Hamilton–d'Alembert principle in phase space.

Proposition 3.15. *The Hamilton–d'Alembert principle in phase space for a curve $(q(t), p(t))$, $t_1 \leq t \leq t_2$, in T^*Q gives the implicit Hamiltonian systems in Eq.* (3.36).

Proof. From Eq. (3.38), the Hamilton–d'Alembert principle in phase space is equivalent to

$$\left(\dot{q} - \frac{\partial H}{\partial p}\right) \delta p + \left(-\dot{p} - \frac{\partial H}{\partial q}\right) \delta q = 0$$

for all $\delta q \in \Delta(q)$, for all δp, and with $\dot{q} \in \Delta(q)$. Thus, we obtain Eq. (3.36). □

3.14. Coordinate representation

Suppose that the dimension of $\Delta(q)$ is $n - m$ at each point q. Let $\Delta^\circ(q)$ be the annihilator of $\Delta(q)$ spanned by m one-forms $\omega^1, \ldots, \omega^m$, and it follows that Eq. (3.36) can be represented, in coordinates, by employing the Lagrange multipliers $\mu_a,\ a = 1, \ldots, m$ such that

$$\begin{pmatrix} \dot{q}^i \\ \dot{p}_i \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \dfrac{\partial H}{\partial q^i} \\ \dfrac{\partial H}{\partial p_i} \end{pmatrix} + \begin{pmatrix} 0 \\ \mu_a\, \omega_i^a(q) \end{pmatrix},$$

$$0 = \omega_i^a(q)\, \frac{\partial H}{\partial p_i},$$

where we use the local expression $\omega^a = \omega_i^a\, \mathrm{d}q^i$.

3.15. Constraint distributions

Let $\Delta_Q \subset TQ$ be a constraint distribution on Q. Define the distribution on T^*Q by

$$\Delta_{T^*Q} = (T\pi_Q)^{-1}(\Delta_Q) \subset TT^*Q,$$

where $\pi_Q : T^*Q \to Q$ and $T\pi_Q : TT^*Q \to TQ$. Let P be defined by the image of Δ_Q under the Legendre transformation $\mathbb{F}L : TQ \to T^*Q$, that is, $P = \mathbb{F}L(\Delta_Q) \subset T^*Q$, and let Δ_P be the restriction of Δ_{T^*Q} to P such that

$$\Delta_P = \Delta_{T^*Q} \cap TP \subset TT^*Q,$$

where we assume that Δ_P is a regular distribution on P. Define the distribution on TT^*Q by

$$\mathcal{I} = (T\pi_Q \circ \tau_{TT^*Q})^{-1}(\Delta_Q) \subset T(TT^*Q),$$

where $\tau_{TT^*Q} : T(TT^*Q) \to TT^*Q$ and hence $T\pi_Q \circ \tau_{TT^*Q} : T(TT^*Q) \to TQ$. Let $\mathcal{J}$ be the restriction of $\mathcal{I}$ to Δ_P such that

$$\mathcal{J} = \mathcal{I} \cap T\Delta_P \subset T(TT^*Q),$$

where $\mathcal{J}$ is assumed to be a regular distribution on Δ_P.

3.16. Intrinsic implicit Hamiltonian systems

Let us see how the *intrinsic implicit Hamiltonian system* is related to the Hamilton–d'Alembert principle in phase space.

Proposition 3.16. *The Hamilton–d'Alembert principle in phase space for a curve $z(t) = (q(t), p(t))$, $t_1 \le t \le t_2$, in T^*Q is intrinsically represented, keeping the endpoints of $q(t)$ fixed, by*

$$\begin{aligned} \delta \int_{t_1}^{t_2} \{p(t) \cdot \dot q(t) - H(q(t), p(t))\}\, \mathrm{d}t &= \delta \int_{t_1}^{t_2} \left\{G(\rho_{TT^*Q}(z, \dot z)) - H(\tau_{T^*Q}(z, \dot z))\right\} \mathrm{d}t \\ &= \int_{t_1}^{t_2} \left\{\chi(z, \dot z) - (\tau_{T^*Q})^* \mathbf{d}H(z, \dot z)\right\} \cdot w \, \mathrm{d}t \\ &= 0 \end{aligned} \tag{3.39}$$

for a chosen variation $w = (q, p, \dot q, \dot p, \delta q, \delta p, \delta \dot q, \delta \dot p) \in \mathcal{J}(z, \dot z) \subset T_{(z,\dot z)}(TT^*Q)$.

Then, the intrinsic Hamilton–d'Alembert principle in phase space is equivalent to the equations

$$\chi(z, \dot z) \cdot w = (\tau_{T^*Q})^* \mathbf{d}H(z, \dot z) \cdot w \tag{3.40}$$

for all variations $w \in \mathcal{J}(z, \dot z)$.

Proof. From Eq. (3.31), it is apparent that Eq. (3.39) is the intrinsic expression of the Hamilton–d'Alembert principle in phase space. Thus, we obtain Eq. (3.40). □

Proposition 3.17. *Let $z(t)$, $t_1 \le t \le t_2$, be the integral curve of a vector field X on T^*Q. If $z(t)$ is a solution curve of the Hamilton–d'Alembert principle, then it satisfies*

$$\chi(X(z)) \cdot w = (\tau_{T^*Q})^* \mathbf{d}H(X(z)) \cdot w \tag{3.41}$$

for all $w \in \mathcal{J}(X(z))$.

Proof. Since $z(t)$ is the integral curve of X, we have $\dot z = X(z)$. By substituting this into Eq. (3.40), we can obtain Eq. (3.41). □

We shall call Eq. (3.41) the *intrinsic implicit Hamiltonian system*.

Proposition 3.18. *If a curve $z(t) = (q(t), p(t))$, $t_1 \le t \le t_2$, is an integral curve of the vector field X associated with Eq. (3.41), then, the curve $z(t) = (q(t), p(t))$ is a solution curve of an implicit Hamiltonian system (H, Δ_Q, X), which satisfies, for each $z(t) = (q(t), p(t))$,*

$$(X(z(t)), \mathbf{d}H(z(t))) \in D_{\Delta_Q}(z(t)),$$

where $D_{\Delta_Q}(z)$ is the induced Dirac structure on T^*Q.

Proof. It is logically obvious that the above proposition holds; however, we shall prove this by direct computations in coordinates.

Recall that the one-form χ is defined, by using the map $\Omega^\flat : TT^*Q \to T^*T^*Q$, such that $\chi = (\Omega^\flat)^*\Theta_{T^*T^*Q}$, and the left-hand side of Eq. (3.41) can be restated as, for each $z = (q, p) \in T^*Q$,

$$\begin{aligned}\chi(X(z)) \cdot w &= (\Omega^\flat)^*\Theta_{T^*T^*Q}(X(z)) \cdot w \\ &= \Theta_{T^*T^*Q}(\Omega^\flat(X(z))) \cdot T\Omega^\flat(w) \\ &= -\dot{p}\,\delta q + \dot{q}\,\delta p \end{aligned} \tag{3.42}$$

for all $w = (q, p, \dot{q}, \dot{p}, \delta q, \delta p, \delta\dot{q}, \delta\dot{p}) \in \mathcal{J}(X(z))$, where $X(z) = (q, p, \dot{q}, \dot{p})$, $\Omega^\flat(X(z)) = (q, p, -\dot{p}, \dot{q})$ and $T\Omega^\flat(w) = (q, p, -\dot{p}, \dot{q}, \delta q, \delta p, -\delta\dot{p}, \delta\dot{q})$. Recall also that the canonical one-form $\Theta_{T^*T^*Q}$ on T^*T^*Q is defined by

$$\Theta_{T^*T^*Q}(\alpha) \cdot V = \langle \alpha,\ T\pi_{T^*Q}(V)\rangle$$

for all $\alpha \in T^*T^*Q$ and $V \in T_\alpha(T^*T^*Q)$. So, it follows that

$$\begin{aligned}\Theta_{T^*T^*Q}(\Omega^\flat(X(z))) \cdot T\Omega^\flat(w) &= \Omega^\flat(X(z)) \cdot T\pi_{T^*Q}(T\Omega^\flat(w)) \\ &= -\dot{p}\,\delta q + \dot{q}\,\delta p, \end{aligned} \tag{3.43}$$

where one can easily check that

$$\begin{aligned}T\pi_{T^*Q}(T\Omega^\flat(w)) &= (T\pi_{T^*Q} \circ T\Omega^\flat)(q, p, \dot{q}, \dot{p}, \delta q, \delta p, \delta\dot{q}, \delta\dot{p}) \\ &= T\pi_{T^*Q}(q, p, -\dot{p}, \dot{q}, \delta q, \delta p, -\delta\dot{p}, \delta\dot{q}) \\ &= (q, p, \delta q, \delta p).\end{aligned}$$

From Eqs. (3.42) and (3.43), it reads that

$$\chi(X(z)) \cdot w = \Omega^\flat(X(z)) \cdot T\pi_{T^*Q}(T\Omega^\flat(w)). \tag{3.44}$$

On the other hand, the right-hand side of Eq. (3.41) is locally expressed by

$$\begin{aligned}(\tau_{T^*Q})^*\mathbf{d}H(X(z)) \cdot w &= \mathbf{d}H(\tau_{T^*Q}(X(z))) \cdot T_{X(z)}(\tau_{T^*Q})(w) \\ &= \frac{\partial H}{\partial q}\delta q + \frac{\partial H}{\partial p}\delta p. \end{aligned} \tag{3.45}$$

Noting the identity

$$T(\tau_{T^*Q}) = T\pi_{T^*Q} \circ T\Omega^\flat,$$

we can set

$$\begin{aligned}\delta z &= T_{X(z)}(\tau_{T^*Q})(w) \\ &= T\pi_{T^*Q} \circ T\Omega^\flat(w) \\ &= (q, p, \delta q, \delta p). \end{aligned} \tag{3.46}$$

From Eqs. (3.44) to (3.46), an integral curve $z(t) = (q(t), p(t)),\ t_1 \le t \le t_2$, of the vector field X on T^*Q satisfies

$$\Omega^\flat(X(z(t))) \cdot \delta z(t) = \mathbf{d}H(z(t)) \cdot \delta z(t)$$

for all $\delta z(t) = (\delta q(t), \delta p(t)) \in \Delta_{T^*Q}(z(t))$. This equation indicates the condition for an implicit Hamiltonian system (H, Δ_Q, X) associated with the induced Dirac structure D_{Δ_Q} on T^*Q, namely,

$$(X, \mathbf{d}H) \in D_{\Delta_Q}. \quad \square$$

We can summarize the results obtained so far in the following theorem:

Theorem 3.19. *Consider a hyperregular Hamiltonian H on T^*Q and with a given distribution Δ_Q on Q. Let X be a vector field on T^*Q such that (H, Δ_Q, X) is an implicit Hamiltonian system. Let $z(t) = (q(t), p(t)),\ t_1 \le t \le t_2$, be a curve in T^*Q. The following statements are equivalent:*

(a) *$z(t)$ is a solution curve of the implicit Hamiltonian system (H, Δ_Q, X);*
(b) *$z(t)$ satisfies the Hamilton–d'Alembert principle in phase space in Eq. (3.39);*
(c) *$z(t)$ is the integral curve of the vector field X on T^*Q.*

4. Nonholonomic systems with external forces

In this section, we demonstrate that nonholonomic mechanical systems with external forces can be naturally incorporated into the context of implicit Lagrangian systems by employing the Lagrange–d'Alembert–Pontryagin principle. Needless to say, nonholonomic mechanical systems have been widely investigated from the viewpoint of geometric mechanics in conjunction with the analysis of stability and control problems (see, for instance, [41,5]). In particular, what is developed in this section will be quite useful in the analysis of interconnected systems, and also relevant to controlled Lagrangian systems (see [6,7]).

4.1. External force fields

Consider a mechanical system with an external force field and let Q be a configuration manifold. Let $\pi_Q : T^*Q \to Q$ be the cotangent projection. Recall that an *external force field* $F : TQ \to T^*Q$ is a fiber-preserving map over the identity, which induces the horizontal one-form F' on T^*Q as

$$F'(z) \cdot \delta z = \langle F(q, v), T_z\pi_Q(\delta z)\rangle,$$

where $z \in T_q^*Q$, $v \in T_qQ$ and $\delta z \in T_zT^*Q$. Further, using the projection $\mathrm{pr}_{T^*Q} : TQ \oplus T^*Q \to T^*Q$, the horizontal one-form F' on T^*Q can be lifted as the horizontal one-form $\widetilde{F}$ on $TQ \oplus T^*Q$ such that, for $x \in TQ \oplus T^*Q$,

$$\begin{aligned}\widetilde{F}(x) \cdot \delta x &= F'(\mathrm{pr}_{T^*Q}(x)) \cdot T_x\mathrm{pr}_{T^*Q}(\delta x)\\ &= F'(z) \cdot \delta z,\end{aligned}$$

where $\delta x \in T_x(TQ \oplus T^*Q)$, $z = \mathrm{pr}_{T^*Q}(x)$ and $\delta z = T_x\mathrm{pr}_{T^*Q}(\delta x)$.

4.2. The Lagrange–d'Alembert–Pontryagin principle

Consider kinematic constraints that are given by a constraint distribution Δ_Q on Q, which is locally represented by $\Delta(q) \subset \mathbb{R}^n$ at each $q \in U \subset \mathbb{R}^n$. We assume that the dimension of $\Delta(q)$ is $n - m$ at each point q and let $\Delta^\circ(q)$ be the annihilator of $\Delta(q)$ spanned by m one-forms $\omega^1, \ldots, \omega^m$.

Recall that the motion of the mechanical system $c : [t_1, t_2] \to Q$ is said to be constrained if $\dot{c}(t) \in \Delta_Q(c(t))$ for all t, $t_1 \leq t \leq t_2$. Further, the distribution Δ_Q is not involutive in general; that is, $[X(q), Y(q)] \notin \Delta(q)$ for any two vector fields X, Y on Q with values in Δ_Q.

Let L be a (possibly degenerate) Lagrangian on TQ and let $F : TQ \to T^*Q$ be an external force field. The Lagrange–d'Alembert–Pontryagin principle for a curve $(q(t), v(t), p(t))$, $t_1 \leq t \leq t_2$, in $TQ \oplus T^*Q$ is represented by

$$\begin{aligned}&\delta \int_{t_1}^{t_2} \{L(q(t), v(t)) + p(t) \cdot (\dot{q}(t) - v(t))\}\, \mathrm{d}t + \int_{t_1}^{t_2} F(q(t), v(t)) \cdot \delta q(t)\, \mathrm{d}t\\ &= \delta \int_{t_1}^{t_2} \{p(t) \cdot \dot{q}(t) - E(q(t), v(t), p(t))\}\, \mathrm{d}t + \int_{t_1}^{t_2} F(q(t), v(t)) \cdot \delta q(t)\, \mathrm{d}t\\ &= 0 \end{aligned} \tag{4.1}$$

for a given variation $\delta q(t) \in \Delta(q(t))$ and with the constraint $v(t) \in \Delta(q(t))$. Keeping the endpoints of $q(t)$ fixed, we have

$$\begin{aligned}\delta \int_{t_1}^{t_2} \{L(q, v) + p \cdot (\dot{q} - v)\}\, \mathrm{d}t &= \delta \int_{t_1}^{t_2} \{p \cdot \dot{q} - E(q, v, p)\}\, \mathrm{d}t\\ &= \int_{t_1}^{t_2} \left\{\left(\frac{\partial L}{\partial q} - \dot{p}\right)\delta q + \left(\frac{\partial L}{\partial v} - p\right)\delta v + (\dot{q} - v)\delta p\right\} \mathrm{d}t.\end{aligned}$$

Hence, the Lagrange–d'Alembert–Pontryagin principle is represented by

$$\int_{t_1}^{t_2} \left\{\left(\frac{\partial L}{\partial q} - \dot{p}\right)\delta q + \left(\frac{\partial L}{\partial v} - p\right)\delta v + (\dot{q} - v)\delta p\right\} \mathrm{d}t + \int_{t_1}^{t_2} F(q, v)\, \delta q\, \mathrm{d}t = 0 \tag{4.2}$$

for a chosen variation $\delta q(t) \in \Delta(q(t))$, for all $\delta v(t)$ and $\delta p(t)$, and with $v(t) \in \Delta(q(t))$.

Proposition 4.1. *The Lagrange–d'Alembert–Pontryagin principle gives the local expressions of equations of motion for a nonholonomic mechanical system with an external force such that*

$$\dot{q} = v, \quad \dot{p} - \frac{\partial L}{\partial q} - F(q, v) \in \Delta^{\circ}(q), \quad p = \frac{\partial L}{\partial v}, \quad v \in \Delta(q). \tag{4.3}$$

Proof. From Eq. (4.2), it reads that

$$\left(\frac{\partial L}{\partial q} - \dot{p} + F(q, v)\right) \delta q + \left(\frac{\partial L}{\partial v} - p\right) \delta v + (\dot{q} - v)\, \delta p = 0,$$

which is satisfied for a given variation $\delta q(t) \in \Delta(q(t))$, for all $\delta v(t)$ and $\delta p(t)$, and with the constraint $v(t) \in \Delta(q(t))$. Thus, we obtain Eq. (4.3). □

4.3. Coordinate representation

Recall the one-forms $\omega^1, \ldots, \omega^m$ span a basis of the annihilator $\Delta^{\circ}(q)$ at each $q \in U \subset \mathbb{R}^n$, and it follows that Eq. (4.3) can be represented, in coordinates, by employing the Lagrange multipliers $\mu_a,\ a = 1, \ldots, m$, such that

$$\begin{aligned}
&\begin{pmatrix} \dot{q}^i \\ \dot{p}_i \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} -\dfrac{\partial L}{\partial q^i} - F_i(q, v) \\ v^i \end{pmatrix} + \begin{pmatrix} 0 \\ \mu_a\, \omega_i^a(q) \end{pmatrix}, \\
&p_i = \frac{\partial L}{\partial v^i}, \\
&0 = \omega_i^a(q)\, v^i,
\end{aligned}$$

where we employ the local expression $\omega^a = \omega_i^a\, \mathrm{d}q^i$.

4.4. Intrinsic formulation

Denote by $\Delta_Q \subset TQ$ a constraint distribution. Let $\mathcal{K} \subset TQ \oplus T^*Q$ be the submanifold defined in Eq. (3.14) and $\mathcal{G} \subset TT(TQ \oplus T^*Q)$ be the regular distribution defined by Eq. (3.15). Let $F : TQ \to T^*Q$ be an external force field and $\widetilde{F}$ on $TQ \oplus T^*Q$ be the horizontal one-form. The Lagrange–d'Alembert–Pontryagin principle for a curve $x(t) = (q(t), v(t), p(t))$ in $TQ \oplus T^*Q$ is intrinsically represented by

$$\begin{aligned}
&\delta \int_{t_1}^{t_2} \{L(q(t), v(t)) + p(t) \cdot (\dot{q}(t) - v(t))\}\ \mathrm{d}t + \int_{t_1}^{t_2} F(q(t), v(t)) \cdot \delta q(t)\, \mathrm{d}t \\
&\quad = \delta \int_{t_1}^{t_2} \{p(t) \cdot \dot{q}(t) - E(q(t), v(t), p(t))\}\ \mathrm{d}t + \int_{t_1}^{t_2} F(q(t), v(t)) \cdot \delta q(t)\, \mathrm{d}t \\
&\quad = \delta \int_{t_1}^{t_2} \{G(\rho_{TT^*Q} \circ T\mathrm{pr}_{T^*Q}(x(t), \dot{x}(t))) - E(\tau_{TQ\oplus T^*Q}(x(t), \dot{x}(t)))\}\, \mathrm{d}t \\
&\qquad + \int_{t_1}^{t_2} \widetilde{F}(\tau_{TQ\oplus T^*Q}(x(t), \dot{x}(t))) \cdot T\tau_{TQ\oplus T^*Q}(w(t))\, \mathrm{d}t \\
&\quad = \int_{t_1}^{t_2} \{(T\mathrm{pr}_{T^*Q})^* \chi(x(t), \dot{x}(t)) - (\tau_{TQ\oplus T^*Q})^* \mathbf{d}E(x(t), \dot{x}(t))\} \cdot w(t)\, \mathrm{d}t \\
&\qquad + \int_{t_1}^{t_2} \left(\tau_{TQ\oplus T^*Q}\right)^* \widetilde{F}(x(t), \dot{x}(t)) \cdot w(t)\, \mathrm{d}t \\
&\quad = 0,
\end{aligned} \tag{4.4}$$

which is satisfied for all variations $w = (q, v, p, \dot{q}, \dot{v}, \dot{p}, \delta q, \delta v, \delta p, \delta\dot{q}, \delta\dot{v}, \delta\dot{p}) \in \mathcal{G}(x, \dot{x}) \subset T_{(x,\dot{x})}T(TQ \oplus T^*Q)$ and with the endpoints of $q(t)$ fixed. Using the projections $\mathrm{pr}_Q : TQ \oplus T^*Q \to Q$ and $\tau_{TQ\oplus T^*Q} : T(TQ \oplus T^*Q) \to TQ \oplus T^*Q$ with their tangent maps, it follows that $\left(T\mathrm{pr}_Q \circ T\tau_{TQ\oplus T^*Q}\right)(w) = (q, \delta q) \in \Delta_Q$.

Proposition 4.2. *The Lagrange–d'Alembert–Pontryagin principle in Eq. (4.4) for a curve $x(t)$ in $TQ \oplus T^*Q$ induces the equation*

$$(T\mathrm{pr}_{T^*Q})^* \chi(x(t), \dot{x}(t)) \cdot w(t) = \left(\tau_{TQ\oplus T^*Q}\right)^* \left(\mathbf{d}E(x(t), \dot{x}(t)) - \widetilde{F}(x(t), \dot{x}(t))\right) \cdot w(t) \tag{4.5}$$

for all $w(t) \in \mathcal{G}(x(t), \dot{x}(t))$.

Proof. Recall that the variation of the Hamilton–Pontryagin integral is given in Eq. (3.4) and recall also that the external force field $F : TQ \to T^*Q$ induces the horizontal one-form $\widetilde{F}$ on $TQ \oplus T^*Q$, and it is apparent that Eq. (4.4) is the intrinsic expression for Eq. (4.1). Thus, we obtain Eq. (4.5). □

Proposition 4.3. *Let X be a vector field on T^*Q, defined at points of $P = \mathbb{F}L(\Delta_Q)$. Let $\widetilde{X}$ be the naturally induced vector field on $TQ \oplus T^*Q$, defined at points of $\mathcal{K}$, as shown in Eq. (3.18). Denote by $x(t)$, $t_1 \le t \le t_2$, an integral curve of $\widetilde{X}$.*

If $x(t)$ is a solution curve of the Lagrange–d'Alembert–Pontryagin principle in Eq. (4.4), then it satisfies

$$(T\mathrm{pr}_{T^*Q})^* \chi(\widetilde{X}(x(t))) \cdot w(t) = (\tau_{TQ\oplus T^*Q})^* \left(\mathbf{d}E(\widetilde{X}(x(t))) - \widetilde{F}(\widetilde{X}(x(t)))\right) \cdot w(t) \tag{4.6}$$

for all $w(t) \in \mathcal{G}(\widetilde{X}(x(t)))$.

Proof. If the curve $x(t)$ in $TQ \oplus T^*Q$ is an integral curve of the induced vector field $\widetilde{X}$, then $\dot{x}(t) = \widetilde{X}(x(t))$. Thus, we obtain Eq. (4.6) by substituting $\dot{x}(t) = \widetilde{X}(x(t))$ into Eq. (4.5). □

Proposition 4.4. *If a curve $x(t) = (q(t), v(t), p(t))$ is an integral curve of the induced vector field $\widetilde{X}$ associated with Eq. (4.6), then, the curve $x(t) = (q(t), v(t), p(t))$ is a solution curve of the implicit Lagrangian system (L, F, Δ_Q, X), which satisfies the condition, for each $(q(t), v(t)) \in \Delta_Q(q(t))$,*

$$(X(q(t), p(t)), \mathfrak{D}L(q(t), v(t)) - \pi_Q^* F(q(t), v(t))) \in D_{\Delta_Q}(q(t), p(t)), \tag{4.7}$$

where $(q(t), p(t)) = \mathbb{F}L(q(t), v(t))$ is an integral curve of X and $\mathfrak{D}L : TQ \to T^*T^*Q$ is the Dirac differential of L, $\pi_Q : T^*Q \to Q$ and D_{Δ_Q} is the induced Dirac structure defined in Eq. (3.20).

Proof. In view of Proposition 3.7, it suffices to check the terms of the external force field in Eqs. (4.6) and (4.7). Recall that the horizontal one-form $\widetilde{F}$ on $TQ \oplus T^*Q$ is defined by lifting the horizontal one-form F' on T^*Q such that

$$\begin{aligned}
(\tau_{TQ\oplus T^*Q})^* \widetilde{F}\left(\widetilde{X}(x)\right) \cdot w &= \widetilde{F}(\tau_{TQ\oplus T^*Q})\left(\widetilde{X}(x)\right) \cdot T_{\widetilde{X}(x)} \tau_{TQ\oplus T^*Q}(w) \\
&= F'\left(\mathrm{pr}_{T^*Q}(x)\right) \cdot T_x \mathrm{pr}_{T^*Q}\left(T_{\widetilde{X}(x)} \tau_{TQ\oplus T^*Q}(w)\right) \\
&= F'(z) \cdot \delta z
\end{aligned}$$

for all $w(t) \in \mathcal{G}(\widetilde{X}(x(t)))$. In the above, notice that $x = \tau_{TQ\oplus T^*Q}(\widetilde{X}(x))$, $z = \mathrm{pr}_{T^*Q}(x)$ and $\delta z = T_x \mathrm{pr}_{T^*Q}\left(T_{\widetilde{X}(x)} \tau_{TQ\oplus T^*Q}(w)\right)$. Furthermore, since the horizontal one-form F' is induced from the external force field $F : TQ \to T^*Q$, it follows that

$$\begin{aligned}
F'(z) \cdot \delta z &= \left\langle F(q, v), T_z \pi_Q(\delta z) \right\rangle \\
&= \pi_Q^* F(q, v) \cdot \delta z.
\end{aligned}$$

In combination with the proof of Proposition 3.7, it follows that a curve $(q(t), v(t))$, $t_1 \le t \le t_2$, in Δ_Q satisfies

$$\Omega_{\Delta_Q}(X(z(t)), \delta z(t)) = \left(\mathfrak{D}L(q(t), v(t)) - \pi_Q^* F(q(t), v(t))\right) \cdot \delta z(t),$$

where $z(t) = (q(t), p(t)) = \mathbb{F}L(q(t), v(t))$ and Ω_{Δ_Q} is the restriction of the canonical symplectic form Ω to Δ_{T^*Q}. This can be restated by the condition of an implicit Lagrangian system (L, F, Δ_Q, X); that is, for each $v(t) \in \Delta_Q(q(t))$,

$$(X(z(t)), \mathfrak{D}L(q(t), v(t)) - \pi_Q^* F(q(t), v(t))) \in D_{\Delta_Q}(z(t)),$$

where $z(t) = \mathbb{F}L(q(t), v(t))$ is an integral curve of X. Thus, we can check that the integral curve $x(t) = (q(t), v(t), p(t))$ of $\widetilde{X}$ associated with Eq. (4.6) is a solution curve of (L, F, Δ_Q, X). □

We can summarize the results obtained so far in the following theorem.

Theorem 4.5. *Consider a Lagrangian L on TQ (possibly degenerate) and with a given distribution Δ_Q on Q. Let $F : TQ \to T^*Q$ be an external force field. Let X be a vector field on T^*Q, defined at points of $P = \mathbb{F}L(\Delta_Q)$ such that the quadruple (L, F, Δ_Q, X) is an implicit Lagrangian system. Denote by $x(t) = (q(t), v(t), p(t))$, $t_1 \le t \le t_2$, a curve in $TQ \oplus T^*Q$. The following statements are equivalent:*

(a) *$x(t)$ is a solution curve of the implicit Lagrangian system (L, F, Δ_Q, X);*
(b) *$x(t)$ satisfies the Lagrange–d'Alembert–Pontryagin principle in Eq. (4.5);*
(c) *$x(t)$ is the integral curve of the vector field $\widetilde{X}$ on $TQ \oplus T^*Q$ naturally induced from X.*

5. Implicit constrained Lagrangian systems

In this section, we investigate a *constrained Dirac structure* D_P on the constraint momentum space $P = \mathbb{F}L(\Delta_Q) \subset T^*Q$ by using an Ehresmann connection and we also develop an *implicit constrained Lagrangian system* associated with D_P.

5.1. Ehresmann connections

We briefly review an Ehresmann connection associated with nonholonomic mechanical systems; for details, refer to Koon and Marsden [20] and Bloch [5].

Assume that there is a bundle structure with a projection $\pi : Q \to \mathcal{R}$ for our space Q; that is, there exists another manifold $\mathcal{R}$ called the base. We call the kernel of $T_q\pi$ at any point $q \in Q$ the *vertical space* denoted by $\mathcal{V}_q$.

Recall that an *Ehresmann connection* A is a vertical vector-valued one-form on Q, which satisfies

1. $A_q : T_qQ \to \mathcal{V}_q$ is a linear map at each point $p \in Q$,
2. A is a projection : $A(v_q) = v_q$, for all $v_q \in \mathcal{V}_q$.

Thus, we can split the tangent space at q such that $T_qQ = \mathcal{H}_q \oplus \mathcal{V}_q$, where $\mathcal{H}_q = \mathrm{Ker}\, A_q$ is the horizontal space at q.

Let $\Delta_Q \subset TQ$ be a constraint distribution, which is locally given by

$$\Delta_Q(q) = \{v_q \in T_qQ \mid \langle \omega^a, v_q \rangle = 0,\ a = 1, \ldots, m\},$$

where ω^a are m independent one-forms that form the basis for the annihilator $\Delta_Q^\circ \subset T^*Q$. Let us choose an Ehresmann connection A in such a way that $\mathcal{H}_q = \Delta_Q(q)$. In other words, we assume that the connection is chosen such that the constraints are written as $A \cdot v_q = 0$.

Using the bundle coordinates $q = (r, s) \in \mathbb{R}^{n-m} \times \mathbb{R}^m$, the coordinate representation of π is just projection onto the factor r, and the connection A can be locally expressed by a vector-valued differential form ω^a as

$$A = \omega^a \frac{\partial}{\partial s^a}, \quad \omega^a(q) = \mathrm{d}s^a + A^a_\alpha(r, s)\, \mathrm{d}r^\alpha, \quad a = 1, \ldots, m;\ \alpha = 1, \ldots, n - m.$$

Let

$$v_q = u^\alpha \frac{\partial}{\partial r^\alpha} + w^a \frac{\partial}{\partial s^a}$$

be an element of T_qQ. Then,

$$\omega^a(v_q) = w^a + A^a_\alpha u^\alpha$$

and

$$A(v_q) = \left(w^a + A^a_\alpha u^\alpha\right) \frac{\partial}{\partial s^a}.$$

5.2. Horizontal lift

Given an Ehresmann connection A, a point $q \in Q$ and a vector $v_r \in T_r\mathcal{R}$ tangent to the base at a point $r = \pi(q) \in \mathcal{R}$, we can define the *horizontal lift* of v_r to be the unique vector v_r^h in $\mathcal{H}_q$ that projects to v_r under $T_q\pi$. If we have a vector $X_q \in T_q Q$, we shall write its vertical part as

$$\operatorname{ver} X_q = A(q) \cdot X_q,$$

and we shall also write its horizontal part as

$$\operatorname{hor} X_q = X_q - A(q) \cdot X_q.$$

In coordinates, the vertical projection is the map $(u^\alpha, w^a) \mapsto (0, w^a + A^a_\alpha u^\alpha)$, while the horizontal projection is the map $(u^\alpha, w^a) \mapsto (u^\alpha, -A^a_\alpha u^\alpha)$.

5.3. The Lagrange–d'Alembert–Pontryagin principle

Let L be a (possibly degenerate) Lagrangian on TQ. Define a generalized energy E by $E(q, v, p) = p \cdot v - L(q, v)$ using local coordinates (q, v, p) for $TQ \oplus T^*Q$. Recall that the Lagrange–d'Alembert–Pontryagin principle is given by

$$\begin{aligned}\delta \int_{t_1}^{t_2} \{L(q, v) + p \cdot (\dot q - v)\}\, \mathrm{d}t &= \delta \int_{t_1}^{t_2} \{p \cdot \dot q - E(q, v, p)\}\, \mathrm{d}t \\ &= \int_{t_1}^{t_2} \left\{ \left(\frac{\partial L}{\partial q^i} - \dot p_i \right) \delta q^i + \left(\frac{\partial L}{\partial v^i} - p_i \right) \delta v^i + (\dot q^i - v^i)\delta p_i \right\} \mathrm{d}t \\ &= 0\end{aligned}$$

for chosen variations $\delta q^i(t) \in \Delta_Q(q(t))$, with the endpoints of $q(t)$ fixed and with the constraint $\omega^a(q) \cdot v_q = 0$. Hence, the Lagrange–d'Alembert–Pontryagin principle is equivalent to the equation

$$\left(\frac{\partial L}{\partial q^i} - \dot p_i \right) \delta q^i + \left(\frac{\partial L}{\partial v^i} - p \right) \delta v^i + (\dot q^i - v^i)\delta p_i = 0,$$

for all $\delta q^i(t) \in \Delta_Q(q(t))$ that satisfy, in coordinates $q^i = (r^\alpha, s^a)$,

$$\delta s^a + A^a_\alpha\, \delta r^\alpha = 0,$$

where the distribution Δ_Q is denoted, in coordinates, by

$$\Delta_Q = \operatorname{span} \left\{ \frac{\partial}{\partial r^\alpha} - A^a_\alpha \frac{\partial}{\partial s^a} \right\}.$$

Since the kinematic constraints are given by $\omega^a(q) \cdot v_q = w^a + A^a_\alpha u^\alpha = 0$, the Lagrange–d'Alembert–Pontryagin principle may be restated as

$$\begin{aligned}&\left\{ -A^a_\alpha \left(\frac{\partial L}{\partial s^a} - \dot p_a \right) + \left(\frac{\partial L}{\partial r^\alpha} - \dot p_\alpha \right) \right\} \delta r^\alpha + \left(\frac{\partial L}{\partial u^\alpha} - p_\alpha \right) \delta u^\alpha \\ &\quad + \left(\frac{\partial L}{\partial w^a} - p_a \right) \delta w^a + (\dot r^\alpha - u^\alpha)\, \delta p_\alpha + (\dot s^a - w^a)\, \delta p_a = 0\end{aligned}$$

for all δr^α and for all $\delta v^i = (\delta u^\alpha, \delta w^a)$ and $\delta p_i = (\delta p_\alpha, \delta p_a)$. Then, the equations of motion are given, in coordinates, by

$$\begin{aligned}&\dot r^\alpha = u^\alpha, \\ &\dot s^a = w^a, \\ &\dot p_\alpha - \frac{\partial L}{\partial r^\alpha} = A^a_\alpha \left(\dot p_a - \frac{\partial L}{\partial s^a} \right),\end{aligned} \tag{5.1}$$

and with the Legendre transformation

$$p_\alpha = \frac{\partial L}{\partial u^\alpha}, \quad p_a = \frac{\partial L}{\partial w^a}. \tag{5.2}$$

Note that the equations of motion in Eq. (5.1) and the Legendre transformation in Eq. (5.2) are to be combined with the kinematic constraints

$$w^a = -A^a_\alpha u^\alpha. \tag{5.3}$$

5.4. The constrained Lagrangian

Define the constrained Lagrangian L_c on $\Delta_Q \subset TQ$ by $L_c(q, v) = L(q, \mathrm{hor}\, v)$ for $(q, v) \in TQ$. By substituting Eq. (5.3) into the Lagrangian L on TQ, we can obtain, in local coordinates,

$$L_c(r^\alpha, s^a, u^\alpha) = L(r^\alpha, s^a, u^\alpha, -A^a_\alpha(r, s)u^\alpha)$$

and define

$$\begin{aligned} \widetilde{p}_\alpha &= \frac{\partial L_c}{\partial u^\alpha} \\ &= \frac{\partial L}{\partial u^\alpha} - A^a_\alpha \frac{\partial L}{\partial w^a}, \end{aligned} \tag{5.4}$$

which is equivalent to

$$\widetilde{p}_\alpha = p_\alpha - A^a_\alpha p_a$$

in view of Eq. (5.2). By computations, it follows that

$$\begin{aligned} \frac{\partial L_c}{\partial r^\alpha} &= \frac{\partial L}{\partial r^\alpha} - \frac{\partial L}{\partial w^a}\left(\frac{\partial A^a_\beta}{\partial r^\alpha} u^\beta\right) \\ &= \frac{\partial L}{\partial r^\alpha} - p_a\left(\frac{\partial A^a_\beta}{\partial r^\alpha} u^\beta\right), \\ \frac{\partial L_c}{\partial s^a} &= \frac{\partial L}{\partial s^a} - \frac{\partial L}{\partial w^a}\left(\frac{\partial A^a_\beta}{\partial s^a} u^\beta\right) \\ &= \frac{\partial L}{\partial s^a} - p_a\left(\frac{\partial A^a_\beta}{\partial s^a} u^\beta\right). \end{aligned} \tag{5.5}$$

Substituting Eqs. (5.4) and (5.5) into Eq. (5.1) together with the kinematic constraints in Eq. (5.3), we can eventually obtain the *implicit Lagrange–d'Alembert equations of motion* for the constrained Lagrangian L_c as

$$\begin{aligned} \dot{r}^\alpha &= u^\alpha, \\ \dot{s}^a &= -A^a_\beta u^\beta, \\ \dot{\widetilde{p}}_\alpha &= \frac{\partial L_c}{\partial r^\alpha} - A^b_\alpha \frac{\partial L_c}{\partial s^b} - p_b K^b_{\alpha\beta} u^\beta, \\ \widetilde{p}_\alpha &= \frac{\partial L_c}{\partial u^\alpha}, \end{aligned} \tag{5.6}$$

where

$$K^b_{\alpha\beta} = \frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} + A^a_\alpha \frac{\partial A^b_\beta}{\partial s^a} - A^a_\beta \frac{\partial A^b_\alpha}{\partial s^a}$$

is the *curvature* of the Ehresmann connection A.

Remarks. The curvature of the Ehresmann connection A is the vertical valued two-form on Q defined by its action on two vector fields Y and Z on Q such that

$$K(Y, Z) = -A([\mathrm{hor}\, Y, \mathrm{hor}\, Z]),$$

where the bracket on the right-hand side is the Jacobi–Lie bracket of vector fields.

Recall that the identity for the exterior derivative $\mathbf{d}\alpha$ of a one-form α on a manifold Q acting on two vector fields $Y,\ Z$ is

$$(\mathbf{d}\alpha)[Y, Z] = Y[\alpha(Z)] - Z[\alpha(Y)] - \alpha([Y, Z]).$$

This identity indicates that one can evaluate, in coordinates, the curvature by writing the connection as a one-form ω^b by computing its exterior derivative (component by component) and restricting the result to horizontal vectors, that is, to the constraint distribution. Then, one has

$$K(Y, Z) = \mathrm{d}\omega^b(\mathrm{hor}\, Y, \mathrm{hor}\, Z)\frac{\partial}{\partial s^b},$$

where the local expression for the curvature is denoted by

$$K^b(Y, Z) = K^b_{\alpha\beta}\, Y^\alpha Z^\beta.$$

Proposition 5.1. *Let $\mathbf{d}\omega^b$ be the exterior derivative of ω^b. The Lagrange–d'Alembert–Pontryagin principle for the constrained Lagrangian L_c may be written as*

$$\delta\{L_c(r^\alpha, s^a, u^\alpha) + \widetilde{p}_\alpha\,(\dot{r}^\alpha - u^\alpha)\} = p_b\,\mathbf{d}\omega^b(v, \delta r), \tag{5.7}$$

which provides the equations of motion in Eq. (5.6).

Proof. The left-hand side of Eq. (5.7) is

$$\delta\{L_c(r^\alpha, s^a, u^\alpha) + \widetilde{p}_\alpha(\dot{r}^\alpha - u^\alpha)\} = \left\{\frac{\partial L_c}{\partial r^\alpha} - A^a_\alpha\,\frac{\partial L_c}{\partial s^a} - \dot{\widetilde{p}}_\alpha\right\}\delta r^\alpha + \left(\frac{\partial L_c}{\partial u^\alpha} - \widetilde{p}_\alpha\right)\delta u^\alpha + (\dot{r}^\alpha - u^\alpha)\delta\widetilde{p}_\alpha,$$

while the direct computation using properties of differential forms shows that

$$\mathbf{d}\omega^b(v, \cdot) = K^b_{\alpha\beta}\, u^\beta\, \mathrm{d}r^\alpha.$$

It follows that

$$\left\{\frac{\partial L_c}{\partial r^\alpha} - A^a_\alpha\,\frac{\partial L_c}{\partial s^a} - \dot{\widetilde{p}}_\alpha\right\}\delta r^\alpha + \left(\frac{\partial L_c}{\partial u^\alpha} - \widetilde{p}_\alpha\right)\delta u^\alpha + (\dot{r}^\alpha - u^\alpha)\delta\widetilde{p}_\alpha = \left(p_b\, K^b_{\alpha\beta}\, u^\beta\right)\delta r^\alpha$$

for all δr^α, δu^α, and $\delta\widetilde{p}_\alpha$. Combining with the kinematic constraints, we obtain Eq. (5.6). □

Remarks. The constrained energy E_c may be defined by

$$E_c(r^\alpha, s^a, u^\alpha, \widetilde{p}_\alpha) = \widetilde{p}_\alpha\, u^\alpha - L_c(r^\alpha, s^a, u^\alpha)$$

and then, the Lagrange–d'Alembert–Pontryagin principle in Eq. (5.7) can be restated as the equivalent form

$$\delta\{\widetilde{p}_\alpha\,\dot{r}^\alpha - E_c(r^\alpha, s^a, u^\alpha, \widetilde{p}_\alpha)\} = p_b\,\mathbf{d}\omega^b(v, \delta r).$$

5.5. Restriction of a Dirac structure

Before going into the construction of a constrained Dirac structure, we briefly discuss the restriction of a Dirac structure on a manifold to its submanifold.

Let M be a manifold and let N be a submanifold of M. Recall the definition of a Dirac structure on a manifold M; that is, a subbundle $D_M \subset TM \oplus T^*M$ is called a Dirac structure if for every fiber $D_M(x) \subset T_xM \times T^*_xM,\ x \in M$, one has $D_M(x) = D^\perp_M(x)$, where

$$D^\perp_M(x) = \{(v_x, \alpha_x) \in T_xM \times T^*_xM \mid \langle\bar{\alpha}_x, v_x\rangle + \langle\alpha_x, \bar{v}_x\rangle = 0,\ \text{for all } (\bar{v}_x, \bar{\alpha}_x) \in D_M(x)\}.$$

Now, let D_M be a Dirac structure on M. Define a map $\sigma : TN \times T^*M \to TN \times T^*N$ as follows: For each $y \in N$, let $\sigma(y) : T_yN \times T_y^*M \to T_yN \times T_y^*N$ be defined by

$$\sigma(y)(v_y, \alpha_y) = (v_y, \alpha_y|T_yN),$$

where $v_y \in T_yN$, $\alpha_y \in T_y^*M$, and $\alpha_y|T_yN$ denotes the restriction of the covector α_y to the subspace T_yN. This map σ is a vector bundle projection.

Assume that $D_M(y) \cap (T_yN \times T_y^*M)$ has constant dimension for each $y \in N$, namely, it is a vector subbundle of $TN \times T^*M$. Define the subbundle $D_N \subset TN \oplus T^*N$ by, for each $y \in N$,

$$D_N(y) = \sigma(y)\left(D_M(y) \cap (T_yN \times T_y^*M)\right).$$

Proposition 5.2. *The subbundle $D_N \subset TN \oplus T^*N$ is a Dirac structure on N.*

Proof. Let us check $D_N(y) = D_N^\perp(y)$ for each $y \in N$. It is obvious that $D_N(y) \subset D_N^\perp(y)$ for $y \in N$. Then, let us check $D_N^\perp(y) \subset D_N(y)$ for $y \in N$. Suppose that $(w_y, \beta_y) \in D_N^\perp(y) \subset T_yN \times T_y^*N$. So, we have $\langle\alpha_y, w_y\rangle + \langle\beta_y, v_y\rangle = 0$ for all $(v_y, \alpha_y) \in D_N(y)$. Then, there exists $\alpha'_y \in T_y^*M$ such that $(v_y, \alpha_y) = \sigma(y)(v_y, \alpha'_y) \in D_N(y)$, where $(v_y, \alpha'_y) \in D_M(y)$ and $\alpha'_y|T_yN = \alpha_y$. Therefore, we have $\langle\alpha_y, w_y\rangle + \langle\beta_y, v_y\rangle = \langle\alpha'_y, w_y\rangle + \langle\beta'_y, v_y\rangle = 0$ for all $(v_y, \alpha'_y) \in D_M(y)$, where β'_y is an arbitrary extension of β_y to T_yM and $v_y \in T_yN$. Hence, one obtains

$$\begin{aligned}(w_y, \beta'_y) \in \left(D_M(y) \cap (T_yN \times T_y^*M)\right)^\perp &= D_M^\perp(y) + (T_yN \times T_y^*M)^\perp \\ &= D_M(y) + (\{0\} \times T_yN^\circ).\end{aligned}$$

Then, there exists $\gamma'_y \in T_yN^\circ \subset T_y^*M$ such that $(w_y, \beta'_y + \gamma'_y) \in D_M(y)$. Noting $\sigma(y)(w_y, \beta'_y + \gamma'_y) = (w_y, (\beta'_y + \gamma'_y)|T_yN) = (w_y, \beta_y) \in D_N(y)$, it follows that $D_N^\perp(y) \subset D_N(y)$ for $y \in N$. □

We call the Dirac structure D_N the *restriction of a Dirac structure D_M* to N.

Remarks. Proposition 5.2 was originally developed by Courant [13]; we follow the exposition in [4].

Proposition 5.3. *Let D_N be constructed as in Proposition 5.2 and let $\iota : N \to M$ denote the inclusion map. Then, (w, β) is a local section of D_N if and only if there exists a local section (v, α) of D_M such that $T\iota \circ w = v \circ \iota$ and $\iota^*\alpha = \beta$. In other words, a Dirac structure D_N on N is represented by, for each $y \in N$,*

$$\begin{aligned}D_N(y) = \{&(w_y, \beta_y) \in T_yN \times T_y^*N |\ \textit{there is a}\ (v_{\iota(y)}, \alpha_{\iota(y)}) \in D_M(\iota(y)) \\ &\textit{such that}\ T\iota \circ w = v \circ \iota\ \textit{and}\ \iota^*\alpha = \beta\}.\end{aligned}$$

Proof. As demonstrated in Part I, given a distribution Δ_M on M and a two-form Ω on M, there exists a Dirac structure $D_M \subset TM \oplus T^*M$ on a manifold M, whose fiber is defined by, for each $x \in M$,

$$D_M(x) = \{(v_x, \alpha_x) \in T_xM \times T_x^*M \mid v_x \in \Delta_M(x) \text{ and } \alpha_x(v'_x) = \Omega_{\Delta_M}(x)(v_x, v'_x) \text{ for all } v'_x \in \Delta_M(x)\},$$

where $\Omega_{\Delta_M} = \Omega|_{\Delta_M \times \Delta_M}$.

Define a distribution Δ_N on N by restricting Δ_M to N such that

$$\Delta_N = TN \cap \Delta_M,$$

where we assume Δ_N to be a regular distribution. By construction, we can define the restricted Dirac structure D_N on N such that, for each $y \in N$,

$$\begin{aligned}D_N(y) &= \sigma(y)\left(D_M(\iota(y)) \cap (T_yN \times T_{\iota(y)}^*M)\right) \\ &= \{(w_y, \beta_y) \in T_yN \times T_y^*N \mid w_y \in \Delta_N(y) \text{ and } \beta_y(w'_y) = \Omega_{\Delta_N}(y)(w_y, w'_y) \text{ for all } w'_y \in \Delta_N(y)\},\end{aligned}$$

where $\beta = \iota^*\alpha$, $\Omega_{\Delta_N} = \iota^*\,\Omega_{\Delta_M}$, $T\iota \cdot w = v \circ \iota$ and $T\iota \cdot w' = v' \circ \iota$. □

5.6. Constrained Dirac structure

Let L be a Lagrangian on TQ and $\Delta_Q \subset TQ$ a constraint distribution. Let $\pi_Q : T^*Q \to Q$ be the cotangent projection. Recall that an induced Dirac structure D_{Δ_Q} on T^*Q is defined such that, for each $z \in T^*Q$,

$$\begin{aligned} D_{\Delta_Q}(z) = \{(v_z, \alpha_z) \in T_z T^*Q \times T_z^* T^*Q \mid v_z \in \Delta_{T^*Q}(z) \text{ and } \\ \alpha_z(v_z') = \Omega_{\Delta_Q}(z)(v_z, v_z') \text{ for all } v_z' \in \Delta_{T^*Q}(z)\}, \end{aligned}$$

where $\Delta_{T^*Q} = (T\pi_Q)^{-1}(\Delta_Q)$ and Ω_{Δ_Q} is the restriction of the canonical symplectic structure Ω on T^*Q to Δ_{T^*Q}.

Our initial goal is to define a Dirac structure on the constraint momentum space $P = \mathbb{F}L(\Delta_Q) \subset T^*Q$. To achieve this, we define a regular distribution on P by restricting Δ_{T^*Q} to P such that

$$\Delta_P = TP \cap \Delta_{T^*Q}.$$

Letting $\iota : P \to T^*Q$ be the inclusion, we can define a *constrained Dirac structure* on $P = \mathbb{F}L(\Delta_Q) \subset T^*Q$ by the subbundle $D_P \subset TP \oplus T^*P$, whose fiber is given such that, for each $y \in P$,

$$D_P(y) = \{(w_y, \beta_y) \in T_yP \times T_y^*P \mid w_y \in \Delta_P(y) \text{ and } \beta_y(w_y') = \Omega_{\Delta_P}(y)(w_y, w_y') \text{ for all } w_y' \in \Delta_P(y)\},$$

where $\beta = \iota^*\alpha$, $\Omega_{\Delta_P} = \iota^*\,\Omega_{\Delta_Q}$, $T\iota \cdot w = v \circ \iota$ and $T\iota \cdot w' = v' \circ \iota$. It goes without saying that D_P is the restriction of the induced Dirac structure D_{Δ_Q}.

It is obvious that the constrained Dirac structure D_P may be restated as follows: for each $y \in P$,

$$D_P(y) = \{(v_y, \alpha_y) \in T_yP \times T_y^*P \mid v_y \in \Delta_P(y) \text{ and } \alpha_y - \Omega_P^\flat(y) \cdot v_y \in \Delta_P^\circ(y)\},$$

where Δ_P° is the annihilator of Δ_P and $\Omega_P^\flat : TP \to T^*P$ is the bundle map associated with the skew-symmetric bilinear form $\Omega_P = \Omega|_{TP\times TP}$.

We can also construct the constrained Dirac structure by employing the canonical Poisson structure. Recall that the canonical Poisson tensor $B : T^*T^*Q \times T^*T^*Q \to \mathbb{R}$ is defined by, for any smooth function F, G on T^*Q,

$$\begin{aligned} B(\mathbf{d}F, \mathbf{d}G) &= \Omega(X_F, X_G) \\ &= \left\langle \mathbf{d}F,\ B^\sharp \mathbf{d}G \right\rangle \\ &= \{F, G\}, \end{aligned}$$

where X_F and X_G are vector fields on T^*Q, $B^\sharp : T^*T^*Q \to TT^*Q$ is the associated bundle map, and $\{,\}$ is the Poisson bracket. Further, $B_P^\sharp : T^*P \to TP$ is the associated bundle map of the contravariant antisymmetric two-tensor $B_P = B|_{T^*P\times T^*P}$. Let $\Delta_P^* = \Omega_P^\flat(\Delta_P) \subset T^*P$ be the codistribution and $(\Delta_P^*)^\circ$ be its annihilator, and then the constrained Dirac structure on P is defined by, for each $y \in P$,

$$D_P(y) = \{(v_y, \alpha_y) \in T_yP \times T_y^*P \mid \alpha_y \in \Delta_P^*(y) \text{ and } v_y - B_P^\sharp(y) \cdot \alpha_y \in (\Delta_P^*)^\circ(y)\}. \tag{5.8}$$

5.7. A local representation using an Ehresmann connection

Let us construct a local representation of a constrained Dirac structure using an Ehresmann connection to represent the set D_P given in Eq. (5.8). Needless to say, T^*Q is naturally equipped with the canonical Poisson bracket $\{,\}$ or the Poisson structure $B : T^*T^*Q \times T^*T^*Q \to \mathbb{R}$ such that, for each $(q, p) \in T^*Q$,

$$\begin{aligned} \{F, G\}(q, p) &= B(q, p)(\mathbf{d}F(q, p), \mathbf{d}G(q, p)) \\ &= \left\langle \mathbf{d}F(q, p), B^\sharp(q, p)\, \mathbf{d}G(q, p) \right\rangle. \end{aligned}$$

In the above, F, G are smooth functions on T^*Q and the canonical Poisson bracket is represented, in local coordinates (q^i, p_i) for T^*Q, by

$$\{F, G\}(q, p) := \left(\frac{\partial F^{\mathrm{T}}}{\partial q^i}, \frac{\partial F^{\mathrm{T}}}{\partial p_i} \right) B^\sharp(q, p) \begin{pmatrix} \dfrac{\partial G}{\partial q^i} \\ \dfrac{\partial G}{\partial p_i} \end{pmatrix},$$

where

$$B^\sharp(q,p)=\begin{pmatrix}\{q^i,q^j\} & \{q^i,p_j\}\\ \{p_i,q^j\} & \{p_i,p_j\}\end{pmatrix}=\begin{pmatrix}0 & \delta^i_j\\ -\delta^j_i & 0\end{pmatrix}.$$

As previously illustrated, we choose an Ehresmann connection A such that $\mathcal{H}_q=\Delta_Q(q)$, where the constraint distribution Δ_Q is spanned by a set of m independent one-forms, which is given, in local coordinates $q^i=(r^\alpha,s^a)$ for Q, by

$$\omega^a=\mathrm{d}s^a+A^a_\alpha(r,s)\mathrm{d}r^\alpha.$$

Define the new coordinates $(q^i,\ \widetilde{p}_i)=(r^\alpha,s^a,\ \widetilde{p}_\alpha,\widetilde{p}_a)$ for T^*Q, as in [34,20], by

$$\widetilde{p}_\alpha=p_\alpha-A^a_\alpha p_a,$$

with some choice of complementary coordinates $\widetilde{p}_a$. We then employ the induced coordinates $(q^i,\ \widetilde{p}_\alpha)=(r^\alpha,s^a,\ \widetilde{p}_\alpha)$ for P.

In this context, the bundle map $B^\sharp_P:T^*P\to TP$ associated with $B_P=B|_{T^*P\times T^*P}$ can be constructed, using local coordinates $(q^i,\ \widetilde{p}_\alpha)$ for P, by computing $\{q^i,q^j\}$, $\{q^i,\widetilde{p}_\alpha\}$, $\{\widetilde{p}_\alpha,\widetilde{p}_\beta\}$; one finds that

$$\{q^i,q^j\}=0,\quad \{r^\beta,\widetilde{p}_\alpha\}=\delta^\beta_\alpha,\quad \{s^b,\widetilde{p}_\alpha\}=-A^b_\alpha,\quad \{\widetilde{p}_\alpha,\widetilde{p}_\beta\}=-K^b_{\alpha\beta}\,p_b.$$

Hence, it follows that

$$B^\sharp_P(q;\ \widetilde{p})=\begin{pmatrix}0 & 0 & \delta^\alpha_\beta\\ 0 & 0 & -A^a_\beta\\ -\delta^\beta_\alpha & (A^b_\alpha)^{\mathrm{T}} & -p_bK^b_{\alpha\beta}\end{pmatrix},\tag{5.9}$$

where δ^α_β is Kronecker's delta and $B^\sharp_P(q,\ \widetilde{p})$ is the $(2n-m)\times(2n-m)$ truncated matrix representation of the bundle map $B^\sharp_P:T^*P\to TP$, for each $(q^i,\ \widetilde{p}_\alpha)\in P$. This bundle map defines the bracket $\{,\}_P$ on the constrained submanifold P such that

$$\{F_P,G_P\}_P\,(q,\widetilde{p}):=\left(\frac{\partial F_P^{\mathrm{T}}}{\partial q^i},\frac{\partial F_P^{\mathrm{T}}}{\partial \widetilde{p}_\alpha}\right)B^\sharp_P(q,\widetilde{p})\begin{pmatrix}\dfrac{\partial G_P}{\partial q^i}\\[2mm] \dfrac{\partial G_P}{\partial \widetilde{p}_\alpha}\end{pmatrix}\tag{5.10}$$

for smooth functions F_P,G_P on P.

Thus, we can construct the constrained Dirac structure D_P in the form of Eq. (5.8).

Remarks. In Eq. (5.9), notice that the curvature $K^b_{\alpha\beta}$ measures the failure of the constraint distribution to be an integrable bundle and then, after restricting all terms to P, the term $-p_b\,K^b_{\alpha\beta}$ should be understood as $-(p_b)_P\,K^b_{\alpha\beta}$. However, we write it as $-p_b\,K^b_{\alpha\beta}$ for simplicity. Needless to say, $p_b=\frac{\partial L}{\partial w^b}$ holds and the Poisson bracket $\{,\}_P$ in Eq. (5.10) *does not satisfy the Jacobi identity* when the distribution is nonholonomic.

5.8. A vector field on the constraint momentum space

Let $\Delta_Q\subset TQ$ be a constraint distribution and $P=\mathbb{F}L(\Delta_Q)\subset T^*Q$. Let V be the vector subbundle of $T_P(T^*Q)$ defined, for each $p\in T^*_qQ$, by

$$V_{(q,p)}=\left\{\mathrm{vert}(\eta,p)\,|\,\eta\in\Delta^\circ_Q(q)\right\}.$$

In the above, $T_P(T^*Q)$ is the restriction of the tangent bundle of $T(T^*Q)$ to the constraint momentum space P and $\mathrm{vert}(\eta,p)$ is the vertical lift of $\eta\in T^*_qQ$ with respect to $p\in T^*_qQ$, which is described in coordinates as

$$\mathrm{vert}(\eta,p)=(q,p,0,\eta).$$

Marle [22] shows that the vector subbundle $T_P(T^*Q)$ is a direct sum of the vector subbundles TP and V:

$$T_P(T^*Q) = TP \oplus V.$$

In this context, the restriction of a vector field X on T^*Q to the constraint momentum space P, that is, $X|_P$, splits into a sum

$$X|_P = X_P + X_V, \tag{5.11}$$

where X_P is the *constrained vector field* that is tangent to P and X_V is a smooth section of the subbundle V, whose negative is called the *constraint force field*.

For details of the above construction, refer to Bloch [5], §5.8.

5.9. Implicit constrained Lagrangian systems

We now develop the notion of an implicit constrained Lagrangian system in the context of the constrained Dirac structure.

Definition 5.4. Let $L : TQ \to \mathbb{R}$ be a Lagrangian and let $\Delta_Q \subset TQ$ be a constraint distribution. The constraint momentum space $P \subset T^*Q$ is defined as $P = \mathbb{F}L(\Delta_Q)$. Let L_c be the constrained Lagrangian on Δ_Q defined by $L_c = L|\Delta_Q$. Let D_P be the constrained Dirac structure defined by the restriction of the induced Dirac structure D_{Δ_Q} on T^*Q to P. Denote by X a vector field on T^*Q, defined at points in P, and denote by X_P the constrained vector field on P given in Eq. (5.11). An *implicit constrained Lagrangian system* is a triple (L_c, Δ_Q, X_P), which satisfies the condition, for each $u \in \Delta_Q$,

$$(X_P(y), \mathfrak{D}L_c(u)) \in D_P(y),$$

where $y = \mathbb{F}L_c(u) \in P$ is the partial Legendre transformation. Here, the *Dirac differential of the constrained Lagrangian*, namely, $\mathfrak{D}L_c : \Delta_Q \to T^*P$, is defined at points $u \in \Delta_Q$ by $\mathfrak{D}L_c(u) = \mathfrak{D}L(u)|TP$, where, recall, $\mathfrak{D}L : TQ \to T^*T^*Q$.

The coordinate expression for $\mathfrak{D}L_c$ given below shows that it, in fact, depends only on derivatives of L_c.

Definition 5.5. A *solution curve* of the implicit constrained Lagrangian system (L_c, Δ_Q, X_P) is a curve $u(t) \in \Delta_Q(q(t)),\ t_1 \le t \le t_2$, such that $y(t)$ is an integral curve in P of X_P, where $y(t) = \mathbb{F}L_c(u(t))$.

5.10. Coordinate representation

Let $L : TQ \to \mathbb{R}$ be a Lagrangian, and let Δ_Q be a distribution on Q given, in coordinates $q = (r^\alpha, s^a)$ for Q, by

$$\omega^a(q) = \mathrm{d}s^a + A^a_\alpha(r,s)\,\mathrm{d}r^\alpha, \quad a = 1,\dots,m;\ \alpha = 1,\dots,n-m,$$

where ω^a indicate m-independent one-forms that consist of the basis of the annihilator of Δ_Q and an Ehresmann connection A is chosen such that $\mathcal{H}_q = \Delta_Q(q)$. The constrained Lagrangian $L_c = L|\Delta_Q$ is given, in coordinates, by

$$L_c(r^\alpha, s^a, u^\alpha) = L(r^\alpha, s^a, u^\alpha, -A^a_\alpha(r,s)u^\alpha),$$

and it follows that

$$\mathbf{d}L_c = \left(r^\alpha, s^a, u^\alpha, \frac{\partial L_c}{\partial r^\alpha}, \frac{\partial L_c}{\partial s^a}, \frac{\partial L_c}{\partial u^\alpha}\right).$$

Hence, we obtain the Dirac differential of L_c as

$$\mathfrak{D}L_c = \left(r^\alpha, s^a, \frac{\partial L_c}{\partial u^\alpha}, -\frac{\partial L_c}{\partial r^\alpha}, -\frac{\partial L_c}{\partial s^a}, u^\alpha\right) \tag{5.12}$$

and with

$$\widetilde{p}_\alpha = \frac{\partial L_c}{\partial u^\alpha}.$$

Let X_P be the constrained vector field on $P = \mathbb{F}L(\Delta_Q)$, which is denoted, by using coordinates $(q, \widetilde{p}) = (r^\alpha, s^a, \widetilde{p}_\alpha)$ for P, such that

$$X_P(r^\alpha, s^a, \widetilde{p}_\alpha) = (\dot{r}^\alpha, \dot{s}^a, \dot{\widetilde{p}}_\alpha). \tag{5.13}$$

Recall that the skew-symmetric bundle map $B_P^\sharp : T^*P \to TP$ can be constructed, by using the Ehresmann connection associated with the constraints, as in Eq. (5.9), and recall also that the constrained Dirac structure on P can be defined such that, for each $y = (r^\alpha, s^a, \widetilde{p}_\alpha) \in P$,

$$D_P(y) = \{(w_y, \alpha_y) \in T_yP \times T_y^*P \,|\, \alpha_y \in \Delta_P^*(y),\ w_y - B_P^\sharp(y)\,\alpha_y \in (\Delta_P^*)^\circ(y)\}. \tag{5.14}$$

An implicit constrained Lagrangian system is a triple (L_c, Δ_Q, X_P) that satisfies the condition, for each $u \in \Delta_Q(q)$,

$$(X_P(y), \mathfrak{D}L_c(u)) \in D_P(y),$$

where $y = \mathbb{F}L_c(u) \in P$. In view of Eqs. (5.9) and (5.12)–(5.14), the implicit constrained Lagrangian system (L_c, Δ_Q, X_P) may be described, in coordinates, by

$$\begin{pmatrix} \dot{r}^\alpha \\ \dot{s}^a \\ \dot{\widetilde{p}}_\alpha \end{pmatrix} = \begin{pmatrix} 0 & 0 & \delta_\beta^\alpha \\ 0 & 0 & -A_\beta^a \\ -\delta_\alpha^\beta & (A_\alpha^b)^{\mathrm{T}} & -p_b\, K_{\alpha\,\beta}^b \end{pmatrix} \begin{pmatrix} -\dfrac{\partial L_c}{\partial r^\beta} \\ -\dfrac{\partial L_c}{\partial s^b} \\ u^\beta \end{pmatrix} \tag{5.15}$$

and with the partial Legendre transform $\widetilde{p}_\alpha = \partial L_c / \partial u^\alpha$.

Then, a solution curve of the implicit constrained Lagrangian system (L_c, Δ_Q, X_P) is given by $u^\alpha(t) \in \Delta_Q(r^\alpha(t), s^a(t))$, $t_1 \le t \le t_2$, such that $y(t) = (r^\alpha(t), s^a(t), \widetilde{p}_\alpha(t))$ is the integral curve in P of X_P, where $y(t) = \mathbb{F}L_c(r^\alpha(t), s^a(t), u^\alpha(t))$. In other words, a solution curve of (L_c, Δ_Q, X_P) may also be described by $\widetilde{x}(t) = (r^\alpha(t), s^a(t), u^\alpha(t), \widetilde{p}_\alpha(t))$ in the subbundle $\Delta_Q \oplus P$ of the Pontryagin bundle $TQ \oplus T^*Q$ such that $y(t) = (r^\alpha(t), s^a(t), \widetilde{p}_\alpha(t))$ is the integral curve of X_P and such that $(X_P(y(t)), \mathfrak{D}L_c(u(t))) \in D_P(y(t))$.

We can summarize the results obtained so far in the following theorem.

Theorem 5.6. *Consider a Lagrangian L (possibly degenerate) on TQ and with a constraint distribution Δ_Q on Q. Let $P = \mathbb{F}L(\Delta_Q)$ be the constraint momentum space. Let $L_c = L|\Delta_Q$ be the constrained Lagrangian and D_P be the constrained Dirac structure on P. Let X_P be the constrained vector field on P such that (L_c, Δ_Q, X_P) is an implicit constrained Lagrangian system. Denote by $\widetilde{x}(t) = (r^\alpha(t), s^a(t), u^\alpha(t), \widetilde{p}_\alpha(t))$, $t_1 \le t \le t_2$, a curve in $\mathcal{K} = \Delta_Q \oplus P \subset TQ \oplus T^*Q$. Then, the following statements are equivalent:*

(a) *$\widetilde{x}(t)$ is a solution curve of the implicit constrained Lagrangian system (L_c, Δ_Q, X_P) in Eq. (5.15);*
(b) *$\widetilde{x}(t)$ satisfies the implicit Lagrange–d'Alembert equations in Eq. (5.6);*
(c) *$y(t) = (r^\alpha(t), s^a(t), \widetilde{p}_\alpha(t))$ is the integral curve of the constrained vector field X_P on P, where the partial Legendre transform $y(t) = \mathbb{F}L_c(r^\alpha(t), s^a(t), u^\alpha(t))$ holds.*

6. Examples

In this section, we demonstrate the implicit constrained Lagrangian system together with two examples. Namely, we illustrate the same examples of a vertical rolling disk on a plane and an L–C circuit as in Part I, for this purpose.

6.1. Example: The vertical rolling disk

Consider a vertical rolling disk on the xy-plane. Recall the configuration space of the system is denoted by $Q = \mathbb{R}^2 \times S^1 \times S^1$, whose coordinates are given by $q = (x, y, \theta, \varphi)$, where x, y indicate the position of the contact point of the disk, θ the rotation angle of the disk and φ the orientation of the disk. Recall that the Lagrangian is given by

$$L(x, y, \theta, \varphi, v_x, v_y, v_\theta, v_\varphi) = \frac{1}{2}m(v_x^2 + v_y^2) + \frac{1}{2}I v_\theta^2 + \frac{1}{2}J v_\varphi^2.$$

In the above, m indicates the mass, and I and J the moment of inertia. Recall also that the constraints are given by the constraint distribution $\Delta_Q \subset TQ$ such that, for each $q \in Q$,

$$\Delta_Q(q) = \left\{v_q \in T_q Q \mid \left\langle \omega^a(q), v_q \right\rangle = 0, a = 1, 2\right\},$$

where $v_q = (v_x, v_y, v_\theta, v_\varphi)$ and the one-forms ω^a are given by

$$\omega^1 = \mathrm{d}x - R\,(\cos\varphi)\,\mathrm{d}\theta,$$
$$\omega^2 = \mathrm{d}y - R\,(\sin\varphi)\,\mathrm{d}\theta,$$

where R denotes radius of the disk.

Let us choose a bundle structure $\pi : Q \to \mathcal{R}$ such that the base $\mathcal{R}$ is to be $S^1 \times S^1$ parameterized by θ and φ together with the projection to $\mathcal{R}$, that is, $(r^1, r^2, s^1, s^2) = (\theta, \varphi, x, y) \mapsto (r^1, r^2) = (\theta, \varphi)$. Then, the Ehresmann connection can be constructed by

$$A = \omega^a \frac{\partial}{\partial s^a}, \quad \omega^a(q) = \mathrm{d}s^a + A^a_\alpha(r, s)\mathrm{d}r^\alpha.$$

The components of the Ehresmann connection are given by

$$A^1_1 = -R\,(\cos\varphi), \quad A^2_1 = -R\,(\sin\varphi)$$

and the remaining components are zero.

As in Eq. (5.9), the bundle map $B^\sharp_P : T^*P \to TP$ is locally denoted by

$$B^\sharp_P(r, s, \widetilde{p}) = \left(\begin{array}{cc|cc|cc} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & R\,(\cos\varphi) & 0 \\ 0 & 0 & 0 & 0 & R\,(\sin\varphi) & 0 \\ \hline -1 & 0 & -R\,(\cos\varphi) & -R\,(\sin\varphi) & 0 & p_b\,K^b_{12} \\ 0 & -1 & 0 & 0 & -p_b\,K^b_{21} & 0 \end{array}\right),$$

and the components of the curvature K is given by

$$K^1_{12} = -K^1_{21} = R\sin\varphi, \quad K^2_{12} = -K^2_{21} = -R\cos\varphi.$$

Thus, the constrained Dirac structure D_P on P can be defined as in Eq. (5.14).

Meanwhile, the constrained Lagrangian $L_c(r^\alpha, s^a, u^\alpha) = L(r^\alpha, s^a, u^\alpha, -A^a_\alpha(r, s)u^\alpha)$ is given by

$$L_c(\theta, \varphi, x, y, v_\theta, v_\varphi) = \frac{1}{2}(mR^2 + I)v_\theta^2 + \frac{1}{2}J v_\varphi^2,$$

where $(r^\alpha, s^a, u^\alpha) = (\theta, \varphi, x, y,\ v_\theta, v_\varphi)$. Then, the differential of L_c is locally expressed by

$$\begin{aligned}\mathbf{d}L_c(\theta, \varphi, x, y, v_\theta, v_\varphi) &= \left(\frac{\partial L_c}{\partial\theta}, \frac{\partial L_c}{\partial\varphi}, \frac{\partial L_c}{\partial x}, \frac{\partial L_c}{\partial y}, \frac{\partial L_c}{\partial v_\theta}, \frac{\partial L_c}{\partial v_\varphi}\right) \\ &= \left(0, 0, 0, 0, (mR^2 + I)v_\theta, J v_\varphi\right)\end{aligned}$$

and hence the Dirac differential of L_c is locally denoted by

$$\begin{aligned}\mathfrak{D}L_c(\theta, \varphi, x, y,\ v_\theta, v_\varphi) &= \left(-\frac{\partial L_c}{\partial\theta}, -\frac{\partial L_c}{\partial\varphi}, -\frac{\partial L_c}{\partial x}, -\frac{\partial L_c}{\partial y},\ v_\theta,\ v_\varphi\right) \\ &= \left(0, 0, 0, 0,\ v_\theta,\ v_\varphi\right),\end{aligned}$$

and with the partial Legendre transform

$$\widetilde{p}_\theta = \frac{\partial L_c}{\partial v_\theta} = (mR^2 + I)v_\theta,$$
$$\widetilde{p}_\varphi = \frac{\partial L_c}{\partial v_\varphi} = J v_\varphi.$$

Since $(r^\alpha, s^a, \widetilde{p}_\alpha) = (\theta, \varphi, x, y, \widetilde{p}_\theta, \widetilde{p}_\varphi)$ are local coordinates for P, the constrained vector field X_P on P is locally denoted by

$$X_P(\theta, \varphi, x, y, \widetilde{p}_\theta, \widetilde{p}_\varphi) = (\dot{\theta}, \dot{\varphi}, \dot{x}, \dot{y}, \dot{\widetilde{p}}_\theta, \dot{\widetilde{p}}_\varphi).$$

Hence, we obtain the coordinate representation of the implicit constrained Lagrangian system (L_c, Δ_Q, X_P) such that

$$\left(\begin{array}{c} \dot{\theta} \\ \dot{\varphi} \\ \hline \dot{x} \\ \dot{y} \\ \hline \dot{\widetilde{p}}_\theta \\ \dot{\widetilde{p}}_\varphi \end{array}\right) = \left(\begin{array}{cc|cc|cc} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & R\,(\cos\varphi) & 0 \\ 0 & 0 & 0 & 0 & R\,(\sin\varphi) & 0 \\ \hline -1 & 0 & -R\,(\cos\varphi) & -R\,(\sin\varphi) & 0 & p_b\,K^b_{12} \\ 0 & -1 & 0 & 0 & -p_b\,K^b_{21} & 0 \end{array}\right) \left(\begin{array}{c} 0 \\ 0 \\ \hline 0 \\ 0 \\ \hline v_\theta \\ v_\varphi \end{array}\right),$$

where $\widetilde{p}_\theta = \left(mR^2 + I\right) v_\theta$ and $\widetilde{p}_\varphi = J v_\varphi$.

By computations, it follows that

$$\begin{aligned} \dot{\widetilde{p}}_\theta &= p_b\,K^b_{12}\,v_\varphi \\ &= \{m\,v_x\,(R\sin\varphi) - m\,v_y\,(R\cos\varphi)\}\,v_\varphi \\ &= \{m\,(R\cos\varphi)\,(R\sin\varphi)\,v_\theta - m\,(R\sin\varphi)\,(R\cos\varphi)\,v_\theta\}\;v_\varphi \\ &= 0, \\ \dot{\widetilde{p}}_\varphi &= -p_b\,K^b_{21}\,v_\theta \\ &= \{m\,v_x\,(R\sin\varphi) - m\,v_y\,(R\cos\varphi)\}\,v_\theta \\ &= \{m\,(R\cos\varphi)\,(R\sin\varphi)\,v_\theta - m\,(R\sin\varphi)\,(R\cos\varphi)\,v_\theta\}\;v_\theta \\ &= 0, \end{aligned}$$

where $p_b = (p_x, p_y)$ are given by

$$p_x = \frac{\partial L}{\partial v_x} = m\,v_x = m\,R\,(\cos\varphi)\,v_\theta,$$
$$p_y = \frac{\partial L}{\partial v_y} = m\,v_y = m\,R\,(\sin\varphi)\,v_\theta.$$

Thus, we obtain the equations of motion in the context of implicit constrained Lagrangian systems as

$$\dot{\theta} = v_\theta, \quad \dot{\varphi} = v_\varphi, \quad \dot{\widetilde{p}}_\theta = 0, \quad \dot{\widetilde{p}}_\varphi = 0,$$
$$\widetilde{p}_\theta = (mR^2 + I)v_\theta, \quad \widetilde{p}_\varphi = J v_\varphi.$$

6.2. Example: L–C circuits

Let us consider the same example of an L–C circuit as illustrated in Part I in the context of implicit constrained Lagrangian systems.

Recall that the configuration space $E = \mathbb{R}^4$, whose element denotes the charge q and its local coordinates are given by $(q_L, q_{C_1}, q_{C_2}, q_{C_3})$ and recall also that the KCL constraints form a constraint subspace called the constraint KCL space $\Delta \subset TE$, which is defined, for each $q \in E$, by

$$\Delta_q = \{f \in T_qE \mid \langle \omega^a, f \rangle = 0,\ a = 1, 2\}.$$

Note that $f = (f_L, f_{C_1}, f_{C_2}, f_{C_3}) \in T_qE$ denotes the current vector and ω^a indicate independent covectors (or one-forms), which form the basis for the annihilator $\Delta_q^\circ \subset T_q^*E$ and are given, in coordinates, by

$$\omega^1 = -\mathrm{d}q_L + \mathrm{d}q_{C_2},$$
$$\omega^2 = -\mathrm{d}q_{C_1} + \mathrm{d}q_{C_2} - \mathrm{d}q_{C_3}.$$

Therefore, the KCL constraints for currents $f = (f_L, f_{C_1}, f_{C_2}, f_{C_3}) \in T_qE$ are represented, in coordinates, by

$$-f_L + f_{C_2} = 0,$$
$$-f_{C_1} + f_{C_2} - f_{C_3} = 0.$$

Choose a bundle structure $\pi : E \to \mathcal{R}$ such that the base $\mathcal{R}$ is to be $\mathbb{R}^2$ parameterized by $(r^1, r^2) = (q_{C_2}, q_{C_3})$ together with the projection to $\mathcal{R}$, that is,

$$(r^1, r^2, s^1, s^2) = (q_{C_2}, q_{C_3}, q_L, q_{C_1}) \mapsto (r^1, r^2) = (q_{C_2}, q_{C_3}),$$

where we choose an Ehresmann connection in such a way that $\mathcal{H}_q = \Delta_Q(q)$. The connection A is described, in local coordinates $q = (r^1, r^2, s^1, s^2) = (q_{C_2}, q_{C_3}, q_L, q_{C_1})$ for $E = \mathbb{R}^4$, by a vertical valued one-form ω^a such that

$$A = \omega^a \frac{\partial}{\partial s^a}, \quad \omega^a = \mathrm{d}s^a + A^a_\alpha \mathrm{d}r^\alpha, \quad a = 1, 2, \quad \alpha = 1, 2,$$

where the components of A are locally represented by

$$A^a_\alpha = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Therefore, the KCL constraints may be rewritten as

$$f^a = -A^a_\alpha f^\alpha,$$

that is, in matrix representation,

$$\begin{pmatrix} f_L \\ f_{C_1} \end{pmatrix} = -\begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} f_{C_2} \\ f_{C_3} \end{pmatrix}.$$

As in Part I, recall that the Lagrangian $\mathcal{L}$ on TE is locally given by

$$\mathcal{L} = \frac{1}{2} L\,(f_L)^2 - \frac{1}{2}\frac{(q_{C_1})^2}{C_1} - \frac{1}{2}\frac{(q_{C_2})^2}{C_2} - \frac{1}{2}\frac{(q_{C_3})^2}{C_3}.$$

Hence, we have the equations of motion

$$\dot{p}_\alpha - \frac{\partial \mathcal{L}}{\partial r^\alpha} = A^a_\alpha \left(\dot{p}_a - \frac{\partial \mathcal{L}}{\partial s^a} \right),$$

which may be denoted, in matrix form, by

$$\begin{pmatrix} \dot{p}_{C_2} + \dfrac{q_{C_2}}{C_2} \\ \dot{p}_{C_3} + \dfrac{q_{C_3}}{C_3} \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \dot{p}_L \\ \dot{p}_{C_1} + \dfrac{q_{C_2}}{C_3} \end{pmatrix}.$$

The flux linkages $p = (p_\alpha, p_a) = (p_{C_2}, p_{C_3}, p_L, p_{C_1})$ are defined by

$$p_\alpha = \frac{\partial \mathcal{L}}{\partial f^\alpha}, \quad p_a = \frac{\partial \mathcal{L}}{\partial f^a},$$

and it reads that

$$p_{C_2} = 0, \quad p_{C_3} = 0, \quad p_L = L\,f_L, \quad p_{C_1} = 0.$$

Note that the constraint flux linkage space is defined by

$$P = \mathbb{F}\mathcal{L}(\Delta) \subset T^*E,$$

where the distribution $\Delta \subset TE$ is represented, in coordinates, by

$$\Delta = \mathrm{span}\left\{\frac{\partial}{\partial r^\alpha} - A^a_\alpha \frac{\partial}{\partial s^a}\right\}.$$

Define the new coordinates $(r^\alpha, s^a, \widetilde{p}_\alpha, \widetilde{p}_a)$ for T^*E such that

$$\widetilde{p}_\alpha = p_a - A^a_\alpha\, p_a$$

with some choice of complementary coordinates $\widetilde{p}_a$. Hence, we employ the induced coordinates $(r^\alpha, s^a, \widetilde{p}_\alpha) = (q_{C_2}, q_{C_3}, q_L, q_{C_1}, \widetilde{p}_{C_2}, \widetilde{p}_{C_3})$ for P.

The bundle map $B^\sharp_P : T^*P \to TP$ associated with the constrained Poisson structure B_P on P can be constructed by computing $\{q^i, q^j\} = 0$, $\{r^\beta, \widetilde{p}_\alpha\} = \delta^\beta_\alpha$, $\{s^b, \widetilde{p}_\alpha\} = -A^b_\alpha$, $\{\widetilde{p}_\alpha, \widetilde{p}_\beta\} = 0$ such that

$$B^\sharp_P(q, \widetilde{p}_\alpha) = \begin{pmatrix} 0 & 0 & \delta^\alpha_\beta \\ 0 & 0 & -A^a_\beta \\ -\delta^\beta_\alpha & (A^b_\alpha)^{\mathrm{T}} & 0 \end{pmatrix}.$$

Since the KCL constraints are holonomic, the curvature $K^d_{\alpha\beta}$ of the Ehresmann connection A does not appear in $B^\sharp_P(q, \widetilde{p}_\alpha)$, and it immediately reads that the Jacobi identity holds.

As previously mentioned, the set of $B^\sharp_P$ and Δ^*_P defines the Dirac structure $D_P \subset TP \oplus T^*P$ on P, whose fiber is given, for each $y \in P$, by

$$D_P(y) = \{(v_y, \alpha_y) \in T_yP \times T^*_yP \mid \alpha_y \in \Delta^*_P(y) \text{ and } v_y - B^\sharp_P(y)\,\alpha_y \in (\Delta^*_P)^\circ(y)\}.$$

The constrained Lagrangian for the L–C circuit may be constructed as

$$\mathcal{L}_c(r^\alpha, s^a, f^\alpha) = \mathcal{L}(r^\alpha, s^a, f^\alpha, -A^a_\alpha f^\alpha)$$

and we obtain

$$\mathcal{L}_c(q_{C_2}, q_{C_3}, q_L, q_{C_1}, f_{C_2}, f_{C_3}) = \frac{1}{2}L\,(f_{C_2})^2 - \frac{1}{2}\frac{(q_{C_1})^2}{C_1} - \frac{1}{2}\frac{(q_{C_2})^2}{C_2} - \frac{1}{2}\frac{(q_{C_3})^2}{C_3}.$$

Then, it follows from the partial Legendre transformation that

$$\widetilde{p}_{C_2} = \frac{\partial \mathcal{L}_c}{\partial f_{C_2}} = L\, f_{C_2}, \quad \widetilde{p}_{C_3} = \frac{\partial \mathcal{L}_c}{\partial f_{C_3}} = 0,$$

which exactly correspond to the equality of the base points. By using local coordinates $(r^\alpha, s^a, \widetilde{p}_\alpha) = (q_{C_2}, q_{C_3}, q_L, q_{C_1}, \widetilde{p}_{C_2}, 0)$ for P, the constrained vector field X_P on P can be denoted as

$$\begin{aligned} X_P(r^\alpha, s^a, \widetilde{p}_\alpha) &= (\dot{r}^\alpha, \dot{s}^a, \dot{\widetilde{p}}_\alpha), \\ &= \left(\dot{q}_{C_2}, \dot{q}_{C_3}, \dot{q}_L, \dot{q}_{C_1}, \dot{\widetilde{p}}_{C_2}, 0\right), \end{aligned}$$

while the Dirac differential of $\mathcal{L}_c$ is given by

$$\begin{aligned} \mathfrak{D}\mathcal{L}_c(r^\alpha, s^a, f^\alpha) &= \left(-\frac{\partial \mathcal{L}_c}{\partial r^\alpha}, -\frac{\partial \mathcal{L}_c}{\partial s^a}, f^\alpha\right) \\ &= \left(\frac{q_{C_2}}{C_2}, \frac{q_{C_3}}{C_3}, 0, \frac{q_{C_1}}{C_1}, f_{C_2}, f_{C_3}\right). \end{aligned}$$

Then, the L–C circuit can be expressed as an implicit constrained Lagrangian system, since the triple $(\mathcal{L}_c, \Delta, X_P)$ satisfies

$$(X_P, \mathfrak{D}\mathcal{L}_c) \in D_P,$$

which is represented, in coordinates, by using the bundle map $B_P^\sharp(q,\ \widetilde{p})$ as

$$\begin{pmatrix} \dot{r}^\alpha \\ \dot{s}^a \\ \dot{\widetilde{p}}_\alpha \end{pmatrix} = \begin{pmatrix} 0 & 0 & \delta^\alpha_\beta \\ 0 & 0 & -A^a_\beta \\ -\delta^\beta_\alpha & (A^b_\alpha)^{\mathrm{T}} & 0 \end{pmatrix} \begin{pmatrix} -\dfrac{\partial \mathcal{L}_c}{\partial r^\beta} \\ -\dfrac{\partial \mathcal{L}_c}{\partial s^b} \\ f^\beta \end{pmatrix}.$$

Therefore, we have

$$\left(\begin{array}{c} \dot{q}_{C_2} \\ \dot{q}_{C_3} \\ \hline \dot{q}_L \\ \dot{q}_{C_1} \\ \hline \dot{\widetilde{p}}_{C_2} \\ 0 \end{array}\right) = \left(\begin{array}{cc|cc|cc} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ \hline -1 & 0 & -1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \end{array}\right) \left(\begin{array}{c} \frac{q_{C_2}}{C_2} \\ \frac{q_{C_3}}{C_3} \\ \hline 0 \\ \frac{q_{C_1}}{C_1} \\ \hline f_{C_2} \\ f_{C_3} \end{array}\right),$$

where $\widetilde{p}_{C_2} = L\, f_{C_2}$ holds. Furthermore, we can eliminate the components associated with the current $\dot{q}_L$ of the inductor L, and it reads

$$\left(\begin{array}{c} \dot{q}_{C_2} \\ \dot{q}_{C_3} \\ \dot{q}_{C_1} \\ \hline \dot{\widetilde{p}}_{C_2} \\ 0 \end{array}\right) = \left(\begin{array}{ccc|cc} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ \hline -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \end{array}\right) \left(\begin{array}{c} \frac{q_{C_2}}{C_2} \\ \frac{q_{C_3}}{C_3} \\ \frac{q_{C_1}}{C_1} \\ \hline f_{C_2} \\ f_{C_3} \end{array}\right).$$

Thus, the reduced equations of motion of the L–C circuit are derived in the context of the implicit constrained Lagrangian system as follows:

$$\begin{aligned} &\dot{q}_{C_1} = f_{C_2} - f_{C_3}, \quad \dot{q}_{C_2} = f_{C_2}, \quad \dot{q}_{C_3} = f_{C_3}, \\ &\dot{\widetilde{p}}_{C_2} = -\frac{q_{C_1}}{C_1} - \frac{q_{C_2}}{C_2}, \quad \widetilde{p}_{C_2} = L\, f_{C_2}, \\ &0 = -\frac{q_{C_3}}{C_3} + \frac{q_{C_1}}{C_1}. \end{aligned}$$

Remarks. Note that the original Lagrangian $\mathcal{L}(q^i, f^i)$ and the constrained Lagrangian $\mathcal{L}_c(r^\alpha, s^a, f^\alpha)$ are independent of q_L, which implies q_L might be a "secret" variable that is related to symmetries.

7. Conclusions

Part I showed that a constrained distribution on a manifold together with the canonical two-form induces a Dirac structure on the cotangent bundle. It was shown that some basic examples, such as KCL and KVL constraints in electric circuits, interconnections, as well as nonholonomic constraints, naturally fit into this context. Utilizing the symplectomorphisms between the iterated tangent and cotangent bundles, Part I also developed the notion of an implicit Lagrangian system (L, Δ_Q, X) in this context.

In Part II of the paper, we established the link between variational structures and implicit Lagrangian systems in mechanics. To do this, use was made of an extension of the variational principle of Hamilton, called the Hamilton–Pontryagin principle, which, in the case $\Delta_Q = TQ$, leads to a set of equations that naturally includes the Legendre transformation as well as the Euler–Lagrange equations themselves. For the case of a general constraint distribution $\Delta_Q \subset TQ$, we showed that an implicit Lagrangian system can be derived from a generalization of the Hamilton–Pontryagin principle, namely an extended Lagrange–d'Alembert principle called the Lagrange–d'Alembert–Pontryagin principle. We also proposed a generalization of Hamilton's phase space principle called the Hamilton–d'Alembert principle in phase space and used this to establish the relationship with implicit Hamiltonian systems for the case of regular Lagrangians.

In conjunction with applications to controlled interconnected systems such as robots and electromechanical systems, we demonstrated that nonholonomic mechanical systems with external forces naturally fall into the context of implicit Lagrangian systems.

Furthermore, we developed the notion of an implicit constrained Lagrangian system (L_c, Δ_Q, X_P), by introducing the constrained Dirac structure D_P on the constraint momentum space P and the constrained vector field X_P on P; the constrained Dirac structure D_P on P can be naturally defined by restricting the induced Dirac structure D_{Δ_Q} on T^*Q to P and we have shown that D_P can be constructed by using an Ehresmann connection associated with the constraint distribution. Also, implicit constrained Lagrangian systems were shown to fit naturally into the context of the Lagrange–d'Alembert–Pontryagin principle. Finally, two examples were given, namely, a vertical rolling disk on a plane as an example of a nonholonomic mechanical system and an L–C circuit as an example of a degenerate Lagrangian system with holonomic constraints.

Some interesting topics for future work are as follows:

- Implicit Lagrangian systems with symmetry, Dirac reduction and links with, for example, [10]. Specifically, it would be interesting to explore the Euler–Poincaré and Lie–Poisson equations from this point of view.
- The momentum equations and the reduced Lagrange–d'Alembert equations in the context of Dirac structures; see [11] and references therein.
- The relationship between implicit Lagrangian systems and implicit Hamiltonian systems for degenerate Lagrangians, using a generalized Legendre transform and the theory of Dirac constraints.
- An analog of controlled Lagrangians and related stability problems for implicit Lagrangian systems [6,7,41].
- Discrete mechanics and variational integrators for implicit Lagrangian systems from the viewpoint of the Hamilton–Pontryagin principle.
- Applications to interconnected systems such as multibody systems, general electric circuits, and networks including sensing and communications, electromechanical systems, biochemical systems.

Acknowledgements

This paper was written during a visit of this author during 2002–2003 in the Department of Control and Dynamical Systems, California Institute of Technology, Pasadena, CA 91125, USA. The first author's research was partially supported by JSPS Grant 16560216. The second author's research was partially supported by NSF-ITR Grant ACI-0204932.

References

[1] R. Abraham, J.E. Marsden, Foundations of Mechanics, 2nd ed., Benjamin-Cummings Publ. Co, 1978 (Updated 1985 version, reprinted by Persius Publishing).
[2] L. Bates, J. Sniatycki, Nonholonomic reduction, Rep. Math. Phys. 32 (1993) 99–115.
[3] G. Blankenstein, Implicit Hamiltonian systems: Symmetry and interconnection, Ph.D. Dissertation, University of Twente, 2000.
[4] G. Blankenstein, T.S. Ratiu, Singular reduction of implicit Hamiltonian systems, Rep. Math. Phys. 53 (2) (2004) 211–260.
[5] A.M. Bloch, Nonholonomic Mechanics and Control, in: Interdisciplinary Applied Mathematics, vol. 24, Springer-Verlag, New York, 2003.With the collaboration of J. Baillieul, P. Crouch and J. Marsden, and with scientific input from P.S. Krishnaprasad, R.M. Murray and D. Zenkov.
[6] A.M. Bloch, N.E. Leonard, J.E. Marsden, Controlled Lagrangians and the stabilization of mechanical systems I: The first matching theorem, IEEE Trans. Automat. Control 45 (2000) 2253–2270.
[7] A.M. Bloch, D.E. Chang, N.E. Leonard, J.E. Marsden, Controlled Lagrangians and the stabilization of mechanical systems II: Potential shaping, IEEE Trans. Automat. Control 46 (2001) 1556–1571.
[8] A.M. Bloch, P.S. Krishnaprasad, J.E. Marsden, R. Murray, Nonholonomic mechanical systems with symmetry, Arch. Ration. Mech. Anal. 136 (1996) 21–99.

[9] H. Cendra, D.D. Holm, M.J.W. Hoyle, J.E. Marsden, The Maxwell–Vlasov equations in Euler–Poincaré form, J. Math. Phys. 39 (1998) 3138–3157.
[10] H. Cendra, J.E. Marsden, S. Pekarsky, T.S. Ratiu, Variational principles for Lie–Poisson and Hamilton–Poincaré equations, Mosc. Math. J. 3 (2003) 833–867.
[11] H. Cendra, J.E. Marsden, T.S. Ratiu, Geometric mechanics, Lagrangian reduction and nonholonomic systems, in: B. Enquist, W. Schmid (Eds.), Mathematics Unlimited-2001 and Beyond, Springer-Verlag, New York, 2001, pp. 221–273.
[12] L.O. Chua, J.D. McPherson, Explicit topological formulation of Lagrangian and Hamiltonian equations for nonlinear networks, IEEE Trans. Circuits and Syst. CAS-21 (1974) 277–286.
[13] T.J. Courant, Dirac manifolds, Trans. Amer. Math. Soc. 319 (1990) 631–661.
[14] T.J. Courant, Tangent Dirac structures, J. Phys. A: Math. Gen. 23 (1990) 5153–5168.
[15] P.A.M. Dirac, Generalized Hamiltonian dynamics, Canad. J. Math. 2 (1950) 129–148.
[16] P.A.M. Dirac, Lectures on Quantum Mechanics, Belfer Graduate School of Science, Yeshiva University, New York, 1964.
[17] I. Dorfman, Dirac structures of integrable evolution equations, Phys. Lett. A 125 (5) (1987) 240–246.
[18] I. Dorfman, Dirac structures and integrability of nonlinear evolution equations, in: Nonlinear Science Theory and Applications, John Wiley & Sons Ltd., Chichester, 1993.
[19] W.S. Koon, J.E. Marsden, The Hamiltonian and Lagrangian approaches to the dynamics of nonholonomic systems, Rep. Math. Phys. 40 (1997) 21–62.
[20] W.S. Koon, J.E. Marsden, Poisson reduction for nonholonomic mechanical systems with symmetry, Rep. Math. Phys. 42 (1998) 101–134.
[21] G.H. Livens, On Hamilton's principle and the modified function in analytical dynamics, Proc. Roy. Soc. Edinburgh 39 (1919) 113.
[22] C.-M. Marle, Various approaches to conservative and nonconservative nonholonomic systems, Rep. Math. Phys. 42 (1998) 211–229.
[23] J.E. Marsden, T.J.R. Hughes, Mathematical Foundations of Elasticity, Prentice-Hall, 1983 (Reprinted by Dover Publications, NY, 1994).
[24] J.E. Marsden, T.S. Ratiu, Reduction of Poisson manifold, Lett. Math. Phys. 11 (1986) 161–169.
[25] J.E. Marsden, T.S. Ratiu, Introduction to Mechanics and Symmetry, 2nd ed., in: Texts in Applied Mathematics, vol. 17, Springer-Verlag, 1999.
[26] L. Moreau, D. Aeyels, A variational principle for nonlinear LC circuits with arbitrary interconnection structure, the 5th IFAC Symposium of Nonlinear Control Systems, NOLCOS '01, Saint-Petersburg, July 2001 (preprint).
[27] L. Moreau, A.J. van der Schaft, Implicit Lagrangian equations and the mathematical modeling of physical systems, in: Proc. 41st IEEE Conf. Decision and Control, Las Vegas, Nevada, December 2002.
[28] L.A. Pars, A Treatise on Analytical Dynamics, Wiley, 1965.
[29] L.S. Pontryagin, V.G. Boltyanskiĭ, R.V. Gamkrelidze, E.F. Mishchenko, Mathematical Theory of Optimal Processes (K.N. Trirogoff, Trans.), Wiley-Interscience, New York, 1962.
[30] H.J. Sussmann, An introduction to the coordinate-free maximum principle, in: B. Jackubczyk, W. Respondek (Eds.), Geometry of Feedback and Optimal Control, in: Monographs Textbooks Pure. Appl. Math., vol. 207, M. Dekker, 1998, pp. 463–557.
[31] W.M. Tulczyjew, Hamiltonian systems, Lagrangian systems and the Legendre transformation, Symp. Math. 14 (1974) 247–258.
[32] W.M. Tulczyjew, P. Urbański, A slow and careful Legendre transformation for singular Lagrangians, Acta Phys. Polon. B 30 (1999) 2909–2978.
[33] A.J. van der Schaft, Implicit Hamiltonian systems with symmetry, Rep. Math. Phys. 41 (1998) 203–221.
[34] A.J. van der Schaft, B.M. Maschke, On the Hamiltonian formulation of nonholonomic mechanical systems, Rep. Math. Phys. 34 (1994) 225–233.
[35] A.J. van der Schaft, B.M. Maschke, The Hamiltonian formulation of energy conserving physical systems with external ports, Arch. Elektronik Übertragungstechnik 49 (1995) 362–371.
[36] A.M. Vershik, L.D. Faddeev, Lagrangian mechanics in invariant form, Sel. Math. Sov. 1 (1981) 339–350.
[37] K. Washizu, Variational Methods in Elasticity and Plasticity, 3rd ed., in: International Series of Monographs in Aeronautics and Astronautics: Solid and Structural Mechanics, Vol. 9, Pergamon Press, Oxford, 1982.
[38] A. Weinstein, Symplectic manifolds and their Lagrangian submanifolds, Adv. Math. 6 (1971) 329–346.
[39] A. Weinstein, Lectures on Symplectic Manifolds, in: Regional Conference Series in Mathematics, vol. 29, Amer. Math, Soc., Providence, RI, 1977.
[40] A. Weinstein, Lagrangian mechanics and groupoids, Fields Inst. Commun. 7 (1996) 207–231.
[41] D.V. Zenkov, A.M. Bloch, J.E. Marsden, The energy momentum method for the stability of nonholonomic systems, Dyn. Stab. Syst. 13 (1998) 123–166.