Available online at www.sciencedirect.com
ScienceDirect
Journal of Geometry and Physics 57 (2006) 133–156
JOURNAL OF GEOMETRY AND PHYSICS
www.elsevier.com/locate/jgp

Dirac structures in Lagrangian mechanics Part I: Implicit Lagrangian systems

Hiroaki Yoshimura[a,1], 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 21 January 2006; accepted 19 February 2006
Available online 29 March 2006

Abstract

This paper develops the notion of implicit Lagrangian systems and presents some of their basic properties in the context of Dirac structures. This setting includes degenerate Lagrangian systems and systems with both holonomic and nonholonomic constraints, as well as networks of Lagrangian mechanical systems. The definition of implicit Lagrangian systems with a configuration space Q makes use of Dirac structures on T^*Q that are induced from a constraint distribution on Q as well as natural symplectomorphisms between the spaces T^*TQ, TT^*Q, and T^*T^*Q. Two illustrative examples are presented; the first is a nonholonomic system, namely a vertical disk rolling on a plane, and the second is an L–C circuit, a degenerate Lagrangian system with holonomic constraints.

Keywords: Dirac structures; Implicit Lagrangian systems; Nonholonomic systems

1. Introduction

1.1. History and background

The idea of a system made up of interconnected subsystems has played a key role in the modeling of complex physical systems as a network; that is, these systems are regarded as an aggregation of subsystems interconnected by physical connections, information exchange, or sensors. This interconnection view was originally developed in [30] for the analysis of electromechanical systems and has been widely employed in various physical and engineering systems such as electric circuits, chemical reaction systems, and mechanical systems (see, for example, [31,11, 36,38,50]). Owing to a large number of (implicit) algebraic equations that appear in interconnecting subsystems, mathematical models of the associated system dynamics will normally be implicit differential-algebraic equations—regardless of which dynamical formalism, the Lagrangian or Hamiltonian, is employed.

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

[1] 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.

0393-0440/$ - see front matter
doi:10.1016/j.geomphys.2006.02.009

In mechanics, interconnected and implicit systems play a key role in, for example, controlled mechanical systems like robots. An important class of implicit mechanical systems is those with nonholonomic constraints, which has a long and rich history; see, for instance, [35] and [8]. The Lagrangian and Hamiltonian approaches for such systems have been extensively developed, including symmetry and reduction [32]. On the Lagrangian side, see, for example, [47,27,9], while on the Hamiltonian side, refer to [48,2,44,29]. One of the interesting aspects of the Hamiltonian side is the fact that the Poisson bracket fails to satisfy the Jacobi identity; such bracket structures are known as *almost Poisson structures*. The equivalence of the Lagrangian and Hamiltonian approaches to nonholonomic systems was shown in [28].

In the field of electric circuits and networks, the idea of interconnected systems, also referred to as nonenergic systems (or multiports), has been often employed in modeling and analysis of the systems, as in the work in [25, 26,31,49,11]. The term *nonenergic system*, which is sometimes used for these systems, was first coined in [3] in the context of Lagrangian mechanics [37].

The ideas underlying interconnected systems, which appear in energy-conserving systems such as L–C circuits and nonholonomic systems, were put into the context of Poisson structures in [44,33], and later in the general context of Dirac structures in [45,7]. The idea of interconnections was investigated in the context of implicit Hamiltonian systems in [22,46] and [4].

An L–C circuit represents a fundamental example in our context since it involves, as we shall see, a system *that has a degenerate Lagrangian as well as nontrivial additional constraints (the Kirchhoff current laws)*. In electric circuits, the Brayton–Moser equations (see [10]) have been known as basic equations of circuit dynamics, which were employed, for instance, in the analysis of jumping phenomena that can appear in nonlinear circuits (see, for example, [40]). Smale [42] investigated the Brayton–Moser equations from a geometric point of view, and in doing so, he introduced a symmetric nondegenerate bilinear form on a subspace of a direct sum of a vector space V and its dual space V^* and then derived equations of motion by a vector field associated with a one-form using this bilinear form. This is reminiscent of a symplectic structure on $V \oplus V^*$, which may be viewed as providing an early version of a geometric formulation that is linked with a Dirac structure on V.

The above works take a Hamiltonian system point of view of interconnected and implicit systems and, in addition, make use of Dirac structures. One of our main points is to extend this theory to also include the variational, that is, the Lagrangian mechanics point of view. Interestingly, this Lagrangian development also makes use of the framework of Dirac structures, and developing this idea is one of our main objectives.

An algebraic theory of Dirac structure associated with formal variational calculus was proposed in [24] in the Hamiltonian framework of integrable evolution equations. One the other hand, the geometric development of a Dirac structure on a manifold M, as a subset $D \subset TM \oplus T^*M$ satisfying certain conditions (reviewed below), was introduced in [21] and [19]. The original goal of these authors was to find a common generalization of Poisson and pre-symplectic structures and also was designed to include constrained systems, including constraints induced by degenerate Lagrangians, as was investigated in [23], which is the reason for the name. It is useful to recall that Dirac was concerned with degenerate Lagrangians, so that the image $P \subset T^*Q$ of the Legendre transformation, called the set of *primary constraints* in the language of Dirac, need not be the whole space. As we have mentioned, electric circuits involve *both* primary constraints as well as constraints coming from Kirchhoff's laws.

Interestingly, despite the fact that Dirac's original idea started off with degenerate Lagrangians and a Lagrangian viewpoint, Dirac structures have not been widely studied in the context of Lagrangian systems, nor in the context of variational principles. In Part II, we will demonstrate how Dirac structures are incorporated into Lagrangian systems in the context of variational structures, and we hope that the present paper helps to fill the gap.

In studying implicit systems, one usually starts with a configuration manifold Q, and for interconnected systems, a collection of such manifolds. Of course Lagrangian mechanics normally deals with the tangent bundle TQ and Hamiltonian mechanics with the cotangent bundle T^*Q. A key point of our development will be making use of what we call the Pontryagin bundle (because of its fundamental role in Pontryagin's maximum principle), namely the fiber product (or Whitney) bundle $TQ \oplus T^*Q$, as well as the iterated tangent and cotangent spaces TT^*Q, T^*TQ, and T^*T^*Q and the relations between these spaces. To the best of our knowledge, the Pontryagin bundle was first investigated in [41] to aid in the study of degenerate Lagrangian systems, which is a case we also treat in the present paper. An important ingredient for studying the iterated tangent bundle and its relation to the Pontryagin bundle is a fundamental diffeomorphism $\kappa_Q : TT^*Q \to T^*TQ$, which is used in the theory of the generalized Legendre transform developed in [43]. Interestingly, Courant [20] started the investigation of iterated spaces in his work on tangent Dirac structures. The relation between these iterated spaces and the Pontryagin bundle was also discussed in [12].

One of the topics that we plan to pursue is reduction theory for implicit Lagrangian systems from the point of view of Lagrangian reduction (see, for example, [13]). On the Hamiltonian side, this theory has been developed in [46] and [5] and in the singular case in [6].

Other recent developments that might be relevant to future directions include the study of critical manifolds and the stability of nonholonomic constrained systems. One of the developments in the context of Hamiltonian systems has been investigated from the viewpoint of generalized Dirac constraints in [16], while, on the Lagrangian side, a geometric description of vakonomic and nonholonomic dynamics was given on the Pontryagin bundle in [18] including the case of L–C circuits. Lagrangian and Hamiltonian models of L–C circuits in conjunction with Dirac structures were studied in the context of Pontryagin's maximum principle in [34]. Both of these papers, however, along with [33], treat L–C circuits in a special way that avoids the degeneracy of the Lagrangian, but which also is, in our opinion, at odds with the standard approaches in the circuit literature and which also makes the coupling to other mechanical systems non-obvious. Our treatment of L–C circuits considers the KCL and KVL constraints as well as the primary constraints due to the degeneracy of the Lagrangian in a more standard and transparent way consistent with standard approaches in mechanics.

Another very interesting direction is the control of implicit Lagrangian systems; in particular, it would be very nice to have an implicit analog of the general theorem of [14], which asserts the equivalence of controlled Hamiltonian and controlled Lagrangian systems (including passivity, structure modification, gyroscopic forces, etc.). Reduction theory for these systems was developed in [15].

1.2. Goals and purpose of this paper

In the present paper, a framework for implicit Lagrangian systems based on variational principles, Dirac structures and the use of the spaces $TQ \oplus T^*Q$, TT^*Q, T^*TQ, and T^*T^*Q, is developed.

Our definition of implicit Lagrangian systems is based directly on the Dirac structure on T^*Q that is naturally induced by a constraint distribution Δ_Q on Q; the definition also involves a natural symplectomorphism $\gamma_Q : T^*TQ \to T^*T^*Q$ to intertwine the differential of the Lagrangian with the Dirac structure.

One of our main results is to provide a basic link between implicit Lagrangian systems and variational principles. We also show, via the example of nonholonomic systems, that the formalism captures systems with velocity constraints and, via the example of electric circuits, that this theory is very suitable for the study of interconnected systems as well as for systems with degenerate Lagrangians.

1.3. Outline of the paper

The paper is constructed as follows. In Part I, we develop the framework of implicit Lagrangian systems associated with Dirac structures induced by constraint distributions. In Section 2, Dirac structures are reviewed and, in particular, an important construction of a Dirac structure on T^*Q induced by a constraint on Q is introduced. In Section 3, basic examples of Dirac structures are presented in conjunction with Kirchhoff's current laws and voltage laws, interconnections, and nonholonomic constraints. The geometry of the iterated tangent and cotangent spaces TT^*Q, T^*TQ and T^*T^*Q, and natural diffeomorphisms between them, as well as some geometry of the Pontryagin bundle $TQ \oplus T^*Q$, is reviewed in Section 4. In Section 5, we study Dirac structures induced by constraint sets. The main goal of Section 6 is to define implicit Lagrangian systems in terms of induced Dirac structures. In Section 7, the vertical rolling disk on a plane, a basic example of a nonholonomic system and an L–C circuit that is a typical degenerate Lagrangian system with holonomic constraints are developed in the context of implicit Lagrangian systems. In Section 8, concluding remarks are given.

In Part II, we link variational principles with Dirac structures; we present an extended variational principle of Hamilton called the Hamilton–Pontryagin principle to formulate implicit Lagrangian systems in relation to the symplectic structures on TT^*Q, T^*TQ and T^*T^*Q. Furthermore, we develop an extended Lagrange–d'Alembert principle and it is shown how implicit Lagrangian systems can be formulated in the variational context. We also demonstrate the constrained implicit Lagrangian systems, where we illustrate how the constrained Dirac structure on the constraint momentum space can be constructed by using an Ehresmann connection.

2. Dirac structures

This section briefly reviews Dirac structures on vector spaces and manifolds, following [21] and [19]. Following this, we shall recall the construction of some Dirac structures which will be important for the definition of an implicit Lagrangian system. We first consider the Dirac structure on a vector space V that is constructed out of a subspace (that will later be a constraint distribution) $\Delta \subset V$ and its annihilator $\Delta^\circ \subset V^*$. Following this, we investigate the Dirac structure on a manifold M constructed from a distribution Δ_M and a two-form Ω on M.

2.1. Dirac structures on vector spaces

Let V be an n-dimensional vector space, V^* be its dual space, and let $\langle\cdot,\cdot\rangle$ be the natural pairing between V^* and V. Define the symmetric pairing $\langle\langle\cdot,\cdot\rangle\rangle$ on $V \oplus V^*$ by

$$\langle\langle(v,\alpha),(\bar{v},\bar{\alpha})\rangle\rangle = \langle\alpha,\bar{v}\rangle + \langle\bar{\alpha},v\rangle,$$

for $(v,\alpha),(\bar{v},\bar{\alpha}) \in V \oplus V^*$. A *Dirac structure* on V is a subspace $D \subset V \oplus V^*$ such that $D = D^\perp$, where $D^\perp$ is the orthogonal of D relative to the pairing $\langle\langle\cdot,\cdot\rangle\rangle$.

One can easily check that a vector subspace $D \subset V \oplus V^*$ is a Dirac structure on V if and only if $\dim D = n$ and $\langle\alpha,\bar{v}\rangle + \langle\bar{\alpha},v\rangle = 0$ for all $(v,\alpha),(\bar{v},\bar{\alpha}) \in D$.

By the definition of a Dirac structure, note that the condition $D = D^\perp$ implies that for each $(v,\alpha) \in D$, we have

$$\langle\alpha, v\rangle = 0 \tag{2.1}$$

(just choose $\bar{v} = v$ and $\bar{\alpha} = \alpha$ in the definition). Even this elementary remark has interesting links with applications. For instance, in the context of electric circuits, it gives *Tellegen's theorem*, which we shall see in the next section.

2.2. Construction of Dirac structures

We begin by reviewing some basic constructions for Dirac structures on vector spaces that will be useful in the examples in the next section. The constructions will be generalized to the context of manifolds in the next paragraph. We include a few proofs so that our exposition is reasonably self-contained.

We now recall a standard fact about annihilators that is useful in examples. First of all, for a vector space V and a subspace $\Delta \subset V$, the *annihilator* $\Delta^\circ \subset V^*$ of Δ is the subspace defined by

$$\Delta^\circ = \left\{\alpha \in V^* \mid \langle\alpha, v\rangle = 0 \text{ for all } v \in \Delta\right\}.$$

Proposition 2.1. *Suppose $\phi : V \to W$ is a linear surjective map of the vector space V to the vector space W, and $\Delta = \operatorname{Ker}\phi$. Then $\Delta^\circ = \operatorname{Im}\phi^T$, where $\phi^T : W^* \to V^*$ is the transpose (that is, the dual) of ϕ.*

Proof. We recall the following standard argument from linear algebra for the reader's convenience. Let $\alpha = \phi^T\mu$, where $\mu \in W^*$ and $\alpha \in V^*$. Then, for $v \in \Delta$,

$$\langle\alpha, v\rangle = \langle\phi^T\mu, v\rangle = \langle\mu, \phi(v)\rangle = 0.$$

Therefore, $\operatorname{Im}\phi^T \subset \Delta^\circ$. Conversely, let $\alpha \in \Delta^\circ$ and let $V = \Delta \oplus U$. Then, for $v \in V$, one may define $v = r + u$ such that $r \in \Delta$ and $u \in U$. Note that $\phi|U : U \to W$ is an isomorphism and hence $(\phi|U)^T : W^* \to U^*$ is also an isomorphism. Then,

$$\begin{aligned}\alpha(v) &= \alpha(r+u) = \langle\alpha, u\rangle = \langle\alpha|U, u\rangle = \langle(\phi|U)^T\mu, u\rangle = \langle\mu, (\phi|U)(u)\rangle \\ &= \langle\mu, \phi(u)\rangle = \langle\mu, \phi(r+u)\rangle = \langle\mu, \phi(v)\rangle = \langle\phi^T\mu, v\rangle.\end{aligned}$$

Since this holds for all $v \in V$, we get $\alpha = \phi^T\mu$. □

Given a subspace $\Delta \subset V$, and applying this proposition to the map $\pi : V^* \to \Delta^*$ that is the dual of the inclusion map $i : \Delta \to V$, and noticing that $\operatorname{Ker}\pi = \Delta^\circ$ (so in the preceding proposition, replace V by V^*, ϕ by π and Δ by Δ°) shows the well-known fact that $(\Delta^\circ)^\circ = \Delta$, which we will need in the following proof.

Proposition 2.2. *Let V be a vector space and $\Delta \subset V$ a subspace. Then $D_V = \Delta \times \Delta^\circ$ is a Dirac structure on V.*

Proof. By definition,

$$D_V^{\perp} = \{(w, \beta) \in V \times V^* \mid \langle \alpha, w \rangle + \langle \beta, v \rangle = 0 \text{ for all } v \in \Delta, \alpha \in \Delta^\circ\}.$$

We need to check that $D_V = D_V^{\perp}$. First, we check $D_V^{\perp} \subset D_V$. Let $(w, \beta) \in D_V^{\perp}$ and in the definition of $D_V^{\perp}$, first choose $\alpha = 0$. This implies that $\langle \beta, v \rangle = 0$ for arbitrary $v \in \Delta$. Hence, $\beta \in \Delta^\circ$. Similarly, let $v = 0$, then $\langle \alpha, w \rangle = 0$ for arbitrary $\alpha \in \Delta^\circ$. Therefore, $w \in (\Delta^\circ)^\circ = \Delta$. Therefore, if $(w, \beta) \in D_V^{\perp}$, then $(w, \beta) \in D_V$. Thus, $D_V^{\perp} \subset D_V$. Conversely, $D_V \subset D_V^{\perp}$ is trivial because if $(w, \beta) \in D_V$, then $\langle \alpha, w \rangle = 0$ for any $\alpha \in \Delta^\circ$, and $\langle \beta, v \rangle = 0$ for any $v \in \Delta$. Thus, $D_V = D_V^{\perp}$. □

2.3. Dirac structures on manifolds

Let M be a smooth differentiable manifold whose tangent bundle is denoted as TM and whose cotangent bundle is denoted as T^*M. Let $TM \oplus T^*M$ denote the Whitney sum bundle over M; that is, it is the bundle over the base M and with fiber over the point $x \in M$ equal to $T_xM \times T_x^*M$. An *almost Dirac structure* on M is a subbundle $D \subset TM \oplus T^*M$ that is a Dirac structure in the sense of vector spaces at each point $x \in M$.

As we shall see shortly, in geometric mechanics, almost Dirac structures provide a simultaneous generalization of both two-forms (not necessarily closed, and possibly degenerate) as well as almost Poisson structures (that is brackets that need not satisfy the Jacobi identity). A *Dirac structure*, which corresponds in geometric mechanics to assuming the two-form is closed or to assuming Jacobi's identity for the Poisson tensor, is one that satisfies

$$\langle \pounds_{X_1} \alpha_2, X_3 \rangle + \langle \pounds_{X_2} \alpha_3, X_1 \rangle + \langle \pounds_{X_3} \alpha_1, X_2 \rangle = 0,$$

for all pairs of vector fields and one-forms (X_1, α_1), (X_2, α_2), (X_3, α_3) that take values in D and where $\pounds_X$ denotes the Lie derivative along the vector field X on M.

In this paper, we will mostly be concerned with almost Dirac structures and hence for brevity, unless there is danger of confusion, we will just say "Dirac structures" even for the case of almost Dirac structures.

Theorem 2.3. *Let M be a manifold and let Ω be a two-form on M. Given a distribution*[2] *Δ_M on M, define the skew-symmetric bilinear form Ω_{Δ_M} on Δ_M by restricting Ω to Δ_M. For each $x \in M$, let*

$$D_M(x) = \{(v_x, \alpha_x) \in T_xM \times T_x^*M \mid v_x \in \Delta_M(x), \textit{ and } \alpha_x(w_x) = \Omega_{\Delta_M}(x)(v_x, w_x) \textit{ for all } w_x \in \Delta_M(x)\}.$$

Then $D_M \subset TM \oplus T^*M$ is a Dirac structure on M.

Proof. The orthogonal of $D_M \subset TM \oplus T^*M$ is given at $x \in M$ by

$$\begin{aligned} D_M^{\perp}(x) = \{&(u_x, \beta_x) \in T_xM \times T_x^*M \mid \alpha_x(u_x) + \beta_x(v_x) = 0 \text{ for all } v_x \in \Delta_M(x), \\ &\text{and } \alpha_x(w_x) = \Omega_{\Delta_M}(x)(v_x, w_x) \text{ for all } w_x \in \Delta_M(x)\}. \end{aligned}$$

Let us first check that $D_M(x) \subset D_M^{\perp}(x)$. To do this, let $(v_x, \alpha_x) \in D_M(x)$ and $(v_x', \alpha_x') \in D_M(x)$. Then,

$$\alpha_x(v_x') + \alpha_x'(v_x) = \Omega_{\Delta_M}(x)(v_x, v_x') + \Omega_{\Delta_M}(x)(v_x', v_x) = 0,$$

by skew-symmetry of Ω and hence of Ω_{Δ_M}; therefore, $D_M(x) \subset D_M^{\perp}(x)$.

Next, let us show that $D_M^{\perp}(x) \subset D_M(x)$. Let $(u_x, \beta_x) \in D_M^{\perp}(x)$. By definition of $D_M^{\perp}$, we have

$$\alpha_x(u_x) + \beta_x(v_x) = 0$$

for all $(v_x, \alpha_x) \in T_xM \times T_x^*M$ such that $v_x \in \Delta_M(x)$ and $\alpha_x(w_x) = \Omega_{\Delta_M}(x)(v_x, w_x)$ for all $w_x \in \Delta_M(x)$. First, choose $v_x = 0$, and $\alpha_x \in \Delta_M^\circ(x)$, so that $(v_x, \alpha_x) \in D_M(x)$. We conclude that $\alpha_x(u_x) = 0$ for all $\alpha_x \in \Delta_M^\circ(x)$, and hence (since the annihilator of $\Delta_M^\circ(x)$ is $\Delta_M(x)$) that $u_x \in \Delta_M(x)$. Second, let $v_x \in \Delta_M(x)$ be arbitrary and suppose that $\alpha_x(w_x) = \Omega_{\Delta_M}(x)(v_x, w_x)$ for all $w_x \in \Delta_M(x)$. Since we have already proved that $u_x \in \Delta_M(x)$, we get $\alpha_x(u_x) = \Omega_{\Delta_M}(x)(v_x, u_x)$ and so from $\alpha_x(u_x) + \beta_x(v_x) = 0$, we get

$$\Omega_{\Delta_M}(x)(v_x, u_x) + \beta_x(v_x) = 0 \quad \text{for all } v_x \in \Delta_M(x).$$

2 In some examples, for instance in nonholonomic systems, the distribution Δ_M is the constraint distribution (see, for instance [8]); in circuits, this distribution corresponds to Kirchhoff's current law, as we shall explain later.

In other words, $\beta_x(v_x) = \Omega_{\Delta_M}(x)(u_x, v_x)$ for all $v_x \in \Delta_M(x)$. Therefore, $(u_x, \beta_x) \in D_M(x)$. Thus, $D_M^{\perp}(x) \subset D_M(x)$.

Since we have proved both inclusions, we get $D_M^{\perp}(x) = D_M(x)$ and so indeed, D_M is a Dirac structure on M. □

Example. Let $\Omega = 0$, so that also $\Omega_{\Delta_M} = 0$. Then, $\Delta_M \times \Delta_M^{\circ} \subset TM \oplus T^*M$ is a Dirac structure on M.

Example. Let M be a symplectic manifold with a two-form $\Omega : TM \times TM \to \mathbb{R}$. In the case $\Delta_M = TM$, the Dirac structure defined by Δ_M and Ω_{Δ_M} according to the preceding Theorem is simply the graph of the symplectic structure, thought of as a map of TM to T^*M.

The preceding theorem has the following dual, which is proved in the same way.

Theorem 2.4. *Let M be a manifold and let $B : T^*M \times T^*M \to \mathbb{R}$ be a skew-symmetric two-tensor (such as a Poisson tensor). Given a codistribution $\Delta_M^* \subset T^*M$ on M, define the skew-symmetric two-tensor $B_{\Delta_M^*}$ on Δ_M^* by restricting B to Δ_M^*. For each $x \in M$, let*

$$D_M(x) = \{(v_x, \alpha_x) \in T_xM \times T_x^*M \mid \alpha_x \in \Delta_M^*(x), \textit{ and } \beta_x(v_x) = B_{\Delta_M^*}(x)(\alpha_x, \beta_x) \textit{ for all } \beta_x \in \Delta_M^*(x)\}.$$

Then $D_M \subset TM \oplus T^*M$ is a Dirac structure on M.

Example. Let M be a Poisson manifold with the Poisson structure $B : T^*M \times T^*M \to \mathbb{R}$. If $\Delta_M^* = T^*M$, then the Dirac structure defined by Δ_M^* and B_M according to the preceding theorem is the graph of the Poisson structure thought of as a map of T^*M to TM.

In the case that $M = T^*Q$ and $\Delta_M = TT^*Q$, let us call the Dirac structure on the manifold T^*Q the *canonical Dirac structure*. One gets the same structure if one uses the Poisson structure and the codistribution $\Delta_M^* = T^*T^*Q$.

Notice that the previous construction of the Dirac structure D_V on the vector space V given in Proposition 2.2 may be regarded as an instance of this more general construction of the Dirac structure D_M on the manifold M, where the two-form Ω (or the form B) is chosen to be zero.

3. Basic examples of Dirac structures

In this section, two basic constructions of induced Dirac structures on vector spaces will be given; these are important for the examples given later in the paper as well as providing insight into the corresponding construction on manifolds.

The first example is a Dirac structure on a vector space motivated by the type of structure encountered in Kirchhoff's current laws for electric circuits. The second is a Dirac structure relevant to the interconnection of systems. These examples are also fundamental for understanding Dirac structures of the sort that appear in mechanical systems with nonholonomic constraints. As will be illustrated in Section 5, these constructions produce Dirac structures on a cotangent bundle T^*Q that is induced from a given constraint distribution Δ_Q on Q as well as Dirac structures induced on submanifolds of Dirac manifolds, which will be important for the consideration of degenerate Lagrangian systems, which correspondingly have primary constraints.

3.1. Example: Kirchhoff's current and voltage laws

Consider a circuit, which is given by a nonenergic n-port with a given circuit topology specified by a graph $\mathcal{G}$.[3] In this case, the configuration manifold Q is an n-dimensional vector space E and we use the notation $q = (q^1, \ldots, q^n) \in E$, where q^k denotes the charge associated with the kth branch.

Given a circuit topology $\mathcal{G}$, Kirchhoff's current law (KCL) may be written in terms of a fixed collection of covectors (or one-forms) ω^a as

$$\langle \omega^a, f \rangle = 0, \quad a = 1, \ldots, m.$$

[3] For the fundamentals of circuit theory, see, for example, [17].

Here, $f = (f^1, \ldots, f^n) \in E$ denote branch currents in the circuit. Write the covectors ω^a as

$$\omega^a = \omega^a_k dq^k, \quad a = 1, \ldots, m; k = 1, \ldots, n,$$

where $m < n$; the coefficients ω^a_k are ± 1 and 0 are determined by the circuit topology.

The set of all branch currents $f = (f^1, \ldots, f^n)$ that satisfy the KCL forms an $(n-m)$-dimensional subspace Δ_q of E (which is identified with the tangent space $T_q E$ to E at $q \in E$) defined by

$$\Delta_q = \left\{ f \in T_q E \mid \left\langle \omega^a, f \right\rangle = 0, \quad a = 1, \ldots, m \right\}.$$

We call $\Delta \subset TE$ the *constraint KCL space*. Its annihilator is defined, for each $q \in E$, as

$$\Delta^\circ_q = \{ e \in T^*_q E \mid \langle e, f \rangle = 0 \text{ for all } f \in \Delta_q \}.$$

In coordinates, any $e \in \Delta^\circ_q$ can be written, by using Lagrange multipliers μ_a, in the form

$$e_k = \mu_a \omega^a_k, \quad a = 1, \ldots, m; k = 1, \ldots, n.$$

We identify $\Delta^\circ \subset T^*E$ with the set of *branch voltages* $e = (e_1, \ldots, e_n)$. It annihilates the KCL space and is an m-dimensional subspace of $T^*_q E$ at each $q \in E$ which is equivalent to Kirchhoff's voltage law. Accordingly, $\Delta^\circ \subset T^*E$ is called the *constraint KVL space*. From Proposition 2.2, we get

Theorem 3.1. *Let $\Delta \subset TE$ be the constraint KCL space and $\Delta^\circ \subset T^*E$ be the constraint KVL space. Then,*

$$D_E = \Delta \times \Delta^\circ \subset TE \times T^*E$$

is a Dirac structure on E.

Using Eq. (2.1), we get the following corollary, a simple but important result known as *Tellegen's theorem*.

Corollary 3.2. *For each $(f, e) \in D_E$, we have $\langle e, f \rangle = 0$.*

3.2. Generalization

Next we give an induced Dirac structure on T^*E, which will be employed in our development of L–C circuits in Section 7.

Theorem 3.3. *Let E be a vector space and E^* be the dual space of E, and hence we can regard $T^*E \cong E \times E^*$. Given $\Delta \subset TE$ and $\Delta^\circ \subset T^*E$, define $D_\Delta \subset TT^*E \times T^*T^*E$ by, for each $(q, p) \in E \times E^*$,*

$$D_\Delta(q, p) = \{((v, \alpha), (-\beta, v)) \mid v \in \Delta_q, \alpha - \beta \in \Delta^\circ_q\}.$$

Then, D_Δ is a Dirac structure on T^*E.

This is a special case of Theorem 2.3 with the choice $M = E \times E^*$ and Ω the standard symplectic structure. This will be generalized to cotangent bundles in Proposition 5.1.

3.3. Example: Interconnection

Another important class of examples of Dirac structures is the *interconnection Dirac structures*, which were developed (using different language of course) in [30] in the context of power invariance under interconnecting electromechanical machines and networks. For the study of such interconnected and nonholonomic systems, another construction will be useful.

Theorem 3.4. *Let E be a vector space, E^* be the dual space of E and $T^*E \cong E \times E^*$. Given $\Delta^* \subset T^*E$, let $(\Delta^*)^\circ \subset TE$ be its annihilator and define $D_\Delta \subset TT^*E \times T^*T^*E$ by, for each $(q, p) \in E \times E^*$,*

$$D_\Delta(q, p) = \{((v, \alpha), (-\alpha, w)) \mid \alpha \in \Delta^*_q, v - w \in (\Delta^*_q)^\circ\}.$$

Then, D_Δ is a Dirac structure on T^*E.

This result is a special instance of the general construction given in Theorem 2.4 with $M = E \times E^* = T^*E$ and with the standard canonical Poisson tensor on T^*E. The reader may regard this special case as a useful precursor of the result given later in Proposition 5.1.

Example. Let $T^*\mathbb{R}^3 = \mathbb{R}^3 \times (\mathbb{R}^3)^*$. Let $D_\Delta \subset TT^*\mathbb{R}^3 \times T^*T^*\mathbb{R}^3$ be the Dirac structure on $T^*\mathbb{R}^3$ defined by, for each $(q, p) \in T^*\mathbb{R}^3$,

$$D_\Delta(q, p) = \{((v, F), (-F, w)) \in T_{(q,p)}T^*\mathbb{R}^3 \times T^*_{(q,p)}T^*\mathbb{R}^3 \mid F \in \Delta^*_q, v - w \in (\Delta^*_q)^\circ\},$$

where $\Delta^* \subset T^*\mathbb{R}^3$ is a given codistribution and $(\Delta^*)^\circ \subset T\mathbb{R}^3$ its annihilator. Let us set $((\hat{v}, F), (\hat{F}, v)) \in D_\Delta(q, p)$, then $F = -\hat{F} \in \Delta^*_q$ and $\hat{v} - v \in (\Delta^*_q)^\circ$. Because $D_\Delta = D_\Delta^\perp$,

$$\langle\langle((\hat{v}, F), (\hat{F}, v)), ((\hat{v}, F), (\hat{F}, v))\rangle\rangle = 2\langle F, v\rangle + 2\langle \hat{F}, \hat{v}\rangle = 2\langle \hat{F}, \hat{v} - v\rangle = 0.$$

Then, we have

$$\langle F, v\rangle + \langle \hat{F}, \hat{v}\rangle = 0,$$

which may be interpreted as a condition related to the interconnection of two subsystems. One may regard the above equation as having the physical interpretation of *power invariance* interrelating two subsystems. In mechanics, a simple example of such an interconnection is given by *Newton's law of action and reaction*, in which case one may consider F and $\hat{F}$ as action and reaction forces associated with adjacent bodies, while, by setting $\hat{v} - v = 0$, one gets $v = \hat{v}$, which is interpreted as the common velocity of those bodies at their point of contact.

4. Iterated tangent and cotangent bundles

In this section we recall some basic geometry of the spaces TT^*Q, T^*T^*Q and T^*TQ, as well as the Pontryagin bundle $TQ \oplus T^*Q$. Those spaces are very important for the understanding of the interrelation between Lagrangian mechanics on the tangent bundle TQ and Hamiltonian mechanics on the cotangent bundle T^*Q. In particular, there are two diffeomorphisms between T^*TQ, TT^*Q and T^*T^*Q that were studied in [43] in the context of the generalized Legendre transform that will be useful in later sections.

4.1. *Diffeomorphism between TT^*Q and T^*TQ*

Given a manifold Q, we are going to define a natural diffeomorphism

$$\kappa_Q : TT^*Q \to T^*TQ.$$

In a local trivialization, Q is represented by an open set U in a linear space E, so that TT^*Q is represented by $(U \times E^*) \times (E \times E^*)$, while T^*TQ is locally given by $(U \times E) \times (E^* \times E^*)$. In this local representation, the map κ_Q will be given by

$$(q, p, \delta q, \delta p) \mapsto (q, \delta q, \delta p, p),$$

where (q, p) are local coordinates of T^*Q and $(q, p, \delta q, \delta p)$ are the corresponding coordinates of TT^*Q, while $(q, \delta q, \delta p, p)$ are the local coordinates of T^*TQ induced by κ_Q. The map κ_Q is the unique map that intertwines two sets of maps given as follows.

4.2. *The first set of maps*

Consider the following two maps:

$$T\pi_Q : TT^*Q \to TQ, \quad \pi_{TQ} : T^*TQ \to TQ,$$

which are the obvious maps and recall that $\pi_Q : T^*Q \to Q$ is the cotangent projection. The first commutation condition that will be used to define κ_Q is that

$$\pi_{TQ} \circ \kappa_Q = T\pi_Q. \tag{4.1}$$

4.3. The second set of maps

The second set of maps is the following:

$$\tau_{T^*Q} : TT^*Q \to T^*Q, \quad \pi^1 : T^*TQ \to T^*Q.$$

Note that $\tau_Q : TQ \to Q$ is the tangent bundle projection. Thus, the second commutation condition is

$$\pi^1 \circ \kappa_Q = \tau_{T^*Q}. \tag{4.2}$$

To explain the map π^1, let $\alpha_{v_q} \in T^*_{v_q}TQ$ and let $u_q \in T_qQ$. Then

$$\langle \pi^1(\alpha_{v_q}), u_q \rangle = \langle \alpha_{v_q}, \mathrm{ver}(u_q, v_q) \rangle,$$

where

$$\mathrm{ver}(u_q, v_q) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (v_q + tu_q) \in T_{v_q}TQ$$

denotes the vertical lift of u_q along v_q.

4.4. Local representation

In a natural local trivialization, these four maps are readily checked to be given by

$$\begin{aligned}
&T\pi_Q(q, p, \delta q, \delta p) = (q, \delta q),\\
&\pi_{TQ}(q, \delta q, \delta p, p) = (q, \delta q),\\
&\tau_{T^*Q}(q, p, \delta q, \delta p) = (q, p),\\
&\pi^1(q, \delta q, \delta p, p) = (q, p),
\end{aligned}$$

from which it is easy to check that the commutation conditions are satisfied and it is clear that this uniquely characterizes the map κ_Q.

Proposition 4.1. *For any manifold Q, there is a unique diffeomorphism*

$$\kappa_Q : TT^*Q \to T^*TQ$$

such that the two commutation conditions (4.1) and (4.2) are satisfied.

4.5. Diffeomorphism between T^*T^*Q and TT^*Q

Let Ω be the canonical symplectic form on the cotangent bundle T^*Q. There is a natural diffeomorphism given by the associated flat map

$$\Omega^\flat : TT^*Q \to T^*T^*Q.$$

If E is the model space of Q and U is an open set of E, then T^*T^*Q is represented by $(U \times E^*) \times (E^* \times E)$, while TT^*Q is represented by $(U \times E^*) \times (E \times E^*)$. The map $\Omega^\flat$ is locally represented by

$$(q, p, \delta q, \delta p) \mapsto (q, p, -\delta p, \delta q).$$

As above, $\Omega^\flat$ is the unique map that also intertwines two sets of maps.

4.6. The first set of maps

Consider the following maps:

$$\tau_{T^*Q} : TT^*Q \to T^*Q, \quad \pi_{T^*Q} : T^*T^*Q \to T^*Q,$$

which are obvious maps and recall that $\tau_Q : TQ \to Q$ and $\pi_Q : T^*Q \to Q$ are respectively the tangent projection and the cotangent projection. The first commutation condition is

$$\pi_{T^*Q} \circ \Omega^\flat = \tau_{T^*Q}.$$

4.7. The second set of maps

The second set of maps is the following:

$$T\pi_Q : TT^*Q \to TQ, \quad \pi^2 : T^*T^*Q \to TQ.$$

Thus, the second commutation condition is

$$\pi^2 \circ \Omega^\flat = T\pi_Q,$$

where π^2 is defined similarly to π^1.

4.8. Local representations of maps

In a local trivialization, these four maps are given by

$$\begin{aligned}
&\pi_{T^*Q}(q, p, -\delta p, \delta q) = (q, p),\\
&\tau_{T^*Q}(q, p, \delta q, \delta p) = (q, p),\\
&\pi^2(q, p, -\delta p, \delta q) = (q, \delta q),\\
&T\pi_Q(q, p, \delta q, \delta p) = (q, \delta q),
\end{aligned}$$

from which it is easy to check that the commutation conditions are satisfied.

4.9. The symplectic form on TT^*Q

The manifold TT^*Q is a symplectic manifold with a particular symplectic form Ω_{TT^*Q} that can be represented by two distinct ways as the exterior derivative of two intrinsic but different one-forms. When the local coordinates of T^*Q and TT^*Q are respectively denoted by (q, p) and $(q, p, \delta q, \delta p)$, these two one-forms are given in local coordinates by:

$$\lambda = \delta p \mathrm{d}q + p \mathrm{d}\delta q, \quad \chi = -\delta p \mathrm{d}q + \delta q \mathrm{d}p.$$

Let Θ_{T^*TQ} denote the canonical one-form on T^*TQ and let $\Theta_{T^*T^*Q}$ denote the canonical one-form on T^*T^*Q. Then we have the following proposition:

Proposition 4.2. *The two one-forms λ and χ are given by*

$$\lambda = (\kappa_Q)^* \, \Theta_{T^*TQ}, \quad \chi = (\Omega^\flat)^* \, \Theta_{T^*T^*Q}.$$

One can easily check the proposition. Thus, the symplectic forms Ω_{TT^*Q} on TT^*Q associated with λ and χ are given by

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

Remarks. It will be useful to establish systematic notation for, and recall the definition of, the canonical one-form and two-form on the cotangent bundle T^*M of a manifold M. Namely, the canonical one-form Θ_{T^*M} on T^*M is defined by

$$\Theta_{T^*M}(\alpha)(w) = \langle \alpha, T\pi_M \cdot w \rangle,$$

where $\alpha \in T^*M$, $w \in T_\alpha T^*M$, $\pi_M : T^*M \to M$ is the cotangent projection and $T\pi_M : TT^*M \to TM$ is the tangent map of π_M. The canonical symplectic two-form is denoted as $\Omega_{T^*M} = -\mathbf{d}\Theta_{T^*M}$.

Thus, in the case $M = TQ$, we have the canonical one-form Θ_{T^*TQ}, while in the case $M = T^*Q$, we have the canonical one-form $\Theta_{T^*T^*Q}$.

4.10. Pontryagin bundle $TQ \oplus T^*Q$

Consider the bundle $TQ \oplus T^*Q$ over Q, that is, the Whitney sum of the tangent bundle and the cotangent bundle over Q, whose fiber at $q \in Q$ is the product $T_qQ \times T_q^*Q$. Let us call $TQ \oplus T^*Q$ the *Pontryagin bundle*. Using a model space E for Q and a chart domain, which is an open set $U \subset E$, then $TQ \times T^*Q$ is locally denoted by $U \times E \times U \times E^*$ and $TQ \oplus T^*Q$ by $U \times E \times E^*$. In this local trivialization, the local coordinates of $TQ \oplus T^*Q$ are written as

$$(q, v, p) \in U \times E \times E^*,$$

from which the following three projections are naturally defined:

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

We now define an embedding $\rho_{T^*TQ} : T^*TQ \to TQ \oplus T^*Q$ which will have the property that $\mathrm{pr}_{TQ} \circ \rho_{T^*TQ} = \pi_{TQ}$. This map is defined intrinsically to be the direct sum of the two maps $\pi_{TQ} : T^*TQ \to TQ$ and $\tau_{T^*Q} \circ \kappa_Q^{-1} : T^*TQ \to T^*Q$ and is easily checked to be locally represented by

$$\rho_{T^*TQ} : T^*TQ \to TQ \oplus T^*Q; \quad (q, \delta q, \delta p, p) \mapsto (q, \delta q, p).$$

Note that this map is a "projection" in the sense that it simply eliminates δp. We likewise define the maps

$$\begin{aligned}
&\rho_{TT^*Q} = \rho_{T^*TQ} \circ \kappa_Q : TT^*Q \to TQ \oplus T^*Q; \quad (q, p, \delta q, \delta p) \mapsto (q, \delta q, p)\\
&\rho_{T^*T^*Q} = \rho_{TT^*Q} \circ (\Omega^\flat)^{-1} : T^*T^*Q \to TQ \oplus T^*Q; \quad (q, p, -\delta p, \delta q) \mapsto (q, \delta q, p),
\end{aligned}$$

each of which also eliminates δp.

5. Induced Dirac structures

In this section, we introduce the notion of an *induced Dirac structure* on the cotangent bundle T^*Q of a configuration manifold Q. This Dirac structure is induced from a given distribution $\Delta_Q \subset TQ$ on Q, which will play an essential role in the definition of implicit Lagrangian systems in the context of Dirac structures.

Let Q be a smooth manifold and let TQ be the tangent bundle of Q and T^*Q the cotangent bundle. Suppose that one is given a regular distribution Δ_Q on Q (that is, a smooth vector subbundle of the tangent bundle TQ) and denote the annihilator of Δ_Q by Δ_Q°. Let $\pi_Q : T^*Q \to Q$ be the canonical projection so that its tangent is a map $T\pi_Q : TT^*Q \to TQ$. Lift the distribution Δ_Q on Q, defining the distribution on T^*Q by

$$\Delta_{T^*Q} = (T\pi_Q)^{-1}(\Delta_Q) \subset TT^*Q. \tag{5.1}$$

Let Ω be the canonical two-form on T^*Q and define a skew-symmetric bilinear form Ω_{Δ_Q} by restricting Ω to Δ_{T^*Q}, that is, $\Omega_{\Delta_Q} = \Omega|_{\Delta_{T^*Q} \times \Delta_{T^*Q}}$. Then, we have the following proposition.

Proposition 5.1. *Let $\Delta_Q \subset TQ$ be a given (regular for simplicity) distribution on Q, and let $\Delta_{T^*Q} = (T\pi_Q)^{-1}(\Delta_Q)$ be the induced distribution on T^*Q. Then, 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, whose fiber is given for each $z \in T^*Q$ as*

$$\begin{aligned}
D_{\Delta_Q}(z) = \{(v_z, \alpha_z) \in T_zT^*Q \times T_z^*T^*Q \mid\ & v_z \in \Delta_{T^*Q}(z), \text{ and}\\
& \alpha_z(w_z) = \Omega_{\Delta_Q}(z)(v_z, w_z) \text{ for all } w_z \in \Delta_{T^*Q}(z)\}.
\end{aligned} \tag{5.2}$$

While this result is a special case of Theorem 2.3, we give the proof from scratch for the convenience of the reader, as this is the situation of direct relevance to this paper.

Proof. The orthogonal of $D_{\Delta_Q} \subset TT^*Q \oplus T^*T^*Q$ at the point $z \in T^*Q$ is given by

$$D^{\perp}_{\Delta_Q}(z) = \{(u_z, \beta_z) \in T_zT^*Q \times T^*_zT^*Q \mid \alpha_z(u_z) + \beta_z(v_z) = 0 \textit{ for all } v_z \in \Delta_{T^*Q}(z), \text{ and } \alpha_z(w_z) = \Omega_{\Delta_Q}(z)(v_z, w_z) \text{ for all } w_z \in \Delta_{T^*Q}(z)\}.$$

Let us first show that $D_{\Delta_Q}(z) \subset D^{\perp}_{\Delta_Q}(z)$. Let $(v_z, \alpha_z) \in D_{\Delta_Q}(z)$ and $(v'_z, \alpha'_z) \in D_{\Delta_Q}(z)$. Then,

$$\alpha_z(v'_z) + \alpha'_z(v_z) = \Omega_{\Delta_Q}(z)(v_z, v'_z) + \Omega_{\Delta_Q}(z)(v'_z, v_z) = 0.$$

Therefore, $D_{\Delta_Q}(z) \subset D^{\perp}_{\Delta_Q}(z)$.

Next, we shall check $D^{\perp}_{\Delta_Q}(z) \subset D_{\Delta_Q}(z)$. Let $(u_z, \beta_z) \in D^{\perp}_{\Delta_Q}(z)$. By definition of $D^{\perp}_{\Delta_Q}$, we have $\alpha_z(u_z) + \beta_z(v_z) = 0$ for all $v_z \in \Delta_{T^*Q}(z)$ and $\alpha_z(w_z) = \Omega_{\Delta_Q}(z)(v_z, w_z)$ for all $w_z \in \Delta_{T^*Q}(z)$. First of all, choose $v_z = 0$, and $\alpha_z \in \Delta^{\circ}_{T^*Q}(z)$, the annihilator of $\Delta_{T^*Q}(z)$; that is, $\alpha_z(w_z) = 0$ for all $w_z \in \Delta_{T^*Q}(z)$. For any $\alpha_z \in \Delta^{\circ}_{T^*Q}(z)$, we have $(0, \alpha_z) \in D_{\Delta_Q}$, and so as (u_z, β_z) is orthogonal to all such elements, we have $\alpha_z(u_z) = 0$ for all $\alpha_z \in \Delta^{\circ}_{T^*Q}(z)$ and so $u_z \in \Delta_{T^*Q}(z)$.

The other condition on (u_z, β_z) we need to check is that $\beta_z(v_z) = \Omega_{\Delta_Q}(z)(u_z, v_z)$ for all $v_z \in \Delta_{T^*Q}(z)$. Thus, let $v_z \in \Delta_{T^*Q}(z)$ be arbitrary and choose α_z such that $\alpha_z(w_z) = \Omega_{\Delta_Q}(z)(v_z, w_z)$ for all $w_z \in \Delta_{T^*Q}(z)$. From $\alpha_z(u_z) + \beta_z(v_z) = 0$ and the fact that we have already proved that $u_z \in \Delta_{T^*Q}(z)$, we get

$$\Omega_{\Delta_Q}(z)(v_z, u_z) + \beta_z(v_z) = 0 \quad \text{for all } v_z \in \Delta_{T^*Q}(z),$$

that is, $\beta_z(v_z) = \Omega_{\Delta_Q}(z)(u_z, v_z)$ for all $v_z \in \Delta_{T^*Q}(z)$. Thus, $(u_z, \beta_z) \in D_{\Delta_Q}(z)$ and hence $D^{\perp}_{\Delta_Q} \subset D_{\Delta_Q}$, as required. Therefore, $D^{\perp}_{\Delta_Q} = D_{\Delta_Q}$. □

5.1. Dirac structures induced from nonholonomic constraints

Let Q be a smooth manifold. Let $\Delta_Q \subset TQ$ be a regular distribution on Q and let Δ°_Q denote the annihilator of Δ_Q. Let Ω be the canonical symplectic form on T^*Q and let $\Omega^{\flat} : TT^*Q \to T^*T^*Q$ be the associated flat map. Recall that the distribution $\Delta_Q \subset TQ$ may be lifted as

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

and that its annihilator is defined, for each $z \in T^*Q$, as

$$\Delta^{\circ}_{T^*Q}(z) = \{\alpha_z \in T^*_zT^*Q \mid \langle w_z, \alpha_z\rangle = 0 \text{ for all } w_z \in \Delta_{T^*Q}(z)\}.$$

The induced Dirac structure D_{Δ_Q} on T^*Q, defined by equation Eq. (5.2), can, by using a little definition chasing, be expressed by the following proposition.

Proposition 5.2. *Given a distribution Δ_Q on Q, the induced Dirac structure D_{Δ_Q} on T^*Q has fiber given, for each $z \in T^*Q$, by*

$$D_{\Delta_Q}(z) = \{(v_z, \alpha_z) \in T_zT^*Q \times T^*_zT^*Q \mid v_z \in \Delta_{T^*Q}(z), \textit{and } \alpha_z - \Omega^{\flat}(z) \cdot v_z \in \Delta^{\circ}_{T^*Q}(z)\}. \tag{5.3}$$

This is a restatement of the preceding proposition, so we can omit the direct proof.

Remarks. Recall the construction of an induced Dirac structure on a vector space that was given in Theorem 3.3. Motivated by the vector space case, we can also describe the induced Dirac structure on T^*Q in equation (5.3) in the following way:

Consider the diffeomorphism defined by $\sigma_Q = \kappa_Q \circ (\Omega^{\flat})^{-1} : T^*T^*Q \to T^*TQ$ and let, as usual, $\pi_Q : T^*Q \to Q$ be the canonical projection. Recall that we have two projections, namely

$$T\pi_Q : TT^*Q \to TQ, \quad \text{and} \quad \pi^2 = T\pi_Q \circ (\Omega^{\flat})^{-1} : T^*T^*Q \to TQ.$$

Consider a pair of points $(v_z, \alpha_z) \in T_zT^*Q \times T^*_zT^*Q$ and also consider the points $\kappa_Q(v_z)$ and $\sigma_Q(\alpha_z)$, each of which lies in T^*TQ. First observe that $\kappa_Q(v_z)$ and $\sigma_Q(\alpha_z)$ have the same base point when $T\pi_Q(v_z) = \pi^2(\alpha_z)$; if this holds, it makes sense to subtract $\kappa_Q(v_z)$ and $\sigma_Q(\alpha_z)$.

Next, let $\Delta_{TQ} = T\tau_Q^{-1}(\Delta_Q)$ and let Δ°_{TQ} be its annihilator. A straightforward check using local representatives then shows that the induced Dirac structure D_{Δ_Q} on T^*Q is given at a point $z \in T^*_q Q$ by

$$D_{\Delta_Q}(z) = \{(v_z, \alpha_z) \in T_z T^*Q \times T^*_z T^*Q \mid T\pi_Q(v_z) = \pi^2(\alpha_z) =: w_q \in \Delta_Q(q), \\ \kappa_Q(v_z) - \sigma_Q(\alpha_z) \in \Delta^{\circ}_{TQ}(w_q)\}. \tag{5.4}$$

5.2. Local representation of the Dirac structure

To study Dirac structures locally, we choose local coordinates q^i on Q so that locally, Q is represented by an open set $U \subset \mathbb{R}^n$. The constraint set Δ_Q defines a subspace of TQ, which we denote by $\Delta(q) \subset \mathbb{R}^n$ at each point $q \in U$. If we let the dimension of the constraint space be $n - m$, then we can choose a basis $e_{m+1}(q), e_{m+2}(q), \ldots, e_n(q)$ of $\Delta(q)$.

It is also common (see, for instance, [8]) to represent constraint sets as the simultaneous kernel of a number of constraint one-forms; that is, the annihilator of $\Delta(q)$, which is denoted by $\Delta^{\circ}(q)$, is spanned by such one-forms, that we write as $\omega^1, \omega^2, \ldots, \omega^m$. Now writing the projection map $\pi_Q : T^*Q \to Q$ locally as $(q, p) \mapsto q$, its tangent map is locally given by $T\pi_Q : (q, p, \dot{q}, \dot{p}) \mapsto (q, \dot{q})$. Thus, we can locally represent Δ_{T^*Q} as

$$\Delta_{T^*Q} \cong \left\{ v_{(q,p)} = (q, p, \dot{q}, \dot{p}) \mid q \in U, \dot{q} \in \Delta(q) \right\}.$$

Letting points in T^*T^*Q be locally denoted by $\alpha_{(q,p)} = (q, p, \alpha, w)$, where α is a covector and w is a vector, notice that the annihilator of Δ_{T^*Q} is locally,

$$\Delta^{\circ}_{T^*Q} \cong \left\{ \alpha_{(q,p)} = (q, p, \alpha, w) \mid q \in U, \alpha \in \Delta^{\circ}(q) \text{ and } w = 0 \right\}.$$

Now we have, from the local expression for the symplectic form,

$$\Omega^{\flat}(z) \cdot v_z = (q, p, -\dot{p}, \dot{q})$$

and we see that the condition $\alpha_z - \Omega^{\flat}(z) \cdot v_z \in \Delta^{\circ}_{T^*Q}$ in Eq. (5.3) reads

$$(q, p, \alpha + \dot{p}, w - \dot{q}) \in \Delta^{\circ}_{T^*Q},$$

that is,

$$\alpha + \dot{p} \in \Delta^{\circ}(q), \quad \text{and} \quad w - \dot{q} = 0.$$

Thus, we get, by Proposition 5.2,

$$D_{\Delta_Q}(z) = \{(v_z, \alpha_z) \in T_z T^*Q \times T^*_z T^*Q \mid v_z \in \Delta_{T^*Q}(z), \text{and } \alpha_z - \Omega^{\flat}(z) \cdot v_z \in \Delta^{\circ}_{T^*Q}(z)\} \\ = \left\{ ((q, p, \dot{q}, \dot{p}), (q, p, \alpha, w)) \mid \dot{q} \in \Delta(q), w = \dot{q}, \text{and } \alpha + \dot{p} \in \Delta^{\circ}(q) \right\}. \tag{5.5}$$

5.3. A Poisson representation of Dirac structures

So far we have used the symplectic structure on T^*Q together with the constraint Δ_Q to define the Dirac structure D_{Δ_Q}. However, as we have already hinted at, there is a dual version of the above constructions in which one takes a Poisson point of view to arrive at a "Poisson version" of Eq. (5.5). Recall that the Poisson structure on T^*Q may be viewed as a bundle map $B^{\sharp} : T^*T^*Q \to TT^*Q$ that is usually thought of as taking the differential of a Hamiltonian and producing the corresponding Hamiltonian vector field. It is associated in a natural way with the canonical symplectic structure $\Omega : TT^*Q \times TT^*Q \to \mathbb{R}$, or its "inverse", the canonical Poisson structure $B : T^*T^*Q \times T^*T^*Q \to \mathbb{R}$ on T^*Q.

Suppose that again, we are given a constraint distribution $\Delta_Q \subset TQ$. Recall that the projection $\pi^2 : T^*T^*Q \to TQ$ is defined by $\pi^2 = T\pi_Q \circ (\Omega^{\flat})^{-1}$, and, analogously to Eq. (5.1), define the induced codistribution $\Delta^*_{T^*Q} \subset T^*T^*Q$ by

$$\Delta^*_{T^*Q} = \Omega^{\flat}(\Delta_{T^*Q}) = (\pi^2)^{-1}(\Delta_Q). \tag{5.6}$$

Recalling that

$$\Delta_{T^*Q} = (T\pi_Q)^{-1}(\Delta_Q) = \{v_{(q,p)} = (q, p, v, \alpha) | q \in U, v \in \Delta(q)\},$$

we see that the codistribution is given, using local coordinates $q \in U \subset \mathbb{R}^n$, by

$$\Delta^*_{T^*Q} = \{\alpha_{(q,p)} = (q, p, -\alpha, v) | q \in U, v \in \Delta(q)\}.$$

The annihilator of $\Delta^*_{T^*Q}$ is given, for each $z \in T^*Q$, by

$$\begin{aligned}(\Delta^*_{T^*Q})^\circ(z) &= \{w_z \in T_zT^*Q | \langle w_z, \alpha_z\rangle = 0 \text{ for all } \alpha_z \in \Delta^*_{T^*Q}(z)\} \\ &= \{w_{(q,p)} = (q, p, w, \beta) | q \in U, w = 0, \text{ and } \beta \in \Delta^\circ(q)\}.\end{aligned} \tag{5.7}$$

We summarize the situation in the following Proposition (compare with Proposition 5.2).

Proposition 5.3. *Using the above notation, the induced Dirac structure D_{Δ_Q} on T^*Q is given, for $z \in T^*Q$, by*

$$D_{\Delta_Q}(z) = \{(v_z, \alpha_z) \in T_zT^*Q \times T^*_zT^*Q \mid \alpha_z \in \Delta^*_{T^*Q}(z) \text{ and } v_z - B^\sharp(z) \cdot \alpha_z \in (\Delta^*_{T^*Q})^\circ(z)\} \tag{5.8}$$

and is given locally by

$$D_{\Delta_Q}(z) = \left\{((q, p, \dot{q}, \dot{p}), (q, p, \alpha, w)) \mid \dot{q} \in \Delta(q), w = \dot{q}, \text{ and } \alpha + \dot{p} \in \Delta^\circ(q)\right\}. \tag{5.9}$$

Of course this last local expression (5.9) agrees with (5.5).

6. Implicit Lagrangian systems

In this section, an implicit Lagrangian system is defined in the context of the induced Dirac structure D_{Δ_Q} on T^*Q. As we shall see, the notion of implicit Lagrangian systems that are developed here can treat with the case of systems with degenerate Lagrangians as well as systems with nonholonomic constraints.

6.1. Legendre transform

Let us recall the Legendre transform. Given a Lagrangian $L : TQ \to \mathbb{R}$, define a map $\mathbb{F}L : TQ \to T^*Q$, called the fiber derivative, by

$$\mathbb{F}L(v) \cdot w = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} L(v + sw),$$

where $v, w \in T_qQ$ and $\mathbb{F}L(v) \cdot w$ is the derivative of L at v along the fiber T_qQ in the direction of w. Note that $\mathbb{F}L$ is fiber-preserving; that is, it maps the fiber T_qQ to the fiber T^*_qQ. In local coordinates (q, v) for TQ, the fiber derivative is denoted by

$$\mathbb{F}L(q, v) = \left(q, \frac{\partial L}{\partial v}\right),$$

that is,

$$p = \frac{\partial L}{\partial v},$$

where (q, p) are local coordinates on T^*Q. What we recalled here is the "naive" definition of the Legendre transform; however, notice that the Legendre transform need not be locally invertible; that is, the Lagrangian is degenerate in general, which is a situation that we wish to cover in this paper. For the case of degenerate Lagrangians, it is useful to have a more sophisticated version — a *generalized Legendre transform*, which was originally developed in [43]. In Part II, we will return to illustrating the generalized Legendre transform in preparation for the variational link with implicit Lagrangian systems.

6.2. The Dirac differential operator

We are working towards the important notion of an implicit Lagrangian system associated with a given Lagrangian (possibly degenerate) and a given constraint distribution. We shall do this using the language of Dirac structures.

In this subsection, we make use of the diffeomorphism between T^*TQ and T^*T^*Q, to introduce what we call the Dirac differential operator $\mathfrak{D}$ that acts on a Lagrangian and that will play an important role in the theory that follows.

Let Q be a manifold and let $L : TQ \to \mathbb{R}$ be a Lagrangian. We employ (q, v) as the variables in a local representation for TQ. Recall that $\mathbf{d}L$ is a one-form on TQ, which is locally expressed by

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

Thus, $\mathbf{d}L : TQ \to T^*TQ$. Now define a differential operator $\mathfrak{D}$ acting on the Lagrangian $L : TQ \to \mathbb{R}$, which we shall call the *Dirac differential* of L, by

$$\mathfrak{D}L = \gamma_Q \circ \mathbf{d}L, \tag{6.1}$$

where γ_Q is the diffeomorphism defined by

$$\gamma_Q = \Omega^\flat \circ (\kappa_Q)^{-1} : T^*TQ \to T^*T^*Q.$$

Using the local representation of $\Omega^\flat$ and of κ_Q, we see that locally $\gamma_Q : (q, \delta q, \delta p, p) \mapsto (q, p, -\delta p, \delta q)$ and so $\mathfrak{D}L : TQ \to T^*T^*Q$ is represented, in local coordinates on T^*T^*Q, by

$$\mathfrak{D}L = \left(q, \frac{\partial L}{\partial v}, -\frac{\partial L}{\partial q}, v\right).$$

6.3. Definition of implicit Lagrangian systems

Now we have all the ingredients we need to define an implicit Lagrangian system. We summarize and extend things in the following definition.

Definition 6.1 (*Implicit Lagrangian Systems*). Let $L : TQ \to \mathbb{R}$ be a given Lagrangian (possibly degenerate) and $\Delta_Q \subset TQ$ be a given regular *constraint distribution* on a configuration manifold Q. Denote by D_{Δ_Q} the induced Dirac structure on T^*Q that is given by Eq. (5.3) and write $\mathfrak{D}L : TQ \to T^*T^*Q$ for the Dirac differential of L, defined by Eq. (6.1). Let $P = \mathbb{F}L(\Delta_Q) \subset T^*Q$, the image of Δ_Q under the Legendre transformation.

An *implicit Lagrangian system* is a triple, (L, Δ_Q, X), where X is a vector field on T^*Q, defined at points of P, together with the condition

$$(X, \mathfrak{D}L) \in D_{\Delta_Q}. \tag{6.2}$$

In other words, we require that for each point $v \in \Delta_Q$ and with $z = \mathbb{F}L(v) \in P$, we have

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

In the case $\Delta_Q = TQ$, P is usually called the *primary constraint set*; thus, P defined here is a generalization of the usual primary constraint set.

Writing $z = (q, p) \in P$, notice that the Legendre transform relation $p = \partial L/\partial v$ corresponds exactly to the equality of the base points in Eq. (6.2).

Definition 6.2. A *solution curve* of an implicit Lagrangian system (L, Δ_Q, X) is a curve $(q(t), v(t)) \in \Delta_Q$, $t_1 \leq t \leq t_2$, such that $z(t) = (q(t), p(t))$ is an integral curve of X, where $z(t) = \mathbb{F}L(v(t))$.

Because of the automatic base point equality, a solution curve of an implicit Lagrangian system (L, Δ_Q, X) may be equivalently defined to be a (smooth) curve $(q(t), v(t), p(t))$, where $t_1 \leq t \leq t_2$, whose image lies in the set

$\Delta_Q \oplus P \subset TQ \oplus T^*Q$ and is such that $z(t) = (q(t), p(t))$ is an integral curve of X and such that

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

6.4. Example: The case $\Delta_Q = TQ$

This is perhaps the simplest case in which one has *no constraints, but the Lagrangian may be degenerate*. In this case, the Dirac structure D_{Δ_Q} is given by

$$D_{\Delta_Q}(z) = \{(v_z, \alpha_z) \in T_zP \times T_z^*P \mid \alpha_z(w_z) = \Omega_z(v_z, w_z) \text{ for all } w_z \in T_zT^*Q\}.$$

Therefore, writing $X(q, p) = (q, p, \dot{q}, \dot{p})$, and using the local expressions

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

for the canonical symplectic form and the local expression

$$\mathfrak{D}L = \left(q, \frac{\partial L}{\partial v}, -\frac{\partial L}{\partial q}, v\right)$$

for the Dirac differential, the condition for an implicit Lagrangian system $(X, \mathfrak{D}L) \in D_{\Delta_Q}$ reads that $p = \partial L/\partial v$ (equality of base points as noted above) and that

$$\left\langle -\frac{\partial L}{\partial q}, u \right\rangle + \langle v, \alpha \rangle = \langle \alpha, \dot{q} \rangle - \langle \dot{p}, u \rangle$$

for all u and all α, where (u, α) are the local representatives of a point in $T_{(q,p)}T^*Q$. Since this holds for all (u, α), this means that in this case, *the condition of an implicit Lagrangian system is equivalent to the Euler–Lagrange equations $\dot{p} = \partial L/\partial q$ together with the condition $v = \dot{q}$.*

In the usual formulation of degenerate Lagrangian systems, one often must simply *assume* that the condition $v = \dot{q}$ holds; for instance, [1], Theorem 3.5.17 is typical. However, in the present context, the condition $v = \dot{q}$ *follows automatically from the definition of an implicit Lagrangian system*. Notice that integral curves automatically satisfy the primary constraints and that this is built into the formalism in a natural way. Of course in typical examples such as when L is linear in the velocity, the actual equations need not be literally of second order.

The above example gives a good indication of the manner in which the general notion of an implicit Lagrangian system includes interesting cases. We will explore examples with a nontrivial constraint set Δ_Q shortly, but we first write out the conditions defining an implicit Lagrangian system in local representation.

These calculations prove the following local representation for an implicit Lagrangian system.

Proposition 6.3. *Locally, and using the preceding notation, the condition $(X, \mathfrak{D}L) \in D_{\Delta_Q}$ defining an implicit Lagrangian system reads*

$$p = \frac{\partial L}{\partial v}, \quad \dot{q} \in \Delta(q), \quad v = \dot{q}, \quad \textit{and} \quad \dot{p} - \frac{\partial L}{\partial q} \in \Delta^\circ(q). \tag{6.3}$$

Let us develop a convenient matrix representation for these equations in the following example.

Example. Let Q be a configuration manifold. Given nonholonomic constraints, denoted by the distribution Δ_Q on Q, we choose local coordinates q^i on Q so that locally, Q is expressed by an open set $U \subset \mathbb{R}^n$. The constraint set Δ_Q defines a subspace of TQ, which we denote by $\Delta(q) \subset \mathbb{R}^n$ at each point $q \in U$. If we let the dimension of the constraint space be $n - m$, then we can choose a basis $e_{m+1}(q), e_{m+2}(q), \ldots, e_n(q)$ of $\Delta(q)$. So, the constraint set denoted by the annihilator $\Delta^\circ(q)$ is spanned by one-forms $\omega^1, \omega^2, \ldots, \omega^m$; that is,

$$\Delta(q) = \{v^i \in \mathbb{R}^n \mid \omega_i^a(q)v^i = 0, a = 1, \ldots, m\},$$

where we employ the local expressions $\omega^a = \omega_i^a(q)\mathrm{d}q^i, a = 1, \ldots, m$. Then, using the Lagrange multipliers $\mu_a, a = 1, \ldots, m$, a nonholonomic mechanical system is locally represented by an implicit Lagrangian system

(L, Δ_Q, X) such that

$$\begin{aligned} &\begin{pmatrix} \dot{q}^i \\ \dot{p}_i \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} -\dfrac{\partial L}{\partial q^i} \\ v^i \end{pmatrix} + \begin{pmatrix} 0 \\ \mu_a \omega_i^a(q) \end{pmatrix}, \\ &p_i = \frac{\partial L}{\partial v^i}, \\ &0 = \omega_i^a(q) v^i. \end{aligned} \tag{6.4}$$

6.5. The "second-order" condition

We pick out the following simple but basic result (consistent with the above example) for special attention to make a point about the usefulness of the present formulation.

Corollary 6.4. *Let $(q(t), v(t))$ be an integral curve of a given implicit Lagrangian system (L, Δ_Q, X). Then $v(t) = \dot{q}(t) \in \Delta_Q$.*

Proof. This is a special case of (6.3). □

In the special case of nonholonomic systems, as we shall examine in more detail later, the condition $v(t) = \dot{q}(t)$ is *normally postulated separately from the equations of motion* (see, for example, [8] and references therein) but the above development shows that in fact, it follows from the condition of having an integral curve of an implicit Lagrangian system.

6.6. Conservation of energy

Define the *generalized energy* $E : TQ \oplus T^*Q \to \mathbb{R}$ by

$$E(q, v, p) = \langle p, v \rangle - L(q, v),$$

where $(q, v) \in \Delta_Q$ and $(q, p) \in P$. As is common in the Pontryagin maximum principle (see, for instance, [8]), stationarity of $E(q, v, p) = \langle p, v \rangle - L(q, v)$ with respect to v defines the submanifold $\mathcal{K}$ obtained by imposing the base point, or the Legendre transform condition $p = \partial L/\partial v$:

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

We now show that energy is conserved for any implicit Lagrangian system as follows.

Proposition 6.5. *Let $(q(t), v(t))$ be an integral curve of a given implicit Lagrangian system (L, Δ_Q, X). Then with $p(t) = (\partial L/\partial v)(t)$, the function $E(q(t), v(t), p(t))$ is constant in time.*

Proof. We give the proof in local representation. From (6.3) and the definition of E, we get

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} E &= \langle \dot{p}, v \rangle + \langle p, \dot{v} \rangle - \frac{\partial L}{\partial q} \dot{q} - \frac{\partial L}{\partial v} \dot{v} \\ &= \left\langle \dot{p} - \frac{\partial L}{\partial q}, v \right\rangle \end{aligned}$$

which vanishes since $v = \dot{q} \in \Delta(q)$ and since $\dot{p} - \frac{\partial L}{\partial q} \in \Delta^\circ(q)$. □

Remarks. Recall that, as was illustrated in Section 3, interconnections can be recognized as Dirac structures and also that the Tellegen's theorem is deduced from the (KCL and KVL) interconnections in circuits. Interestingly, in mechanics, Birkhoff [3] introduced an analogous idea of interconnections called *nonenergic systems* in the context of Lagrangian systems. The nonenergicness is the condition of energy conservation such that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}E(q,v) &= \left\langle \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial v}, v\right\rangle + \left\langle \frac{\partial L}{\partial v}, \dot{v}\right\rangle - \left\langle \frac{\partial L}{\partial q}, v\right\rangle - \left\langle \frac{\partial L}{\partial v}, \dot{v}\right\rangle \\ &= \left\langle \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial v} - \frac{\partial L}{\partial q}, v\right\rangle = 0,\end{aligned}$$

where $E(q,v) = \langle \partial L/\partial v, v\rangle - L(q,v)$, $v = \dot{q} \in \Delta(q)$ and $\mathrm{d}(\partial L/\partial v)/\mathrm{d}t - \partial L/\partial q \in \Delta^{\circ}(q)$. Duinker [25,26] investigated nonenergic Lagrangian circuits called *nonenergic multiports* including an element of *traditor*, whose Lagrangian is given by $L(q,\dot{q}) = \langle \alpha(q), \dot{q}\rangle - V(q)$, where α is a one-form on W which is associated with the traditor and V is a potential function on W associated with conservative elements. Notice that the above Lagrangian is linear in $\dot{q}$ and is degenerate, and this is quite similar to the case of point vortex systems (see, for instance, [39]). Birkhoff [3] actually considered the special case of nonenergic Lagrangian systems of $L(q,\dot{q}) = \langle \alpha(q), \dot{q}\rangle$ that satisfy $E(q,\dot{q}) = 0$; that is, the energy itself is zero; for example, an electrically charged particle of negligible mass moving in a static magnetic field. The nonenergic multiports or energy-conserving interconnection multiports were investigated in the context of reciprocal and nonreciprocal circuits in [11] and [49], and were also studied in the general context of Dirac structures with (implicit) Hamiltonian systems in [33,45,22], where the system with nonenergic multiports was called an *energy conserving system or port-controlled system*.

7. Examples

In this section, we show how the proposed idea of an implicit Lagrangian system can be applied to a nonholonomic mechanical system and an electric circuit with two illustrative examples; the first example is a vertical rolling disk on a plane as a nonholonomic system and the second one is an L–C circuit for a degenerate Lagrangian system with holonomic constraints.

7.1. Example: The vertical rolling disk

Let us consider the vertical rolling disk on a plane, which is one of the most fundamental and essential examples of mechanical systems with nonholonomic constraints; see [9] and [8].

Consider a homogeneous vertical disk with a radius R rolling on the xy-plane about its vertical axis. Let x and y denote the position of contact of the disk in the xy-plane. Let θ be the rotation angle of the disk and φ be the heading angle of the disk. Then, the configuration space of the vertical rolling disk is $Q = \mathbb{R}^2 \times S^1 \times S^1$, whose local coordinates are given by $q = (x, y, \theta, \varphi)$. Let $(q, v) = (x, y, \theta, \varphi, v_x, v_y, v_\theta, v_\varphi)$ be the local coordinates for TQ. The Lagrangian is given by

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

where m is the mass of the disk and I and J are the moments of inertia. Note that we consider the full configuration space and also that the Lagrangian is the standard free Lagrangian without constraints.

The kinematic constraints due to the rolling contact without slipping on the plane are denoted by

$$\begin{aligned}\dot{x} &= R(\cos\varphi)\dot{\theta},\\ \dot{y} &= R(\sin\varphi)\dot{\theta}.\end{aligned}$$

The constraint may be represented by the distribution $\Delta_Q \subset TQ$ such that, for each $q \in Q$,

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

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

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

Note that ω^a span the annihilator Δ_Q° of the distribution Δ_Q.

The differential of the Lagrangian is denoted by

$$\begin{aligned}\mathbf{d}L &= \left(x, y, \theta, \varphi, v_x, v_y, v_\theta, v_\varphi, \frac{\partial L}{\partial x}, \frac{\partial L}{\partial y}, \frac{\partial L}{\partial \theta}, \frac{\partial L}{\partial \varphi}, \frac{\partial L}{\partial v_x}, \frac{\partial L}{\partial v_y}, \frac{\partial L}{\partial v_\theta}, \frac{\partial L}{\partial v_\varphi}\right)\\ &= \left(x, y, \theta, \varphi, v_x, v_y, v_\theta, v_\varphi, 0, 0, 0, 0, mv_x, mv_y, Iv_\theta, Jv_\varphi\right),\end{aligned}$$

and therefore the Dirac differential of the Lagrangian is given by using the diffeomorphism $\gamma_Q = \Omega^\flat \circ (\kappa_Q)^{-1} : T^*TQ \to T^*T^*Q$ such that

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

In the above, the Legendre transformation $\mathbb{F}L : TQ \to T^*Q$ gives

$$\begin{aligned}\left(x, y, \theta, \varphi, p_x, p_y, p_\theta, p_\varphi\right) &= \left(x, y, \theta, \varphi, \frac{\partial L}{\partial v_x}, \frac{\partial L}{\partial v_y}, \frac{\partial L}{\partial v_\theta}, \frac{\partial L}{\partial v_\varphi}\right)\\ &= \left(x, y, \theta, \varphi, m v_x, m v_y, I v_\theta, J v_\varphi\right),\end{aligned}$$

where $(q, p) = (x, y, \theta, \varphi, p_x, p_y, p_\theta, p_\varphi)$ are local coordinates for T^*Q. The distribution Δ_Q on Q is locally given by

$$\Delta_Q(q) = \operatorname{span}\left\{\frac{\partial}{\partial \varphi}, R\cos\varphi\frac{\partial}{\partial x} + R\sin\varphi\frac{\partial}{\partial y} + \frac{\partial}{\partial \theta}\right\}.$$

Recall the distribution Δ_Q is to be lifted to T^*Q such that

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

where $\pi_Q : T^*Q \to Q$ is the cotangent projection and $T\pi_Q : TT^*Q \to T^*Q$ is its tangent map.

As in Eqs. (5.6) and (5.7), we can define the distribution $\Delta^*_{T^*Q} = \Omega^\flat(\Delta_{T^*Q})$ and its annihilator $(\Delta^*_{T^*Q})^\circ$. Then, using the canonical Poisson structure $B : T^*T^*Q \times T^*T^*Q \to \mathbb{R}$, the induced Dirac structure on T^*Q is expressed, as shown in Eq. (5.8), such that, for each $(q, p) \in T^*Q$,

$$\begin{aligned}D_{\Delta_Q}(q, p) = \{&(v_{(q,p)}, \alpha_{(q,p)}) \in T_{(q,p)}T^*Q \times T^*_{(q,p)}T^*Q \mid \alpha_{(q,p)} \in \Delta^*_{T^*Q}(q, p)\\ &\text{and } v_{(q,p)} - B^\sharp(q, p)\cdot\alpha_{(q,p)} \in (\Delta^*_{T^*Q})^\circ(q, p)\}.\end{aligned}$$

Let X denote the vector field on T^*Q, which is locally denoted by, for each $(q, p) \in T^*Q$,

$$X(q, p) = (\dot{x}, \dot{y}, \dot{\theta}, \dot{\varphi}, \dot{p}_x, \dot{p}_y, \dot{p}_\theta, \dot{p}_\varphi).$$

Then, the triple (L, Δ_Q, X) is an implicit Lagrangian system that satisfies, for each $v_q \in \Delta_Q(q)$ and with $(q, p) = \mathbb{F}L(q, v)$,

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

Recall the local expressions for an implicit Lagrangian system are given by Eq. (6.4), and we obtain the following matrix representation:

$$\begin{pmatrix}\dot{x}\\ \dot{y}\\ \dot{\theta}\\ \dot{\varphi}\\ \hline \dot{p}_x\\ \dot{p}_y\\ \dot{p}_\theta\\ \dot{p}_\varphi\end{pmatrix} = \left(\begin{array}{cccc|cccc} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1\\ \hline -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0\end{array}\right)\begin{pmatrix}0\\ 0\\ 0\\ 0\\ \hline v_x\\ v_y\\ v_\theta\\ v_\varphi\end{pmatrix} + \left(\begin{array}{cc} 0 & 0\\ 0 & 0\\ 0 & 0\\ 0 & 0\\ \hline 1 & 0\\ 0 & 1\\ -R\cos\varphi & -R\sin\varphi\\ 0 & 0\end{array}\right)\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},$$

together with the Legendre transformation

$$\begin{pmatrix} p_x \\ p_y \\ p_\theta \\ p_\varphi \end{pmatrix} = \begin{pmatrix} m v_x \\ m v_y \\ I v_\theta \\ J v_\varphi \end{pmatrix},$$

and with the kinematic constraints

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -R\cos\varphi & 0 \\ 0 & 1 & -R\sin\varphi & 0 \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ v_\theta \\ v_\varphi \end{pmatrix}.$$

Then, we obtain the implicit differential-algebraic equations for the system as

$$\begin{aligned} &\dot{p}_x = \mu_1,\ \dot{p}_y = \mu_2,\ \dot{p}_\theta = -(R\cos\varphi)\mu_1 - (R\sin\varphi)\mu_2,\ \dot{p}_\varphi = 0, \\ &\dot{x} = v_x,\ \dot{y} = v_y,\ \dot{\theta} = v_\theta,\ \dot{\varphi} = v_\varphi,\ v_x = (R\cos\varphi)v_\theta,\ v_y = (R\sin\varphi)v_\theta, \\ &p_x = m v_x,\ p_y = m v_y,\ p_\theta = I v_\theta,\ p_\varphi = J v_\varphi. \end{aligned}$$

7.2. Example: L–C circuits

Let us investigate an example of L–C circuits, i.e., a lossless circuit comprising inductors (L's) and capacitors (C's) in the framework of implicit Lagrangian systems. The system of L–C circuits is a typical degenerate Lagrangian system that is accompanied with KCL (holonomic) constraints as well as primary constraints in the sense of Dirac's theory. In analogy with mechanics, the KCL constraints may be recognized as kinematic constraints. Then, we demonstrate how a degenerate Lagrangian system with the KCL (or equivalently kinematic) constraints can be incorporated into the context of implicit Lagrangian systems by an illustrative example of L–C circuits. For comparison with an approach of implicit Hamiltonian systems, refer to [7,46].

In order to investigate L–C circuits, let us employ an analogy with mechanics in such a way that a configuration space Q of circuits may be recognized as the *charge space* whose point indicates a charge associated with every element of the circuit. This implies that local coordinates for the charge q^i as to every element correspond to local coordinates of a point of the configuration space in mechanics. In this analogy, the tangent bundle TQ may be regarded as the *current space*, whose local coordinates are denoted by (q, f), while the cotangent bundle T^*Q corresponds to the *flux linkage space*, whose local coordinates are given by (q, p). In this context, $f \in T_qQ$ indicates the current, while $p \in T_q^*Q$ denotes the flux linkage.

Consider a four-port L–C circuit shown in Fig. 7.1, which was also investigated in [46]. In the illustrative example of L–C circuits, the configuration manifold Q is a four-dimensional vector space $E = \mathbb{R}^4$, that is, $Q = E$. Then, we have $TQ = TE$ $(\cong E \times E)$ and $T^*Q = T^*E$ $(\cong E \times E^*)$. Let $q = (q_L, q_{C_1}, q_{C_2}, q_{C_3}) \in E$ and $f = (f_L, f_{C_1}, f_{C_2}, f_{C_3}) \in T_qE$.

Let $T : TE\ (= E \times E) \to \mathbb{R}$ be the magnetic energy of the L–C circuit, which is defined by the inductance L such that

$$T_q(f) = \frac{1}{2}L(f_L)^2,$$

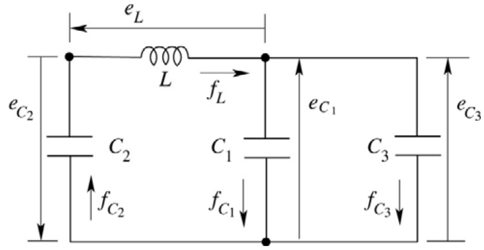

Fig. 7.1. L–C Circuit.

and let $V : E \to \mathbb{R}$ be the electric potential energy of the L–C circuit, which is defined by capacitors C_1, C_2, and C_3 such that

$$V(q) = \frac{1}{2}\frac{(q_{C_1})^2}{C_1} + \frac{1}{2}\frac{(q_{C_2})^2}{C_2} + \frac{1}{2}\frac{(q_{C_3})^2}{C_3}.$$

Then, we can define the Lagrangian of the L–C circuit $\mathcal{L} : TE\ (\cong E \times E) \to \mathbb{R}$ by

$$\begin{aligned}\mathcal{L}(q, f) &= T_q(f) - V(q)\\ &= \frac{1}{2}L(f_L)^2 - \frac{1}{2}\frac{(q_{C_1})^2}{C_1} - \frac{1}{2}\frac{(q_{C_2})^2}{C_2} - \frac{1}{2}\frac{(q_{C_3})^2}{C_3}.\end{aligned} \tag{7.1}$$

The KCL constraints indicate constraints of the current given, in coordinates, by

$$\begin{aligned}&-f_L + f_{C_2} = 0,\\ &-f_{C_1} + f_{C_2} - f_{C_3} = 0,\end{aligned} \tag{7.2}$$

which are represented by

$$\langle \omega^a, f\rangle = 0, \quad a = 1, 2,$$

where $f = (f_L, f_{C_1}, f_{C_2}, f_{C_3}) \in T_qE$ and ω^a denote 2-independent covectors (or one-forms) represented, in coordinates, by

$$\omega^a = \omega^a_k \mathrm{d}q^k, \quad a = 1, 2; k = 1, \ldots, 4,$$

where $q = (q^1, q^2, q^3, q^4) = (q_L, q_{C_1}, q_{C_2}, q_{C_3})$. In this example, the coefficients ω^a_k are given in matrix representation by

$$\omega^a_k = \begin{pmatrix} -1 & 0 & 1 & 0\\ 0 & -1 & 1 & -1\end{pmatrix}.$$

Therefore, the set of currents that satisfy the KCL constraints forms a constraint subspace $\Delta \subset TE$, which we shall call the *constraint KCL space* that is defined, for each $q \in E$, by

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

On the other hand, the annihilator Δ° of Δ is the *constraint KVL space*, which is defined, for each $q \in E$, by

$$\Delta^\circ_q = \{e \in T^*_qE \mid \langle e, f\rangle = 0,\ \text{for all } f \in \Delta_q\}.$$

In the above, $e \in \Delta^\circ_q$ denotes the voltage of the circuit that is the covector in the constraint KVL space, which is represented, in coordinates, by using Lagrange multipliers μ_a such that

$$e_k = \mu_a \omega^a_k, \quad a = 1, 2,\ k = 1, \ldots, 4,$$

where $e = (e_1, e_2, e_3, e_4) = (e_L, e_{C_1}, e_{C_2}, e_{C_3})$.

Setting $f = (f_1, f_2, f_3, f_4) = (f_L, f_{C_1}, f_{C_2}, f_{C_3})$, it is obvious that the Lagrangian $\mathcal{L} : TE \to \mathbb{R}$ of the L–C circuit defined by Eq. (7.1) is degenerate, since

$$\det\left[\frac{\partial^2 \mathcal{L}}{\partial f^i \partial f^j}\right] = 0.$$

The *constraint flux linkage subspace* is defined by the Legendre transform such that

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

In coordinates, $(q, f) = (q_L, q_{C_1}, q_{C_2}, q_{C_3}, f_L, f_{C_1}, f_{C_2}, f_{C_3}) \in \Delta \subset TE$, where the current $f = (f_L, f_{C_1}, f_{C_2}, f_{C_3})$ satisfies the KCL constraints in Eq. (7.2). So we can define $(q, p) = \mathbb{F}\mathcal{L}(q, f) \in T^*E$. Since the Legendre transform is expressed, in coordinates, by

$$(q_L, q_{C_1}, q_{C_2}, q_{C_3}, p_L, p_{C_1}, p_{C_2}, p_{C_3}) = \left(q_L, q_{C_1}, q_{C_2}, q_{C_3}, \frac{\partial \mathcal{L}}{\partial f_L}, \frac{\partial \mathcal{L}}{\partial f_{C_1}}, \frac{\partial \mathcal{L}}{\partial f_{C_2}}, \frac{\partial \mathcal{L}}{\partial f_{C_3}}\right),$$

by computing, we obtain

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

The constraints of the flux linkages, that is, $p_{C_1} = p_{C_2} = p_{C_3} = 0$ correspond to *primary constraints* in the sense of Dirac, which form the constraint flux linkage subspace $P \subset T^*E = E \times E^*$, and it immediately reads

$$(q, p) = \left(q_L, q_{C_1}, q_{C_2}, q_{C_3}, p_L, 0, 0, 0\right) \in P \subset T^*E.$$

Let X be the vector field on T^*E, defined at each point in P such that

$$X = \left(q_L, q_{C_1}, q_{C_2}, q_{C_3}, p_L, 0, 0, 0, \dot{q}_L, \dot{q}_{C_1}, \dot{q}_{C_2}, \dot{q}_{C_3}, \dot{p}_L, 0, 0, 0\right). \tag{7.3}$$

Since the differential of the Lagrangian $\mathbf{d}\mathcal{L} = (q, f, \partial\mathcal{L}/\partial q, \partial\mathcal{L}/\partial f)$ is given by

$$\mathbf{d}\mathcal{L} = \left(q_L, q_{C_1}, q_{C_2}, q_{C_3}, f_L, f_{C_1}, f_{C_2}, f_{C_3}, 0, -\frac{q_{C_1}}{C_1}, -\frac{q_{C_2}}{C_2}, -\frac{q_{C_3}}{C_3}, Lf_L, 0, 0, 0\right),$$

the Dirac differential of the Lagrangian $\mathfrak{D}\mathcal{L} = (q, \partial\mathcal{L}/\partial f, -\partial\mathcal{L}/\partial q, f)$ may be denoted by

$$\mathfrak{D}\mathcal{L} = \left(q_L, q_{C_1}, q_{C_2}, q_{C_3}, Lf_L, 0, 0, 0, 0, \frac{q_{C_1}}{C_1}, \frac{q_{C_2}}{C_2}, \frac{q_{C_3}}{C_3}, f_L, f_{C_1}, f_{C_2}, f_{C_3}\right). \tag{7.4}$$

As was illustrated, the induced distribution Δ_{T^*E} on T^*E is defined by the KCL constraint distribution $\Delta \subset TE$ such that

$$\Delta_{T^*E} = (T\pi_E)^{-1}(\Delta) \subset TT^*E,$$

where $\pi_E : T^*E \to E$ is the canonical projection and $T\pi_E : TT^*E \to TE$. Let $\Delta^\circ_{T^*E}$ be the annihilator of Δ_{T^*E} and let Ω be the canonical symplectic structure on T^*E, and then the Dirac structure D_Δ on T^*E induced from the KCL constraint distribution Δ can be defined by, for each $z = (q, p) \in T^*E$,

$$D_\Delta(z) = \{(w_z, \alpha_z) \in T_zT^*E \times T^*_zT^*E \mid w_z \in \Delta_{T^*E}(z), \text{and } \alpha_z - \Omega^\flat(z) \cdot w_z \in \Delta^\circ_{T^*E}(z)\}.$$

As shown in Eq. (5.4), we can also define the induced Dirac structure in a different way; that is, let $\tau_E : TE \to E$ be the canonical projection and define

$$\Delta_{TE} = T\tau_E^{-1}(\Delta) \subset TTE.$$

Letting Δ°_{TE} be its annihilator, the induced Dirac structure can be given, by employing the symplectomorphisms $\kappa_E : T_zT^*E \to T^*_zTE$ and $\sigma_E : T^*_zT^*E \to T^*_zTE$, such that, for each $z \in T^*_qE$,

$$D_\Delta(z) = \{(w_z, \alpha_z) \in T_zT^*E \times T^*_zT^*E \mid T\pi_E(w_z) = \pi^2(\alpha_z) =: f_q \in \Delta_q, \kappa_E(w_z) - \sigma_E(\alpha_z) \in \Delta^\circ_{TE}(f_q)\}.$$

By Eqs. (7.3) and (7.4), it follows that the condition

$$(X(q, p), \mathfrak{D}\mathcal{L}(q, f)) \in D_\Delta(q, p)$$

holds for each $(q, f) \in \Delta \subset TE$ and with $(q, p) = \mathbb{F}\mathcal{L}(q, f)$.

Thus, L–C circuits can be understood in the context of the implicit Lagrangian system $(\mathcal{L}, \Delta, X)$, which is locally described by Eq. (6.4). It immediately follows that, in matrix representation,

$$\begin{pmatrix} \dot{q}_L \\ \dot{q}_{C_1} \\ \dot{q}_{C_2} \\ \dot{q}_{C_3} \\ \hline \dot{p}_L \\ 0 \\ 0 \\ 0 \end{pmatrix} = \left(\begin{array}{cccc|cccc} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{array}\right) \begin{pmatrix} 0 \\ \frac{q_{C_1}}{C_1} \\ \frac{q_{C_2}}{C_2} \\ \frac{q_{C_3}}{C_3} \\ \hline f_L \\ f_{C_1} \\ f_{C_2} \\ f_{C_3} \end{pmatrix} + \left(\begin{array}{cc} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \hline -1 & 0 \\ 0 & -1 \\ 1 & 1 \\ 0 & -1 \end{array}\right) \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix},$$

together with the Legendre transformation

$$p_L = Lf_L,$$

and with the KCL (holonomic) constraints

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

Thus, we obtain the implicit differential-algebraic equations as

$$\begin{aligned} &\dot{q}_L = f_L, \dot{q}_{C_1} = f_{C_1}, \dot{q}_{C_2} = f_{C_2}, \dot{q}_{C_3} = f_{C_3}, \\ &\dot{p}_L = -\mu_1, \mu_2 = -\frac{q_{C_1}}{C_1}, \mu_1 = -\mu_2 + \frac{q_{C_2}}{C_2}, \mu_2 = -\frac{q_{C_3}}{C_3}, \\ &p_L = Lf_L, f_L = f_{C_2}, f_{C_1} = f_{C_2} - f_{C_3}. \end{aligned}$$

8. Conclusions

In this Part I of a two-part paper, we have proposed a framework for implicit Lagrangian systems, which makes use of Dirac structures on the cotangent bundles that are induced by constraint distributions on the underlying configuration manifold. We have shown that systems such as nonholonomic mechanical systems and L–C circuits (which often have degenerate Lagrangians) are natural examples of implicit Lagrangian systems.

We showed how a Dirac structure D_{Δ_Q} on a cotangent bundle T^*Q is naturally induced by a distribution Δ_Q on Q, and we also showed some basic examples of these induced Dirac structures, such as the KCL and KVL constraints in electric circuits, interconnections, and constraints in nonholonomic mechanics.

The general notion of an implicit Lagrangian system makes use of the Dirac differential $\mathfrak{D}L : TQ \to T^*T^*Q$ of a Lagrangian $L : TQ \to \mathbb{R}$. The definition of the Dirac differential makes use of the diffeomorphisms between the spaces TT^*Q, T^*TQ, and T^*T^*Q as well as the Legendre transformation. An implicit Lagrangian system (L, Δ_Q, X) then is defined by the property that $(X(z), \mathfrak{D}L(v)) \in D_{\Delta_Q}(z)$ for each $v \in \Delta_Q \subset TQ$.

Finally, we illustrated the theory with two specific examples; the first is a vertical rolling disk on a plane as a nonholonomic mechanical system and the second is an L–C circuit, which has a degenerate Lagrangian as well as holonomic constraints.

In Part II, we explore links between variational principles, Dirac structures and implicit Lagrangian systems. To establish these links, use is made of an extension of Hamilton's principle, one version of which is known as the Hamilton–Pontryagin principle, or sometimes, Liven's principle. We also investigate implicit Lagrangian systems in the general context of an extended Lagrange–d'Alembert principle. Furthermore, we address how implicit Hamiltonian systems can be formulated in the context of variational principles.

Acknowledgements

We are very grateful to Tudor Ratiu, Alan Weinstein and Tom Hughes, who kindly provided several very useful remarks. We also thank Jeff Lawson, Shawn Shadden, and Troy Smith for pointing out several improvements.

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, Benjamin-Cummings Publ. Co, 1978, Updated 1985 version, reprinted by Persius Publishing, second ed.

[2] L. Bates, J. Sniatycki, Nonholonomic reduction, Rep. Math. Phys. 32 (1993) 99–115.

[3] G.D. Birkhoff, Dynamical Systems, in: Colloquium Publications, vol. 9, Amer. Math. Soc., Providence, RI, 1927.

[4] G. Blankenstein, Implicit Hamiltonian Systems: Symmetry and Interconnection, Ph.D. Dissertation, University of Twente, 2000.

[5] G. Blankenstein, A.J. van der Schaft, Symmetry and reduction in implicit generalized Hamiltonian systems, Rep. Math. Phys. 47 (1) (2001) 57–100.

[6] G. Blankenstein, T.S. Ratiu, Singular reduction of implicit Hamiltonian systems, Rep. Math. Phys. 53 (2) (2004) 211–260.
[7] A.M. Bloch, P.E. Crouch, Representations of Dirac structures on vector spaces and nonlinear L–C circuits, in: Differential Geometry and Control (Boulder, CO, 1997), in: Proc. Sympos. Pure Math., vol. 64, Amer. Math. Soc., Providence, RI, 1997, pp. 103–117.
[8] 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.
[9] A.M. Bloch, P.S. Krishnaprasad, J.E. Marsden, R. Murray, Nonholonomic mechanical systems with symmetry, Arch. Ration. Mech. Anal. 136 (1996) 21–99.
[10] R.K. Brayton, J.K. Moser, A theory of nonlinear networks. I, II, Quart. Appl. Math. 22 (1964) 1–33, 81–104.
[11] R.K. Brayton, Nonlinear reciprocal networks, in: Mathematical Aspects of Electrical Network Analysis, Amer. Math. Soc., Providence, RI, 1971, pp. 1–15.
[12] 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.
[13] H. Cendra, J.E. Marsden, T.S. Ratiu, Lagrangian Reduction By Stages, in: Memoirs, vol. 152, American Mathematical Society, Providence, RI, 2001.
[14] D. Chang, A.M. Bloch, N. Leonard, J.E. Marsden, C. Woolsey, The equivalence of controlled Lagrangian and controlled Hamiltonian systems, Control Calc. Var. 8 (2002) 393–422 (special issue dedicated to J.L. Lions).
[15] D.E. Chang, J.E. Marsden, Reduction of controlled Lagrangian and Hamiltonian systems with symmetry, SIAM. J. Control Optim. 43 (2004) 277–300.
[16] T. Chen, Critical manifolds and stability in Hamiltonian systems with non-holonomic constraints, J. Geom. Phys. 49 (2004) 418–462.
[17] L.O. Chua, C.A. Desoer, D.A. Kuh, Linear and Nonlinear Circuits, McGraw-Hill Inc., 1987.
[18] J. Cortés, M.D. León, D.M.D. Diego, S. Martínez, Geometric description of vakonomic and nonholonomic dynamics. Comparison of solutions, SIAM J. Control Optim. 41 (5) (2003) 1389–1412.
[19] T.J. Courant, Dirac manifolds, Trans. Amer. Math. Soc. 319 (1990) 631–661.
[20] T.J. Courant, Tangent Dirac structures, J. Phys. A: Math. Gen. 23 (1990) 5153–5168.
[21] T. Courant, A. Weinstein, Beyond Poisson structures, in: Action hamiltoniennes de groupes. Troisième théorème de Lie (Lyon, 1986), in: Travaux en Cours, vol. 27, Hermann, Paris, 1988, pp. 39–49.
[22] M. Dalsmo, A.J. van der Schaft, On representations and integrability of mathematical structures in energy-conserving physical systems, SIAM J. Control Optim. 37 (1) (1998) 54–91.
[23] P.A.M. Dirac, Generalized Hamiltonian dynamics, Canad. J. Math. 2 (1950) 129–148.
[24] I. Dorfman, Dirac structures of integrable evolution equations, Phys. Lett. A 125 (5) (1987) 240–246.
[25] S. Duinker, Traditors, a new class of non-energic non-linear network elements, Philips Res. Rep. 14 (1959) 29–51.
[26] S. Duinker, Conjunctors, another new class of non-energic non-linear network elements, Philips Res. Rep. 17 (1962) 1–19.
[27] J. Koiller, Reduction of some classical nonholonomic systems with symmetry, Arch. Rat. Mech. Anal. 118 (1992) 113–148.
[28] W.S. Koon, J.E. Marsden, The Hamiltonian and Lagrangian approaches to the dynamics of nonholonomic systems, Rep. Math. Phys. 40 (1997) 21–62.
[29] W.S. Koon, J.E. Marsden, Poisson reduction for nonholonomic mechanical systems with symmetry, Rep. Math. Phys. 42 (1998) 101–134.
[30] G. Kron, Tensor Analysis of Networks, John Wiley & Sons, New York, 1939.
[31] G. Kron, Diakoptics—The Piecewise Solution of Large-Scale Systems, MacDonald, London, 1963.
[32] J.E. Marsden, T.S. Ratiu, Introduction to Mechanics and Symmetry, second ed., in: Texts in Applied Mathematics, vol. 17, Springer-Verlag, 1999.
[33] B.M. Maschke, A.J. van der Schaft, P.C. Breedveld, An intrinsic Hamiltonian formulation of the dynamics of L–C circuits, IEEE Trans. Circuits Syst. 42 (2) (1995) 73–82.
[34] L. Moreau, D. Aeyels, A novel variational method for deriving Lagrangian and Hamiltonian models of inductor–capacitor circuits, SIAM Rev. 46 (2004) 59–84.
[35] J.I. Neimark, N.A. Fufaev, Dynamics of Nonholonomic Systems, in: Translations of Mathematical Monographs, vol. 33, AMS, 1972.
[36] G.F. Oster, A.S. Perelson, Chemical reaction dynamics. I: Geometrical structure, Arch. Ration. Mech. Anal. 55 (1974) 230–274.
[37] W.M. Oliva, Lagrangian systems on manifolds. I, Celestial Mech. 1 (1970) 491–511.
[38] A.S. Perelson, G.F. Oster, Chemical reaction dynamics. II: Reaction networks, Arch. Ration. Mech. Anal. 57 (1974) 31–98.
[39] C.W. Rowley, J.E. Marsden, Variational integrators for point vortices, Proc. CDC 40 (2002) 1521–1527.
[40] S.S. Sastry, C.A. Desoer, Jump behavior of circuits and systems, IEEE Trans. Circuits Syst. 28 (1981) 1109–1124.
[41] R. Skinner, R. Rusk, Generalized Hamiltonian dynamics. I: Formulation on $T^*Q \oplus TQ$, J. Math. Phys. 24 (1983) 2589–2594. See also the same issue, pages 2581–2588 and 2595–2601.
[42] S. Smale, On the mathematical foundations of electrical circuit theory, J. Differential Geom. 7 (1972) 193–210.
[43] W.M. Tulczyjew, The Legendre transformation, Ann. Inst. H. Poincaré A 27 (1) (1977) 101–114.
[44] A.J. van der Schaft, B.M. Maschke, On the Hamiltonian formulation of nonholonomic mechanical systems, Rep. Math. Phys. 34 (1994) 225–233.
[45] A.J. van der Schaft, B.M. Maschke, The Hamiltonian formulation of energy conserving physical systems with external ports, Arch. Elektr. Übertrag. 49 (1995) 362–371.
[46] A.J. van der Schaft, Implicit Hamiltonian systems with symmetry, Rep. Math. Phys. 41 (1998) 203–221.
[47] A.M. Vershik, L.D. Faddeev, Lagrangian mechanics in invariant form, Sel. Math. Sov. 1 (1981) 339–350.
[48] R.W. Weber, Hamiltonian systems with constraints and their meaning in mechanics, Arch. Ration. Mech. Anal. 91 (4) (1986) 309–335.
[49] J.L. Wyatt, L.O. Chua, A theory of nonenergic n-ports, Circuit Theory Appl. 5 (1977) 181–208.
[50] H. Yoshimura, Dynamics of Flexible Multibody Systems, Doctoral Dissertation, Waseda University, 1995.