

# A Novel Variational Method for Deriving Lagrangian and Hamiltonian Models of Inductor-Capacitor Circuits\*

Luc Moreau<sup>†</sup>  
Dirk Aeyels<sup>‡</sup>

**Abstract.** We study the dynamical equations of nonlinear inductor-capacitor circuits. We present a novel Lagrangian description of the dynamics and provide a variational interpretation, which is based on the maximum principle of optimal control theory. This gives rise to an alternative method for deriving the dynamic equations. We show how this generalized Lagrangian description is related to generalized Hamiltonian models discussed in the literature by means of a Legendre transformation. Some distinctive features of the present approach are that it is applicable to circuits with arbitrary topology and that the variational principle and the resulting equations do not involve nonphysical inductor charges or capacitor fluxes.

**Key words.** electric circuits, variational modeling, Lagrangian dynamics, Hamiltonian dynamics, Dirac structures, optimal control, maximum principle

**AMS subject classifications.** 70H03, 49S05, 94C05

**DOI.** 10.1137/S0036144502409020

**1. Introduction.** In theoretical mechanics, it is a fundamental observation that a broad class of mechanical systems may be modeled by equations of the form

$$(1.1) \quad \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right) = \frac{\partial L}{\partial x},$$

where  $L$  is a function of  $x$  and  $\dot{x}$ , called the Lagrangian, and  $x$  and  $\dot{x} = dx/dt$  are, respectively, generalized coordinates and generalized velocities. This observation yields an alternative approach for deriving the equations of motion and has many important consequences, including a convenient method for writing down equations of motion in curvilinear coordinate systems. From a philosophical perspective, (1.1)

\*Received by the editors May 29, 2002; accepted for publication (in revised form) July 3, 2003; published electronically February 2, 2004. This paper presents research results of the Belgian Programme on Inter-University Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with its authors. A preliminary version of the present study appeared as a conference paper [12].

<http://www.siam.org/journals/sirev/46-1/40902.html>

<sup>†</sup>Dynamics and Control Group, Mechanical Engineering Department, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands (L.P.M.Moreau@tue.nl). This author is a Postdoctoral Fellow of the Fund for Scientific Research—Flanders (Belgium).

<sup>‡</sup>EESA-SYSTEms, Ghent University, Technologiepark 914, 9052 Zwijnaarde, Belgium (dirk.aeyels@ugent.be).

has a remarkable conceptual interpretation: it is the well-known first-order necessary condition for a curve  $x$  to be a minimizer for the functional

$$(1.2) \quad \int_{t_0}^{t_1} L(x, \dot{x}) dt$$

on the space of curves joining  $x(t_0)$  at  $t_0$  with  $x(t_1)$  at  $t_1$ . Equation (1.1) is called the Euler–Lagrange equation for the functional (1.2). See, for example, [1] for a recent exposition of these topics.

Motivated by the success of Lagrangian mechanics, there have been several attempts to develop a similar theory for electric circuits; see, for example, [7, 4, 19]. The present paper aims to contribute to this line of research. Our interest in electric circuits is motivated by their important role in power electronics and electromechanical engineering. In addition, the class of electric circuits is representative of a much broader class of *network models*, with applications ranging from civil engineering to chemical engineering [15].

In the present article, we restrict our attention to inductor-capacitor circuits (LC circuits). The inductors and capacitors are allowed to be nonlinear, but they are assumed to be characterized by a single-valued function, thus excluding, for example, hysteresis effects. The analogy between electric LC circuits and mechanical mass-spring systems suggests magnetic coenergy minus electric energy as Lagrangian or, in a dual approach, electric coenergy minus magnetic energy. (Mixed formulations were considered in [4, 19].) The natural expression of these Lagrangians is in terms of capacitor charges and inductor currents or, in a dual approach, inductor fluxes and capacitor voltages, but these variables are not related by simple differentiation. Therefore, a central issue in formulating Euler–Lagrange equations for LC circuits is the selection of generalized coordinates and velocities in terms of which the Lagrangian may be re-expressed. Chua and McPherson [4] considered inductor currents or capacitor voltages as generalized velocities and their integrals as generalized coordinates. In that approach the generalized coordinates do not have a physical interpretation. Alternatively, Szatkowski [19] considered capacitor charges or inductor fluxes as generalized coordinates and their derivatives as generalized velocities. The approaches presented in these references are not applicable to general inductor-capacitor circuits: both references make restrictive assumptions about the topology of the circuit.

An alternative approach was proposed by Kwatny, Massimo, and Bahar [10]. In that reference, a generalized Euler–Lagrange formalism [13] was used that enabled the authors to consider generalized velocities that are not simply the derivatives of the generalized coordinates. This approach circumvents the above-mentioned problems, but its implementation seems to give rise to computations (involving pseudoinverses and Lagrange multipliers) that are significantly more complex than in the classical Lagrangian formalism. Although no variational interpretation was given in [10] and although the results in that reference bear little similarity to the present work, the basic idea expounded in that paper is important for the present study.

We associate a variational problem with the circuit that seeks to minimize the time integral of magnetic coenergy minus electric energy subject to Kirchhoff’s current law as a dynamic constraint—as well as some other constraints; see below for details. We interpret the variational problem as an optimal control problem and apply the Pontryagin maximum principle [14] to derive its normal Pontryagin extremals. This leads to a set of generalized Euler–Lagrange equations for the circuit. Some distinctive features of the present approach are that (i) it is applicable to nonlinear LC circuits

with arbitrary interconnection structure; (ii) its implementation is not significantly more complex than in the classical Lagrangian formalism; (iii) the variational formulation and Lagrangian equations of the present paper do not involve nonphysical inductor charges or capacitor fluxes, but are instead stated directly in terms of physically meaningful variables such as capacitor charges and inductor currents; and finally (iv) the generalized Lagrangian description of the present paper is fully compatible with the structural properties of the circuit (energy and topology).

A different line of ideas was put forward in a series of papers by van der Schaft, Maschke, and coworkers [11, 20]. (See also Bloch and Crouch [2].) These papers are not concerned with variational or Lagrangian models for LC circuits; instead they focus upon Hamiltonian modeling. It is recognized in these references that the dynamic equations of a nonlinear LC circuit obtained by algebraic elimination of the branch currents and voltages are of a Hamiltonian nature with respect to a *constant Dirac structure* on the vector space of capacitor charges and inductor fluxes. Our approach intersects this line of research: the generalized Euler–Lagrange equations that we obtain may be converted into the generalized Hamiltonian equations of [11, 20, 2] by means of a Legendre transformation, thus providing a variational interpretation for these Hamiltonian models.

This article is organized as follows. Sections 2 and 3 contain the theoretical foundations of the present work. The Lagrangian and Hamiltonian nature of LC circuits is discussed and a novel variational formalism is proposed. The language of (differential) geometry plays an important role in these sections: it provides a natural framework for a clear and concise formulation of the results. Sections 4 and 5 are more oriented toward applications. Section 4 discusses the novel variational formalism from a purely algorithmic point of view, and section 5 illustrates how the current approach may be applied to electromechanical systems. The paper is completed with a conclusion section and an appendix, where the maximum principle is reviewed.

We end this introduction with some references to other, related work. In a recent book by Jurdjevic [9], the classical variational principle for the heavy top is interpreted as an optimal control problem on the configuration manifold of the rigid body. A canonical system of Hamiltonian equations describing the dynamics of the heavy top is then obtained by applying the Pontryagin maximum principle on manifolds. An alternative approach to the variational modeling of electric circuits has, independently of the present research, been proposed by Scherpen, Jeltsema, and Klaassens [16] and Clemente-Gallardo and Scherpen in [5]. Whereas the present paper makes use of the Pontryagin maximum principle, the approach in these papers is based on the constrained variational principles from holonomic mechanics.

## 2. Preliminaries.

**2.1. Geometry.** In order to identify and characterize the Lagrangian and Hamiltonian structure of nonlinear electric circuits, it is convenient to use the language of (differential) geometry. For a precise definition of such concepts as dual vector spaces, natural pairing, differential, etc., we refer to [3]. Here we provide only an informal introduction, the main purpose of this subsection being to fix the notation. Throughout this subsection,  $U$ ,  $V$ , and  $W$  are finite-dimensional, real vector spaces. Throughout the paper, *smooth* means infinitely times continuously differentiable.

**2.1.1. Dual Vector Spaces and Natural Pairing.** A vector space  $V$  gives rise to a *dual vector space*  $V^*$ , consisting of all linear maps from  $V$  to  $\mathbb{R}$ . Although  $V$  and  $V^*$  have the same dimension, they are really two different vector spaces: there is no

natural (preferred) identification between elements of  $V$  and  $V^*$ .<sup>1</sup> There is, however, a natural identification between elements of  $(V^*)^*$  and  $V$ , and from now on we always identify  $(V^*)^* = V$ .

The image of a vector  $v \in V$  under a linear map  $w \in V^*$  is by definition a real number, which we denote either by  $\langle w, v \rangle$  or by  $\langle v, w \rangle$ . We refer to the operation  $\langle \cdot, \cdot \rangle$  as the *natural pairing* between elements of  $V$  and  $V^*$ .

Any basis for the vector space  $V$  gives rise to a corresponding *dual basis* for  $V^*$ . Dual bases have the following property (which may actually serve as a definition): if we represent  $v \in V$  and  $w \in V^*$  by column vectors with respect to dual bases, then the natural pairing  $\langle v, w \rangle$  corresponds to the inner product of these column vectors.

**2.1.2. Differentials of Real-Valued Functions.** Consider a smooth, nonlinear function  $f$  from  $V$  to  $\mathbb{R}$ . The linear approximation of  $f$  at  $v \in V$  is a linear map from  $V$  to  $\mathbb{R}$  or, equivalently, an element of  $V^*$ . We refer to this linear approximation as the *differential* of  $f$  at  $v$  and denote it by  $(\partial f / \partial v)(v)$ . If we introduce a basis for  $V$  and hence linear coordinates on  $V$ , then the components of  $(\partial f / \partial v)(v)$  with respect to the dual basis correspond to the partial derivatives of  $f$  with respect to the coordinate components on  $V$ .

Consider a smooth, nonlinear function  $f : V \times W \rightarrow \mathbb{R} : (v, w) \mapsto f(v, w)$ . The differential of  $f(\cdot, w) : V \rightarrow \mathbb{R}$  at  $v \in V$  is denoted by  $(\partial f / \partial v)(v, w)$ . The differential of  $f(v, \cdot) : W \rightarrow \mathbb{R}$  at  $w \in W$  is denoted by  $(\partial f / \partial w)(v, w)$ .

**2.1.3. Dual Linear Maps and Skew-Symmetry.** A linear map  $A$  from  $V$  to  $W$  gives rise to a *dual linear map*  $A^*$  from  $W^*$  to  $V^*$ . If we introduce bases for  $V$  and  $W$  and corresponding dual bases for  $W^*$  and  $V^*$ , then the matrix expressions of  $A$  and  $A^*$  with respect to these bases are related by transposition.

The dual of a linear map from  $V$  to  $V^*$  is again a linear map from  $V$  to  $V^*$ . Such a linear map is called *skew-symmetric* if it is equal to minus its dual. If we introduce a basis for  $V$  and a corresponding dual basis for  $V^*$ , then the matrix expression of a skew-symmetric map with respect to these bases is a skew-symmetric matrix.

**2.1.4. Combined Mappings.** Consider linear maps  $A$  and  $B$  from  $U$  to, respectively,  $V$  and  $W$ . The *combined mapping*  $(A, B)$  is a linear map from  $U$  to  $V \times W$  defined by

$$(2.1) \quad (A, B)(u) = (A(u), B(u)).$$

**2.2. Dirac Structures and Hamiltonian Systems.** Let  $V$  be an  $n$ -dimensional, real vector space ( $n \in \mathbb{N}$ ). Consider a skew-symmetric linear map  $\Omega^\# : V^* \rightarrow V$  and a smooth function  $H : V \rightarrow \mathbb{R}$ . This gives rise to a dynamical system on  $V$ :

$$(2.2) \quad \dot{v} = \Omega^\# \frac{\partial H}{\partial v}(v).$$

Here and in the remainder of the paper, an overdot indicates time derivation. We may rewrite (2.2) by introducing the graph  $\Omega$  of the map  $\Omega^\#$ , defined as the  $n$ -dimensional

<sup>1</sup>For some purposes, it may be convenient to think of elements of  $V$  as column vectors and elements of  $V^*$  as row vectors. Because this interpretation has its limitations, we will not use this analogy in this article.

subspace  $\{(\Omega^\#(\alpha), \alpha) : \alpha \in V^*\}$  of  $V \times V^*$ .<sup>2</sup> Equation (2.2) is then equivalent to

$$(2.3) \quad \left( \dot{v}, \frac{\partial H}{\partial v}(v) \right) \in \Omega.$$

It is well known that the skew-symmetry of  $\Omega^\#$  is equivalent to the following property for  $\Omega$ :<sup>3</sup>

$$(2.4) \quad \langle v, w \rangle = 0 \quad \forall (v, w) \in \Omega.$$

This leads to the following fundamental definitions.

DEFINITION 1 (constant Dirac structure [6, 8]). *A constant Dirac structure on an  $n$ -dimensional, real vector space  $V$  is an  $n$ -dimensional subspace  $\Omega \subset V \times V^*$  satisfying (2.4).*

DEFINITION 2 (Dirac Hamiltonian system [20]). *A Dirac Hamiltonian system is a triple  $(V, \Omega, H)$  with*

- (1)  $V$  an  $n$ -dimensional, real vector space;
- (2)  $\Omega$  a constant Dirac structure on  $V$ ;
- (3)  $H : V \rightarrow \mathbb{R}$  a smooth function.

*The dynamics of this Hamiltonian system are governed by equation (2.3).*

It is a fundamental property of Dirac Hamiltonian systems that the Hamiltonian itself is a conserved quantity:  $\dot{H} = \langle \dot{v}, \partial H / \partial v \rangle = 0$ . This follows readily from (2.3) and (2.4).

The above definitions allow for more general systems than those we initially had in mind, since  $\Omega$  need not be the graph of a linear map  $\Omega^\# : V^* \rightarrow V$ . If  $\Omega$  actually is the graph of a linear map  $\Omega^\#$ , then we refer to the triple  $(V, \Omega, H)$  as a *Poisson Hamiltonian system*. If in addition the map  $\Omega^\#$  is invertible, we say that  $(V, \Omega, H)$  is a *symplectic Hamiltonian system*, and for the particular case  $V = W \times W^*$  and  $\Omega^\# : W^* \times W \rightarrow W \times W^* : (\xi, x) \mapsto (x, -\xi)$  for some  $m$ -dimensional vector space  $W$ , the triple  $(V, \Omega, H)$  is called a *canonical symplectic Hamiltonian system*. In this case the dynamics are governed by the well-known set of equations

$$(2.5) \quad \dot{x} = \frac{\partial H}{\partial \xi}(x, \xi),$$

$$(2.6) \quad \dot{\xi} = -\frac{\partial H}{\partial x}(x, \xi).$$

REMARK 1. *The class of Dirac Hamiltonian systems considered in [20] is more general than that introduced in Definition 2. That reference studies Hamiltonian systems with respect to nonconstant Dirac structures on manifolds, whereas we restrict attention to constant Dirac structures on vector spaces.*

<sup>2</sup>The convention of regarding the graph of a function from  $V^*$  to  $V$  as a subset of  $V \times V^*$  (instead of  $V^* \times V$ ) is not standard but is consistent with the notation used in the literature on Dirac structures.

<sup>3</sup>By definition,  $\Omega^\#$  is skew-symmetric iff  $(\Omega^\#)^* + \Omega^\# = 0$ . We have

$$(2.a) \quad (\Omega^\#)^* + \Omega^\# = 0$$

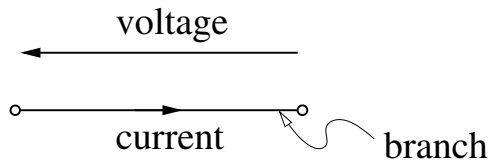
$$(2.b) \quad \Leftrightarrow \langle (\Omega^\#)^*(w), \tilde{w} \rangle + \langle \Omega^\#(w), \tilde{w} \rangle = 0 \quad \forall w, \tilde{w} \in V^*$$

$$(2.c) \quad \Leftrightarrow \langle \Omega^\#(\tilde{w}), w \rangle + \langle \Omega^\#(w), \tilde{w} \rangle = 0 \quad \forall w, \tilde{w} \in V^*$$

$$(2.d) \quad \Leftrightarrow \langle \tilde{v}, w \rangle + \langle v, \tilde{w} \rangle = 0 \quad \forall (v, w), (\tilde{v}, \tilde{w}) \in \Omega$$

$$(2.e) \quad \Leftrightarrow \langle v + \tilde{v}, w + \tilde{w} \rangle - \langle v, w \rangle - \langle \tilde{v}, \tilde{w} \rangle = 0 \quad \forall (v, w), (\tilde{v}, \tilde{w}) \in \Omega.$$

It is not difficult to see that (2.e) is equivalent to (2.4).



**Fig. 1** Associated reference directions.

**2.3. Circuit Variables.** We introduce a natural and consistent notational scheme for writing down the equations of an electric circuit. The proposed notation may be seen as a combination of the matrix notation from [7, 4, 19, 10] and the abstract vector space approach from [11, 20, 2]. The main features of the present notation are that (i) it enables us to assign variables (current, voltage, charge, flux) directly to the branches of a circuit—without having to number the branches first; and (ii) it takes full account of the natural pairing between branch currents and branch voltages.

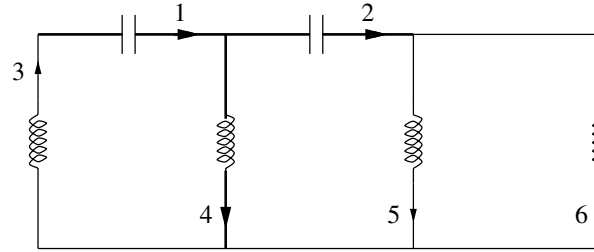
It is natural to consider the current and the voltage associated with one particular branch as elements of dual one-dimensional vector spaces. In this way, only the pairing of the current and the voltage is mathematically well defined—it corresponds to the physical notion of power—and this convention prohibits us from accidentally multiplying the current (or the voltage) with itself, which would be physically meaningless. We consider the current as an element of  $R$  and the voltage as an element of  $R^*$ . Here, and in the remainder of the article,  $R$  denotes the set of real numbers equipped (only) with its natural vector space structure.<sup>4</sup>

In order to assign currents and voltages directly to the branches, we introduce the following notation and terminology. Consider the graph of a lumped circuit and let  $\mathcal{B}$  be the collection of its branches. Consider subsets  $\mathcal{M} \subset \mathcal{N} \subset \mathcal{B}$ . An  $\mathcal{N}$ -vector is a map from  $\mathcal{N}$  to  $R$ , that is, a map that associates with every branch in  $\mathcal{N}$  an element of  $R$ . The set of all  $\mathcal{N}$ -vectors carries a natural vector space structure and is denoted by  $R_{\mathcal{N}}$ . Similarly, an  $\mathcal{N}$ -covector is a map from  $\mathcal{N}$  to  $R^*$  and the set of all  $\mathcal{N}$ -covectors equipped with its natural vector space structure is denoted by  $(R^*)_{\mathcal{N}}$ . Notice that there is a natural pairing between elements of  $R_{\mathcal{N}}$  and  $(R^*)_{\mathcal{N}}$ . In other words,  $(R^*)_{\mathcal{N}}$  is naturally identifiable with  $(R_{\mathcal{N}})^*$  and we may thus unambiguously write  $R_{\mathcal{N}}^*$ . For simplicity of notation, if  $\mathcal{N} = \{j\}$ , we write  $R_j$  (resp.,  $R_j^*$ ) instead of  $R_{\{j\}}$  (resp.,  $R_{\{j\}}^*$ ). For a given  $\mathcal{N}$ -(co)vector  $x$ , the  $\mathcal{M}$ -(co)vector  $x_{\mathcal{M}}$  is defined as the restriction of the map  $x$  to the domain  $\mathcal{M}$ . For simplicity, we write  $x_j$  instead of  $x_{\{j\}}$ .

In order to illustrate the notation, we consider a dynamic circuit consisting of two-terminal capacitors and inductors. Let  $\mathcal{B}$  be the collection of its branches. We denote by  $\mathcal{C}$  (resp.,  $\mathcal{L}$ ) the collection of all capacitor (resp., inductor) branches. To each branch we assign a current and voltage reference direction. Here and in the remainder of the paper, we choose *associated reference directions*<sup>5</sup> (Figure 1). Let  $i$  be the  $\mathcal{B}$ -vector of branch currents and  $v$  the  $\mathcal{B}$ -covector of branch voltages. With the present notation we have, for example, that the voltage over a particular branch

<sup>4</sup>Because  $R$  is only equipped with a vector space structure (and not with a metric structure), there is no natural identification between  $R$  and its dual  $R^*$ .

<sup>5</sup>Associated reference directions are characterized by the following property: the power delivered to a two-terminal element is given by the pairing of the branch voltage and the branch current of that element.



**Fig. 2** *LC circuit with algebraic constraint on inductor currents.*

$j \in \mathcal{B}$  is given by  $v_j$  and that the total power delivered to all capacitors is given by  $\langle v_C, i_C \rangle$ .

**2.4. Kirchhoff's Laws.** Three types of equations determine the dynamical behavior of an electric circuit: those arising from Kirchhoff's current law, those arising from Kirchhoff's voltage law, and the element equations. A fundamental feature of Kirchhoff's laws is that they depend solely on the topology of the circuit; i.e., they hold regardless of the nature of the elements. They may be expressed as follows. Let  $\mathcal{B}$  denote the collection of circuit branches. To each branch we assign a current reference direction. Let  $i$  be the  $\mathcal{B}$ -vector of branch currents and  $v$  the  $\mathcal{B}$ -covector of branch voltages. Then Kirchhoff's laws may be written as

$$(2.7) \quad i \in \text{Image } A,$$

$$(2.8) \quad v \in \text{Kernel } A^*,$$

where  $A$  is a full rank linear map from a vector space  $U$  to  $R_{\mathcal{B}}$  characterizing the topology of the circuit and the chosen current reference directions. Although  $U$  may be an abstract vector space without any physical interpretation, it is convenient to think of its elements as vectors of independent branch currents. Methods for selecting the vector space  $U$  and constructing the linear map  $A$  are well known in the literature on electric circuit theory. Here we outline one possible approach, which is based on the notion of tree and cotree [17]. A *tree* is a connected subgraph of the (connected) network graph which contains all the vertices of the graph but does not contain any loops. The complement of a tree is a *cotree*. It is a fundamental result in circuit theory that the cotree branch currents are independent and determine all other currents in the circuit. We may thus choose  $U$  to be the vector space of cotree branch currents. The linear map  $A$  then follows readily from inspection of the circuit graph.

**EXAMPLE 1.** Consider the LC circuit of Figure 2. Each branch has been assigned a reference number and a current reference direction. Branches 1, 2, and 4 constitute a tree. Hence the cotree branch currents  $i_3$ ,  $i_5$ , and  $i_6$  are independent and determine all other branch currents. Kirchhoff's laws may be written as

$$(2.9) \quad \begin{pmatrix} i_1 \\ i_2 \\ i_3 \\ i_4 \\ i_5 \\ i_6 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} i_3 \\ i_5 \\ i_6 \end{pmatrix},$$

$$(2.10) \quad \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{pmatrix} = 0.$$

It is clear from (2.7)–(2.8) that  $\langle v, i \rangle = 0$  for all possible voltage and current vectors that are consistent with Kirchhoff's laws. Stated in terms of the geometric language of Dirac structures, we have the following result, which is a paraphrase of Tellegen's theorem.

**THEOREM 1 (Tellegen).** *The subspace of  $R_{\mathcal{B}} \times R_{\mathcal{B}}^*$  consisting of all pairs  $(i, v)$  that satisfy Kirchhoff's laws is a constant Dirac structure on  $R_{\mathcal{B}}$ .*

Theorem 1 implies that the interconnection structure of an electric circuit is lossless. The Dirac structure featuring in this theorem will play a central role in the Lagrangian and Hamiltonian description of electric circuits. We refer to it as the *Tellegen Dirac structure*.

**2.5. Classical Lagrangian and Hamiltonian Formalism.** In this subsection, we review the classical Lagrangian and Hamiltonian formalism from the modern perspective of optimal control theory. We show how this modern perspective, when applied to the classical problem of variational calculus, leads to the standard and well-known models of Lagrangian and Hamiltonian dynamics. As we shall see in section 3, the main advantage of this modern perspective is that it lends itself naturally to the consideration of more general variational problems, leading to generalized Lagrangian and Hamiltonian models that are useful for the modeling of electric circuits. The discussion of the present section is based on ideas put forward by Sussmann and Willems in [18].

**2.5.1. The Optimal Control Problem.** Let  $X$  be a finite-dimensional, real vector space and  $L$  a smooth function from  $X \times X$  to  $\mathbb{R}$ . The classical problem of the calculus of variations, which plays a central role in Lagrangian mechanics, may be reformulated as an optimal control problem as follows. Let  $t_0 < t_1 \in \mathbb{R}$  and  $x_0, x_1 \in X$  be given. Among all piecewise continuous controls  $u : [t_0, t_1] \rightarrow X$  that steer  $x(t_0) = x_0$  to  $x(t_1) = x_1$  according to the dynamics

$$(2.11) \quad \dot{x} = u,$$

find the particular control that minimizes the cost functional

$$(2.12) \quad \int_{t_0}^{t_1} L(x, u) \, dt.$$

**2.5.2. Applying the Maximum Principle.** We apply the *Pontryagin maximum principle* to the above-mentioned optimal control problem (see the appendix for background and terminology). Here and in the remainder of the paper, we restrict our attention to the normal Pontryagin extremals. We also replace the maximum condition of the maximum principle by its first-order necessary condition. Having said this, it is important to notice that we do not make any claim about possible optimality properties of the resulting extremals. Rather, we use the ingredients of the maximum principle as an algorithm for deriving Lagrangian and Hamiltonian models.



The normal Pontryagin extremals of the optimal control problem above satisfy the control Hamiltonian equations

$$(2.13) \quad \dot{x} = u,$$

$$(2.14) \quad \dot{\xi} = \frac{\partial L}{\partial x}(x, u),$$

$$(2.15) \quad \frac{\partial L}{\partial u}(x, u) = \xi$$

associated with the control Hamiltonian

$$(2.16) \quad \mathcal{H} : X \times X^* \times X \rightarrow \mathbb{R} : (x, \xi, u) \mapsto \mathcal{H}(x, \xi, u) = -L(x, u) + \langle \xi, u \rangle.$$

**2.5.3. The Euler–Lagrange Equations.** Elimination of  $\xi$  from (2.13)–(2.15) leads to

$$(2.17) \quad \dot{x} = u,$$

$$(2.18) \quad \frac{d}{dt} \left( \frac{\partial L}{\partial u}(x, u) \right) = \frac{\partial L}{\partial x}(x, u),$$

which is equivalent to the classical Euler–Lagrange equation (1.1).

**2.5.4. The Legendre Transformation and Hamiltonian Description.** If the Legendre transformation

$$(2.19) \quad X \times X \rightarrow X \times X^* : (x, u) \mapsto (x, \xi) = \left( x, \frac{\partial L}{\partial u}(x, u) \right)$$

is a smooth, global diffeomorphism, then it converts the Euler–Lagrange equations (2.17)–(2.18) into the canonical symplectic Hamiltonian system

$$(2.20) \quad \dot{x} = \frac{\partial H}{\partial \xi}(x, \xi),$$

$$(2.21) \quad \dot{\xi} = -\frac{\partial H}{\partial x}(x, \xi)$$

with Hamiltonian  $H(x, \xi) = -L(x, u(x, \xi)) + \langle \xi, u(x, \xi) \rangle$ , where  $u(x, \xi)$  is obtained by inverting (2.19).

REMARK 2. *Alternatively, (2.20)–(2.21) may be obtained directly from the control Hamiltonian equations (2.13)–(2.15) by elimination of  $u$ .*

**3. LC Circuits.** In the present section, we study the dynamical equations of nonlinear inductor-capacitor circuits. We associate an optimal control problem with the circuit that seeks the minimization of the time integral of magnetic coenergy minus electric energy subject to Kirchhoff’s current law as a dynamic constraint—as well as some other constraints. This leads to a new, generalized Lagrangian description for the dynamics of the circuit. Under appropriate invertibility conditions, this Lagrangian description may be converted into the Dirac Hamiltonian model from [11, 20, 2] by means of a Legendre transformation. The present approach thus provides an alternative method for deriving the dynamical equations of an LC circuit, by means of a variational principle. Unlike most other variational approaches mentioned in the literature, the present variational formulation and the resulting equations are formulated directly in terms of physically meaningful variables.

Consider a dynamic circuit consisting of two-terminal capacitors and inductors, and let  $\mathcal{B}$  be the collection of its branches. Denote by  $\mathcal{C}$  (resp.,  $\mathcal{L}$ ) the collection of all capacitor (resp., inductor) branches. To each branch we assign a current reference direction. Let  $i$  be the  $\mathcal{B}$ -vector of branch currents,  $v$  the  $\mathcal{B}$ -covector of branch voltages,  $q$  the  $\mathcal{C}$ -vector of capacitor charges, and  $\phi$  the  $\mathcal{L}$ -covector of inductor fluxes.<sup>6</sup> We make the following element assumption.

ASSUMPTION 1. *Each capacitor is charge-controlled:*

$$(3.1) \quad v_j = S_j(q_j), \quad j \in \mathcal{C},$$

with  $S_j : R_j \rightarrow R_j^*$  a smooth function satisfying  $S_j(0) = 0$ . Each inductor is current-controlled:

$$(3.2) \quad \phi_j = L_j(i_j), \quad j \in \mathcal{L},$$

with  $L_j : R_j \rightarrow R_j^*$  a smooth function satisfying  $L_j(0) = 0$ .

Associated with the capacitors and inductors we introduce the electric energy  $W_e$  and magnetic coenergy  $W_m^*$  defined by

$$(3.3) \quad W_e : R_{\mathcal{C}} \rightarrow \mathbb{R} : q_{\mathcal{C}} \mapsto W_e(q_{\mathcal{C}}) = \sum_{j \in \mathcal{C}} \int_0^{q_j} S_j(s) \, ds,$$

$$(3.4) \quad W_m^* : R_{\mathcal{L}} \rightarrow \mathbb{R} : i_{\mathcal{L}} \mapsto W_m^*(i_{\mathcal{L}}) = \sum_{j \in \mathcal{L}} \int_0^{i_j} L_j(s) \, ds.$$

Kirchhoff's laws were introduced in (2.7)–(2.8) in section 2.3. For the purpose of the present section, it is convenient to write Kirchhoff's current law as

$$(3.5) \quad i_{\mathcal{C}} = A_{\mathcal{C}}u,$$

$$(3.6) \quad i_{\mathcal{L}} = A_{\mathcal{L}}u,$$

where  $A_{\mathcal{C}}$  and  $A_{\mathcal{L}}$  are linear maps from  $U$  to  $R_{\mathcal{C}}$ , respectively,  $R_{\mathcal{L}}$ , related to  $A$  by the equality

$$(3.7) \quad A = (A_{\mathcal{C}}, A_{\mathcal{L}}).$$

In the above formulation  $u \in U$  has been introduced as a new variable. As mentioned before, it is convenient to think of  $u$  as a vector of independent branch currents.

**3.1. The Optimal Control Problem.** We introduce the Lagrangian

$$(3.8) \quad L : R_{\mathcal{C}} \times R_{\mathcal{L}} \rightarrow \mathbb{R} : (q_{\mathcal{C}}, i_{\mathcal{L}}) \mapsto L(q_{\mathcal{C}}, i_{\mathcal{L}}) = W_m^*(i_{\mathcal{L}}) - W_e(q_{\mathcal{C}})$$

and associate the following optimal control problem with the circuit. Let  $t_0 < t_1 \in \mathbb{R}$ ,  $q_{\mathcal{C}0}, q_{\mathcal{C}1} \in R_{\mathcal{C}}$ , and  $\lambda \in U$  be given. Among all piecewise continuous controls  $u : [t_0, t_1] \rightarrow U$  that steer  $q_{\mathcal{C}}(t_0) = q_{\mathcal{C}0}$  to  $q_{\mathcal{C}}(t_1) = q_{\mathcal{C}1}$  according to the dynamics

$$(3.9) \quad \dot{q}_{\mathcal{C}} = A_{\mathcal{C}}u$$

<sup>6</sup>When we write down differential equations for these variables, time derivation is indicated by an overdot, except for the case of current variables  $i$ , whose time derivative is denoted by  $di/dt$ . No confusion should arise.

while satisfying

$$(3.10) \quad \int_{t_0}^{t_1} u \, dt = \lambda,$$

find the particular control for which the cost functional

$$(3.11) \quad \int_{t_0}^{t_1} L(q_C, A_{\mathcal{L}} u) \, dt$$

is minimized. We study this optimal control problem for all possible values of  $t_0$ ,  $t_1$ ,  $q_{C0}$ ,  $q_{C1}$ , and  $\lambda$ .

EXAMPLE 2 (continued from Example 1). *We continue our study of the LC circuit of Figure 2. For simplicity of the exposition, we assume that the constitutive relations for the capacitors and inductors are linear and bijective; they are, respectively, characterized by their nonzero capacitance ( $C_1$  and  $C_2$ ) and nonzero inductance ( $L_3$ ,  $L_4$ ,  $L_5$ , and  $L_6$ ). The considered optimal control problem seeks the minimization of the time integral of magnetic coenergy minus electric energy*

$$(3.12) \quad \int_{t_0}^{t_1} \left\{ \frac{1}{2} L_3 i_3^2 + \frac{1}{2} L_4 (i_3 - i_5 - i_6)^2 + \frac{1}{2} L_5 i_5^2 + \frac{1}{2} L_6 i_6^2 - \frac{1}{2C_1} q_1^2 - \frac{1}{2C_2} q_2^2 \right\} dt$$

subject to the dynamics

$$(3.13) \quad \dot{q}_1 = i_3, \quad \dot{q}_2 = i_5 + i_6$$

on a fixed time interval  $[t_0, t_1]$  with fixed terminal points  $q_1(t_0)$ ,  $q_2(t_0)$ ,  $q_1(t_1)$ , and  $q_2(t_1)$  and integral constraints on the inputs

$$(3.14) \quad \int_{t_0}^{t_1} i_3 \, dt = \lambda_3, \quad \int_{t_0}^{t_1} i_5 \, dt = \lambda_5, \quad \int_{t_0}^{t_1} i_6 \, dt = \lambda_6.$$

In the following subsections, we apply the Pontryagin maximum principle to this optimal control problem and derive a Lagrangian and Hamiltonian description for the circuit. In doing so, we try to proceed along the lines of section 2.5. In particular, we require that, for the special case of both  $A_C$  and  $A_{\mathcal{L}}$  being the identity map, the models obtained coincide with the classical Lagrangian and Hamiltonian models discussed in that section. (For the particular case of both  $A_C$  and  $A_{\mathcal{L}}$  being the identity map, the present optimal control problem essentially reduces to the classical variational problem of section 2.5; see the comment at the end of section 3.5.2.)

**3.2. Applying the Maximum Principle.** In order to characterize the normal Pontryagin extremals, we consider the control Hamiltonian with abnormal multiplier equal to  $-1$ :

$$(3.15) \quad \mathcal{H}^\zeta : R_C \times R_C^* \times U \rightarrow \mathbb{R} : (q_C, \xi_C, u) \mapsto -L(q_C, A_{\mathcal{L}} u) + \langle \xi_C, A_C u \rangle + \langle \zeta, u \rangle,$$

where  $\zeta \in U^*$  is a parameter covector. The associated control Hamiltonian equations are

$$(3.16) \quad \dot{q}_C = A_C u,$$

$$(3.17) \quad \dot{\xi}_C = \frac{\partial L}{\partial q_C}(q_C, A_{\mathcal{L}} u),$$

$$(3.18) \quad A_{\mathcal{L}}^* \frac{\partial L}{\partial i_{\mathcal{L}}}(q_C, A_{\mathcal{L}} u) = A_C^* \xi_C + \zeta.$$

With the introduction of  $i_{\mathcal{L}} = A_{\mathcal{L}}u$  and the subsequent elimination of  $u$  these equations may be rewritten as

$$(3.19) \quad (q_{\mathcal{C}}, i_{\mathcal{L}}) \in \text{Image}(A_{\mathcal{C}}, A_{\mathcal{L}}),$$

$$(3.20) \quad \dot{\xi}_{\mathcal{C}} = \frac{\partial L}{\partial q_{\mathcal{C}}}(q_{\mathcal{C}}, i_{\mathcal{L}}),$$

$$(3.21) \quad A_{\mathcal{L}}^* \frac{\partial L}{\partial i_{\mathcal{L}}}(q_{\mathcal{C}}, i_{\mathcal{L}}) = A_{\mathcal{C}}^* \xi_{\mathcal{C}} + \zeta.$$

EXAMPLE 3 (continued from Example 2). *We continue our study of the LC circuit of Figure 2. The control Hamiltonian with abnormal multiplier equal to  $-1$  is given by*

$$(3.22) \quad \begin{aligned} \mathcal{H}^{\zeta_3, \zeta_5, \zeta_6}(q_1, q_2, \xi_1, \xi_2, i_3, i_5, i_6) \\ = -\frac{1}{2}L_3 i_3^2 - \frac{1}{2}L_4(i_3 - i_5 - i_6)^2 - \frac{1}{2}L_5 i_5^2 - \frac{1}{2}L_6 i_6^2 + \frac{1}{2C_1}q_1^2 + \frac{1}{2C_2}q_2^2 \\ + \xi_1 i_3 + \xi_2(i_5 + i_6) + \zeta_3 i_3 + \zeta_5 i_5 + \zeta_6 i_6. \end{aligned}$$

The corresponding control Hamiltonian equations are

$$(3.23) \quad \begin{aligned} \dot{q}_1 &= i_3, \\ \dot{q}_2 &= i_5 + i_6, \end{aligned}$$

$$(3.24) \quad \begin{aligned} \dot{\xi}_1 &= -\frac{q_1}{C_1}, \\ \dot{\xi}_2 &= -\frac{q_2}{C_2}, \end{aligned}$$

$$(3.25) \quad \begin{aligned} L_3 i_3 + L_4(i_3 - i_5 - i_6) &= \xi_1 + \zeta_3, \\ -L_4(i_3 - i_5 - i_6) + L_5 i_5 &= \xi_2 + \zeta_5, \\ -L_4(i_3 - i_5 - i_6) + L_6 i_6 &= \xi_2 + \zeta_6. \end{aligned}$$

Introduction of  $i_4 = i_3 - i_5 - i_6$  yields

$$(3.26) \quad \begin{aligned} \dot{q}_1 &= i_3, \\ \dot{q}_2 &= i_5 + i_6, \\ i_4 &= i_3 - i_5 - i_6, \end{aligned}$$

$$(3.27) \quad \begin{aligned} \dot{\xi}_1 &= -\frac{q_1}{C_1}, \\ \dot{\xi}_2 &= -\frac{q_2}{C_2}, \end{aligned}$$

$$(3.28) \quad \begin{aligned} L_3 i_3 + L_4 i_4 &= \xi_1 + \zeta_3, \\ -L_4 i_4 + L_5 i_5 &= \xi_2 + \zeta_5, \\ -L_4 i_4 + L_6 i_6 &= \xi_2 + \zeta_6. \end{aligned}$$

**3.3. The Generalized Euler–Lagrange Equations.** Elimination of  $\xi_{\mathcal{C}}$  from (3.19)–(3.21) yields

$$(3.29) \quad (q_{\mathcal{C}}, i_{\mathcal{L}}) \in \text{Image}(A_{\mathcal{C}}, A_{\mathcal{L}}),$$

$$(3.30) \quad \left( -\frac{\partial L}{\partial q_{\mathcal{C}}}(q_{\mathcal{C}}, i_{\mathcal{L}}), \frac{d}{dt} \left( \frac{\partial L}{\partial i_{\mathcal{L}}}(q_{\mathcal{C}}, i_{\mathcal{L}}) \right) \right) \in \text{Kernel}(A_{\mathcal{C}}, A_{\mathcal{L}})^*.$$

These equations reduce to the classical Euler–Lagrange equations if  $A_C$  and  $A_L$  are identity maps. We refer to (3.29)–(3.30) as *generalized Euler–Lagrange equations*, a notion that is made more precise by the following definition.

DEFINITION 3. Consider a triple  $(X \times Y, \Omega, L)$  with

- (1)  $X$  and  $Y$  finite-dimensional, real vector spaces;
- (2)  $\Omega$  a constant Dirac structure on  $X \times Y$ ;
- (3)  $L : X \times Y \rightarrow \mathbb{R} : (x, y) \mapsto L(x, y)$  a smooth function.

The generalized Euler–Lagrange equations on  $X \times Y$  with Lagrangian  $L$  and with respect to the Dirac structure  $\Omega$  are the differential and algebraic equations

$$(3.31) \quad \left( \dot{x}, y, -\frac{\partial L}{\partial x}, \frac{d}{dt} \left( \frac{\partial L}{\partial y} \right) \right) \in \Omega.$$

According to this definition, (3.29)–(3.30) are the generalized Euler–Lagrange equations on  $R_C \times R_L$  with Lagrangian equal to magnetic coenergy minus electric energy and with respect to the Tellegen Dirac structure. They describe the dynamical behavior of the circuit as follows.

THEOREM 2. If Assumption 1 is satisfied, then the dynamics of the LC circuit are governed by the generalized Euler–Lagrange equations on  $R_C \times R_L$  with Lagrangian equal to magnetic coenergy minus electric energy and with respect to the Tellegen Dirac structure.

*Proof.* The dynamics of the circuit are governed by three types of equations: the element equations, the equations arising from Kirchhoff’s current law, and those arising from Kirchhoff’s voltage law. It is clear from (3.1)–(3.4) that the element equations of the capacitors and inductors may be written as

$$(3.32) \quad v_C = \frac{\partial W_e}{\partial q_C}(q_C),$$

$$(3.33) \quad \dot{q}_C = i_C$$

and

$$(3.34) \quad \phi_L = \frac{\partial W_m^*}{\partial i_L}(i_L),$$

$$(3.35) \quad \dot{\phi}_L = v_L,$$

respectively. According to (3.5)–(3.6) Kirchhoff’s laws are given by

$$(3.36) \quad (i_C, i_L) \in \text{Image}(A_C, A_L),$$

$$(3.37) \quad (v_C, v_L) \in \text{Kernel}(A_C, A_L)^*.$$

By algebraic elimination, (3.32)–(3.37) may be reduced to a closed system of differential and algebraic equations for  $q_C$  and  $i_L$ :

$$(3.38) \quad (\dot{q}_C, i_L) \in \text{Image}(A_C, A_L),$$

$$(3.39) \quad \left( \frac{\partial W_e}{\partial q_C}(q_C), \frac{d}{dt} \left( \frac{\partial W_m^*}{\partial i_L}(i_L) \right) \right) \in \text{Kernel}(A_C, A_L)^*.$$

It is clear from the definition of the Lagrangian (3.8) that (3.38)–(3.39) are equivalent to (3.29)–(3.30).  $\square$

EXAMPLE 4 (continued from Example 3). *We continue our study of the LC circuit of Figure 2. Elimination of the  $\xi$  variables from (3.26)–(3.28) yields the generalized Euler–Lagrange equations*

$$\begin{aligned}
 (3.40) \quad & \dot{q}_1 = i_3, \\
 & \dot{q}_2 = i_5 + i_6, \\
 & i_4 = i_3 - i_5 - i_6, \\
 (3.41) \quad & L_3 \frac{di_3}{dt} + L_4 \frac{di_4}{dt} = -\frac{q_1}{C_1}, \\
 & -L_4 \frac{di_4}{dt} + L_5 \frac{di_5}{dt} = -\frac{q_2}{C_2}, \\
 & -L_4 \frac{di_4}{dt} + L_6 \frac{di_6}{dt} = -\frac{q_2}{C_2},
 \end{aligned}$$

which govern the dynamics of the circuit.

**3.4. The Legendre Transformation and Dirac Hamiltonian Description.** We start this subsection with a review of the Dirac Hamiltonian description of LC circuits from [11, 20, 2]. We make an extra element assumption.

ASSUMPTION 2. *Each function  $L_j$  ( $j \in \mathcal{L}$ ) is a smooth diffeomorphism.*

This assumption implies that the inductors are also flux-controlled:  $i_j = \Gamma_j(\phi_j)$  ( $j \in \mathcal{L}$ ) with  $\Gamma_j : R_j^* \rightarrow R_j$  a smooth function, the inverse of  $L_j$ , satisfying  $\Gamma_j(0) = 0$ . We introduce the *magnetic energy*  $W_m$ ,

$$(3.42) \quad W_m : R_{\mathcal{L}}^* \rightarrow \mathbb{R} : \phi_{\mathcal{L}} \mapsto W_m(\phi_{\mathcal{L}}) = \sum_{j \in \mathcal{L}} \int_0^{\phi_j} \Gamma_j(s) ds,$$

which is related to  $W_m^*$  by a Legendre transformation. This enables us to introduce the Hamiltonian

$$(3.43) \quad H : R_{\mathcal{C}} \times R_{\mathcal{L}}^* \rightarrow \mathbb{R} : (q_{\mathcal{C}}, \phi_{\mathcal{L}}) \mapsto W_e(q_{\mathcal{C}}) + W_m(\phi_{\mathcal{L}}).$$

It is recognized in [11, 20, 2] that the dynamics of the LC circuit are governed by the equations

$$(3.44) \quad \left( \dot{q}_{\mathcal{C}}, \frac{\partial H}{\partial \phi_{\mathcal{L}}}(q_{\mathcal{C}}, \phi_{\mathcal{L}}) \right) \in \text{Image}(A_{\mathcal{C}}, A_{\mathcal{L}}),$$

$$(3.45) \quad \left( \frac{\partial H}{\partial q_{\mathcal{C}}}(q_{\mathcal{C}}, \phi_{\mathcal{L}}), \dot{\phi}_{\mathcal{L}} \right) \in \text{Kernel}(A_{\mathcal{C}}, A_{\mathcal{L}})^*,$$

and that these equations form a Dirac Hamiltonian system. In these references, equations (3.44)–(3.45) are obtained directly from the element equations and Kirchhoff's laws by algebraic elimination. We show that the Dirac Hamiltonian description (3.44)–(3.45) may be obtained from the Lagrangian description (3.29)–(3.30) by means of a Legendre transformation. This provides an alternative method for deriving the Dirac Hamiltonian system (3.44)–(3.45).

THEOREM 3. *If Assumptions 1 and 2 are satisfied, then the Legendre transformation*

$$(3.46) \quad R_{\mathcal{C}} \times R_{\mathcal{L}} \rightarrow R_{\mathcal{C}} \times R_{\mathcal{L}}^* : (q_{\mathcal{C}}, i_{\mathcal{L}}) \mapsto (q_{\mathcal{C}}, \phi_{\mathcal{L}}) = \left( q_{\mathcal{C}}, \frac{\partial W_m^*}{\partial i_{\mathcal{L}}}(i_{\mathcal{L}}) \right)$$

is a smooth, global diffeomorphism which converts the generalized Euler–Lagrange equations (3.29)–(3.30) into the Dirac Hamiltonian system (3.44)–(3.45).

Theorem 3 is proven below.

EXAMPLE 5 (continued from Example 4). We continue our study of the LC circuit of Figure 2. The generalized Euler–Lagrange equations (3.40)–(3.41) may be converted into a Dirac Hamiltonian system by means of a Legendre transformation. Introduction of  $\phi_j = L_j i_j$  and subsequent elimination of  $i_j$  ( $j = 3, 4, 5, 6$ ) yields the Dirac Hamiltonian system

$$\begin{aligned}
 (3.47) \quad & \dot{q}_1 = \frac{\phi_3}{L_3}, \\
 & \dot{q}_2 = \frac{\phi_5}{L_5} + \frac{\phi_6}{L_6}, \\
 & \frac{\phi_4}{L_4} = \frac{\phi_3}{L_3} - \frac{\phi_5}{L_5} - \frac{\phi_6}{L_6}, \\
 (3.48) \quad & \dot{\phi}_3 + \dot{\phi}_4 = -\frac{q_1}{C_1}, \\
 & -\dot{\phi}_4 + \dot{\phi}_5 = -\frac{q_2}{C_2}, \\
 & -\dot{\phi}_4 + \dot{\phi}_6 = -\frac{q_2}{C_2},
 \end{aligned}$$

which governs the dynamics of the circuit.

We end this subsection with a proof of Theorem 3. It is obtained as a consequence of the following more general result.

THEOREM 4. Consider finite-dimensional, real vector spaces  $X$  and  $Y$ , a constant Dirac structure  $\Omega$  on  $X \times Y$ , and a smooth function  $L : X \times Y \rightarrow \mathbb{R} : (x, y) \mapsto L(x, y)$ . If the Legendre transformation

$$(3.49) \quad X \times Y \rightarrow X \times Y^* : (x, y) \mapsto (x, z) = \left( x, \frac{\partial L}{\partial y}(x, y) \right)$$

is a smooth, global diffeomorphism, then it converts the generalized Euler–Lagrange system  $(X \times Y, \Omega, L)$  into the Dirac Hamiltonian system  $(X \times Y^*, \tilde{\Omega}, H)$ , where

- (1)  $\tilde{\Omega}$  is the subspace of  $X \times Y^* \times X^* \times Y$  corresponding to the subspace  $\Omega$  of  $X \times Y \times X^* \times Y^*$  under the natural identification of  $X \times Y^* \times X^* \times Y$  with  $X \times Y \times X^* \times Y^*$  (via permutation of the second and fourth components);
- (2)  $H : X \times Y^* \rightarrow \mathbb{R} : (x, z) \mapsto H(x, z) = -L(x, y(x, z)) + \langle z, y(x, z) \rangle$ , where  $y(x, z)$  is obtained by inverting (3.49).

*Proof* (of Theorem 4). First, notice that the triple  $(X \times Y^*, \tilde{\Omega}, H)$  actually is a Dirac Hamiltonian system. Indeed, it is easy to see that  $\tilde{\Omega}$  as introduced in the theorem is a constant Dirac structure on  $X \times Y^*$ .

The generalized Euler–Lagrange equations corresponding to the triple  $(X \times Y, \Omega, L)$  are given by

$$(3.50) \quad \left( \dot{x}, y, -\frac{\partial L}{\partial x}, \frac{d}{dt} \left( \frac{\partial L}{\partial y} \right) \right) \in \Omega,$$

or, via permutation of the second and fourth components, by

$$(3.51) \quad \left( \dot{x}, \frac{d}{dt} \left( \frac{\partial L}{\partial y} \right), -\frac{\partial L}{\partial x}, y \right) \in \tilde{\Omega}.$$

From the definitions of the Legendre transformation (3.49) and the Hamiltonian  $H$ , it is clear that (3.51) may be rewritten as

$$(3.52) \quad \left( \dot{x}, \dot{z}, \frac{\partial H}{\partial x}, \frac{\partial H}{\partial z} \right) \in \tilde{\Omega},$$

which are the equations corresponding to the Dirac Hamiltonian system  $(X \times Y^*, \tilde{\Omega}, H)$ .  $\square$

*Proof* (of Theorem 3). First, notice that if Assumption 2 is satisfied, then the Legendre transformation (3.46) is indeed a smooth, global diffeomorphism with inverse

$$(3.53) \quad R_{\mathcal{C}} \times R_{\mathcal{L}}^* \rightarrow R_{\mathcal{C}} \times R_{\mathcal{L}} : (q_{\mathcal{C}}, \phi_{\mathcal{L}}) \mapsto (q_{\mathcal{C}}, i_{\mathcal{L}}) = \left( q_{\mathcal{C}}, \frac{\partial W_{\mathbf{m}}}{\partial \phi_{\mathcal{L}}}(\phi_{\mathcal{L}}) \right).$$

In view of Theorem 4, it only remains to verify that the Hamiltonian (3.43) and the Lagrangian (3.8) are related by

$$(3.54) \quad H(q_{\mathcal{C}}, \phi_{\mathcal{L}}) = -L(q_{\mathcal{C}}, i_{\mathcal{L}}) + \langle \phi_{\mathcal{L}}, i_{\mathcal{L}} \rangle$$

when  $i_{\mathcal{L}}$  and  $\phi_{\mathcal{L}}$  are related by (3.46). With the expressions (3.8) and (3.43), equation (3.54) may be rewritten as

$$(3.55) \quad W_{\mathbf{m}}(\phi_{\mathcal{L}}) + W_{\mathbf{m}}^*(i_{\mathcal{L}}) = \langle \phi_{\mathcal{L}}, i_{\mathcal{L}} \rangle.$$

It is not difficult to see from the explicit expressions for  $W_{\mathbf{m}}$  and  $W_{\mathbf{m}}^*$  that (3.55) holds when  $i_{\mathcal{L}}$  and  $\phi_{\mathcal{L}}$  are related by (3.46).  $\square$

### 3.5. Notes.

**3.5.1. Existence and Uniqueness of Solutions.** Equations (3.29)–(3.30) (resp., (3.44)–(3.45)) constitute mathematical models for the electric circuit. Their solution requires the specification of initial conditions  $q_{\mathcal{C}}(0)$  and  $i_{\mathcal{L}}(0)$  (resp.,  $q_{\mathcal{C}}(0)$  and  $\phi_{\mathcal{L}}(0)$ ). In a well-posed problem, these initial conditions must be compatible with the algebraic constraints embedded in (3.29)–(3.30) (resp., (3.44)–(3.45)). From a physical point of view, it seems natural to assume that there is a unique solution and that it is defined for all times  $t$ . The specification of conditions under which the mathematical models (3.29)–(3.30) and (3.44)–(3.45) meet these requirements is not the subject of the present study. Instead, the purpose of the present paper is to provide a variational interpretation for these models and to study their geometric properties.

**3.5.2. Integral Constraint on the Input.** The occurrence of an integral constraint (3.10) on the input  $u$  in the formulation of the optimal control problem may seem surprising. Without this constraint, the variational principle would give rise to the set of equations

$$(3.56) \quad (\dot{q}_{\mathcal{C}}, i_{\mathcal{L}}) \in \text{Image}(A_{\mathcal{C}}, A_{\mathcal{L}}),$$

$$(3.57) \quad \dot{\xi}_{\mathcal{C}} = \frac{\partial L}{\partial q_{\mathcal{C}}}(q_{\mathcal{C}}, i_{\mathcal{L}}),$$

$$(3.58) \quad A_{\mathcal{L}}^* \frac{\partial L}{\partial i_{\mathcal{L}}}(q_{\mathcal{C}}, i_{\mathcal{L}}) = A_{\mathcal{C}}^* \xi_{\mathcal{C}},$$

which corresponds to (3.19)–(3.21) with  $\zeta$  set equal to zero. In general, the dynamics of the electric circuit are only partially described by (3.56)–(3.58) because (3.58) imposes a constraint on the initial conditions:

$$(3.59) \quad A_{\mathcal{L}}^* \frac{\partial L}{\partial i_{\mathcal{L}}}(q_{\mathcal{C}}(0), i_{\mathcal{L}}(0)) \in \text{Image } A_{\mathcal{C}}^*.$$



These considerations suggest that, for the particular case of  $A_C^*$  being onto, the integral constraint (3.10) on the input is redundant. That this is actually the case follows, for example, from the variational interpretation. Indeed, notice that the integral constraint on  $u$  and the terminal constraints on  $q_C$  are related by the following equation, which is obtained by integrating (3.9):

$$(3.60) \quad q_C(t_1) - q_C(t_0) = A_C \int_{t_0}^{t_1} u \, dt.$$

If  $A_C^*$  is onto, then  $A_C$  has a trivial kernel, and thus the terminal constraints for  $q_C$  uniquely determine the value of  $\int_{t_0}^{t_1} u \, dt$ . As a consequence, it is also clear that for the special case of both  $A_C$  and  $A_L$  being the identity map, the considered optimal control problem essentially reduces to the classical variational problem of Lagrangian mechanics.<sup>7</sup>

**3.5.3. Nonphysical Capacitor Fluxes.** We may interpret  $\xi_C$  and  $\zeta$  in (3.19)–(3.21) in terms of nonphysical capacitor fluxes. In order to see this, introduce capacitor fluxes  $\phi_C$  according to

$$(3.61) \quad \dot{\phi}_C = v_C.$$

Then  $\phi_C$  and  $\phi_L$  are related by

$$(3.62) \quad A_C^* \phi_C + A_L^* \phi_L = A_C^* \phi_C(0) + A_L^* \phi_L(0),$$

which is obtained by integrating Kirchhoff's voltage law (3.37). With (3.8), (3.20), and (3.32), equation (3.61) may be rewritten as

$$(3.63) \quad \dot{\phi}_C = -\dot{\xi}_C.$$

It is thus clear that if we initially choose  $\xi_C(0) = -\phi_C(0)$ , then the equality  $\xi_C(t) = -\phi_C(t)$  holds for all times  $t$ . In this case  $\zeta$  corresponds to the constant of integration  $A_C^* \phi_C(0) + A_L^* \phi_L(0)$  that features in the integrated form (3.62) of Kirchhoff's voltage law. This follows from the equation

$$(3.64) \quad A_C^* \phi_C + A_L^* \phi_L = \zeta,$$

which we obtain from (3.21) and the definition of  $\phi_L$ .

**3.5.4. LC Circuits without Excess Elements.** In the remainder of this section we restrict attention to LC circuits without excess elements; that is, capacitor-only loops and inductor-only cutsets are not allowed. Under this restriction, there is a tree  $\mathcal{C} \subset \mathcal{B}$  with cotree  $\mathcal{L} \subset \mathcal{B}$  such that  $\mathcal{C}$  contains all of the capacitors and  $\mathcal{L}$  all of the inductors. This class of circuits has attracted considerable attention [4, 10, 11]. It has the important property that Kirchhoff's laws do not impose algebraic constraints on the capacitor voltages or inductor currents. For this class of LC circuits, Kirchhoff's current law may be formulated as

$$(3.65) \quad i_C = D i_L,$$

<sup>7</sup>Although, strictly speaking,  $A_C$  and  $A_L$  cannot both be equal to the identity map when we consider LC circuits (because  $i_C$  and  $i_L$  belong to different vector spaces), we may still formally consider the optimal control problem in this case, without any reference to electric circuits.

where  $D$  is a linear map from  $R_{\mathcal{L}}$  to  $R_{\mathcal{C}}$  defined by the topology of the circuit and the chosen current reference directions. This corresponds to the general formulation (3.5)–(3.6) with  $U = R_{\mathcal{L}}$ ,  $A_{\mathcal{C}} = D$ , and  $A_{\mathcal{L}}$  equal to the identity map. Applying the general theory of the previous section to this particular case, we obtain generalized Euler–Lagrange equations

$$(3.66) \quad \dot{q}_{\mathcal{C}} = D i_{\mathcal{L}},$$

$$(3.67) \quad \frac{d}{dt} \left( \frac{\partial L}{\partial i_{\mathcal{L}}} (q_{\mathcal{C}}, i_{\mathcal{L}}) \right) = D^* \frac{\partial L}{\partial q_{\mathcal{C}}} (q_{\mathcal{C}}, i_{\mathcal{L}})$$

and, if the inductors are also flux-controlled, Hamiltonian equations

$$(3.68) \quad \dot{q}_{\mathcal{C}} = D \frac{\partial H}{\partial \phi_{\mathcal{L}}} (q_{\mathcal{C}}, \phi_{\mathcal{L}}),$$

$$(3.69) \quad \dot{\phi}_{\mathcal{L}} = -D^* \frac{\partial H}{\partial q_{\mathcal{C}}} (q_{\mathcal{C}}, \phi_{\mathcal{L}}),$$

which actually form a *Poisson* Hamiltonian system.

We may also consider an alternative approach, which enables us to establish an interesting link with the literature. If the inductors are also flux-controlled, then the variable  $i_{\mathcal{L}}$  may be eliminated from the control Hamiltonian equations (3.19)–(3.21) (with  $A_{\mathcal{C}} = D$  and  $A_{\mathcal{L}}$  equal to the identity map), yielding the canonical symplectic Hamiltonian system

$$(3.70) \quad \dot{q}_{\mathcal{C}} = \frac{\partial \bar{H}}{\partial \xi_{\mathcal{C}}} (q_{\mathcal{C}}, \xi_{\mathcal{C}}),$$

$$(3.71) \quad \dot{\xi}_{\mathcal{C}} = -\frac{\partial \bar{H}}{\partial q_{\mathcal{C}}} (q_{\mathcal{C}}, \xi_{\mathcal{C}})$$

with Hamiltonian  $\bar{H} : R_{\mathcal{C}} \times R_{\mathcal{C}}^* \rightarrow \mathbb{R} : (q_{\mathcal{C}}, \xi_{\mathcal{C}}) \mapsto W_e(q_{\mathcal{C}}) + W_m(D^* \xi_{\mathcal{C}} + \zeta)$ .

With the interpretation of  $-\xi_{\mathcal{C}}$  as capacitor fluxes (section 3.5.3), the Hamiltonian system (3.70)–(3.71) has been studied in the literature; see, for example, [4]. Usually it is derived as follows. First, nonphysical capacitor fluxes are introduced by integrating Kirchhoff's voltage law. Second, the Lagrangian, equal to electric coenergy minus magnetic energy, is expressed in terms of these capacitor fluxes and their time derivatives, capacitor voltages. Next, a standard variational principle is invoked leading to a classical Euler–Lagrange equation. Finally, this is converted into the Hamiltonian system (3.70)–(3.71) by a Legendre transformation. Two important features of that approach are as follows: (i) nonphysical capacitor fluxes have to be introduced a priori; and (ii) the variational principle is expressed in terms of these nonphysical variables. This is in contrast with the present approach, where the variational principle is stated completely in terms of physically meaningful quantities and the introduction of capacitor fluxes is not an arbitrary step but instead dictated by the maximum principle.

**REMARK 3.** *The canonical symplectic Hamiltonian system (3.70)–(3.71) and the Poisson Hamiltonian system (3.68)–(3.69) are related: the map*

$$(3.72) \quad \Pi^{\zeta} : R_{\mathcal{C}} \times R_{\mathcal{C}}^* \rightarrow R_{\mathcal{C}} \times R_{\mathcal{L}}^* : (q_{\mathcal{C}}, \xi_{\mathcal{C}}) \mapsto (q_{\mathcal{C}}, \phi_{\mathcal{L}}) = (q_{\mathcal{C}}, D^* \xi_{\mathcal{C}} + \zeta)$$

*maps system (3.70)–(3.71) to (3.68)–(3.69).<sup>8</sup> This is easily seen in terms of the explicit expressions for the governing differential equations.*

<sup>8</sup>For one fixed value of  $\zeta$ , the map  $\Pi^{\zeta}$  does not necessarily map (3.70)–(3.71) onto (3.68)–(3.69) but only onto a subsystem of it. However, as the parameter  $\zeta$  varies over  $R_{\mathcal{L}}^*$ , these subsystems constitute the complete system (3.68)–(3.69). This follows from the fact that the map  $\Pi^{(\cdot)}(\cdot, \cdot) : R_{\mathcal{C}} \times R_{\mathcal{C}}^* \times R_{\mathcal{L}}^* \rightarrow R_{\mathcal{C}} \times R_{\mathcal{L}}^* : (q_{\mathcal{C}}, \xi_{\mathcal{C}}, \zeta) \mapsto \Pi^{\zeta}(q_{\mathcal{C}}, \xi_{\mathcal{C}})$  is onto.

**4. Algorithm for Variational Modeling.** The previous section introduced a novel variational formalism for LC circuits. The emphasis in that section was on the structural, geometric properties of the circuit enabling this variational approach. In the present section, we turn our attention to the practical application of the variational formalism as a means of deriving dynamic equations for the circuit. For brevity of the exposition, we restrict our attention to the derivation of the generalized Euler–Lagrange equations (thus omitting the transformation of these equations into a Dirac Hamiltonian system by means of a Legendre transformation).

An important difference between the algorithm of the present section and the theory of the previous section concerns the parameter vector  $\zeta$ . Recall from section 3.5.2 that the parameter vector  $\zeta$  is needed to ensure that the control Hamiltonian equations (3.16)–(3.18) or (3.19)–(3.21) constitute a mathematical model for the circuit. However, it is important to notice that the parameter vector  $\zeta$  does not appear in the generalized Euler–Lagrange equations (3.29)–(3.30), since it is eliminated when differentiating (3.21) with respect to time. Hence, in the present section, where the only purpose of the control Hamiltonian equations is to serve as an intermediate step in deriving the generalized Euler–Lagrange equations, the parameter vector  $\zeta$  may be omitted.

A second, minor difference between the present section and the previous section concerns the notational convention for the magnetic coenergy. In the previous section, when formulating the cost functional (3.11) and the control Hamiltonian (3.15), we have expressed the magnetic coenergy as a function of the independent branch currents by replacing  $i_{\mathcal{L}}$  with  $A_{\mathcal{L}}u$ . For the purpose of a systematic and algorithmic derivation of the dynamic equations, we choose to keep the natural expression of the magnetic coenergy as a function of the inductor currents and to augment the expression of the control Hamiltonian with an additional equation  $i_{\mathcal{L}} = A_{\mathcal{L}}u$ . As indicated below, this notational convention has to be taken into account when calculating the partial derivatives of the control Hamiltonian with respect to the independent branch currents (Step 4 of the algorithm).

With these modifications, the variational approach amounts to the following steps.

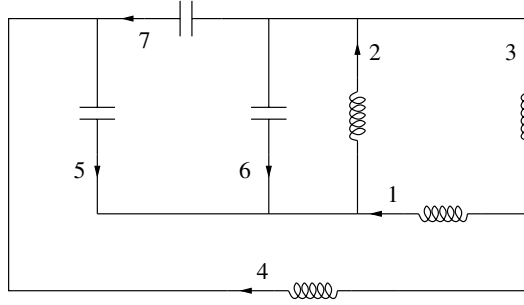
- Step 1. Select a set of independent branch currents that determine all the other currents in the circuit.
- Step 2. (a) Write the minus magnetic coenergy term, expressed as a function of the inductor currents.  
 (b) Write the electric energy term, expressed as a function of the capacitor charges.  
 (c) Associate with each capacitor (say, capacitor  $j$ ) a costate variable  $\xi_j$  and write the product of  $\xi_j$  with the corresponding capacitor current, expressed as a linear combination of the independent branch currents.

Denote the sum of these terms by  $H$ .

- Step 3. Augment this expression for  $H$  with an expression for the inductor currents as a linear combination of the independent branch currents.
- Step 4. For each capacitor (say, capacitor  $j$ ), write two differential equations:

$$(4.1) \quad \dot{q}_j = \frac{\partial H}{\partial \xi_j},$$

$$(4.2) \quad \dot{\xi}_j = -\frac{\partial H}{\partial q_j}.$$



**Fig. 3** *LC circuit with algebraic constraints on inductor currents and capacitor charges.*

For each of the independent branch currents (say,  $i_k$ ), write the algebraic equation

$$(4.3) \quad 0 = \frac{\partial H}{\partial i_k}.$$

Step 5. Differentiate (4.3) with respect to time and eliminate the  $\xi$  variables by means of (4.2).

The equations obtained in the last step, together with (4.1) and the expression of the inductor currents in terms of the independent branch currents, constitute a mathematical model for the circuit.

We briefly comment on some steps of this algorithm. A systematic, graph-theoretic procedure for the selection of independent branch currents (Step 1) is based on the notion of tree and cotree [17]; see section 2.4. If the circuit has no excess elements, then the inductor currents may be chosen as independent branch currents. Augmenting the expression for  $H$  with an expression for the inductor currents as a linear combination of the independent branch currents (Step 3) is only needed for those inductor currents that have not been selected as part of the independent branch currents. In particular, if the circuit has no excess elements and the inductor currents have been chosen as independent branch currents, then this step becomes redundant. For the calculation of  $\partial H / \partial i_k$  (Step 4), the functional dependence of the inductor currents on the independent branch currents has to be taken into account via the chain rule. This becomes trivial when the circuit has no excess elements and the inductor currents have been chosen as independent branch currents.

**EXAMPLE 6.** Consider the LC circuit of Figure 3. Each branch has been assigned a reference number and a current reference direction. For simplicity of the exposition, we assume that the constitutive relations for the inductors and capacitors are linear and bijective; they are, respectively, characterized by their nonzero inductance ( $L_1$ ,  $L_2$ ,  $L_3$ , and  $L_4$ ) and their nonzero capacitance ( $C_5$ ,  $C_6$ , and  $C_7$ ).

Step 1. Branches 3, 6, and 7 constitute a tree. Hence the cotree branch currents  $i_1$ ,  $i_2$ ,  $i_4$ , and  $i_5$  are independent and determine all other branch currents.

Steps 2–3.

$$(4.4) \quad \begin{aligned} H = & -\frac{L_1 i_1^2}{2} - \frac{L_2 i_2^2}{2} - \frac{L_3 i_3^2}{2} - \frac{L_4 i_4^2}{2} + \frac{q_5^2}{2C_5} + \frac{q_6^2}{2C_6} + \frac{q_7^2}{2C_7} \\ & + \xi_5 i_5 + \xi_6 (i_2 - i_1 - i_5) + \xi_7 (i_5 - i_4) \end{aligned}$$

with

$$(4.5) \quad i_3 = -i_1 - i_4.$$

Step 4.

$$(4.6) \quad \begin{aligned} \dot{q}_5 &= i_5, \\ \dot{q}_6 &= i_2 - i_1 - i_5, \\ \dot{q}_7 &= i_5 - i_4, \end{aligned}$$

$$(4.7) \quad \begin{aligned} \dot{\xi}_5 &= -\frac{q_5}{C_5}, \\ \dot{\xi}_6 &= -\frac{q_6}{C_6}, \\ \dot{\xi}_7 &= -\frac{q_7}{C_7}, \end{aligned}$$

$$(4.8) \quad \begin{aligned} L_1 i_1 - L_3 i_3 &= -\xi_6, \\ L_2 i_2 &= \xi_6, \\ -L_3 i_3 + L_4 i_4 &= -\xi_7, \\ \xi_5 - \xi_6 + \xi_7 &= 0. \end{aligned}$$

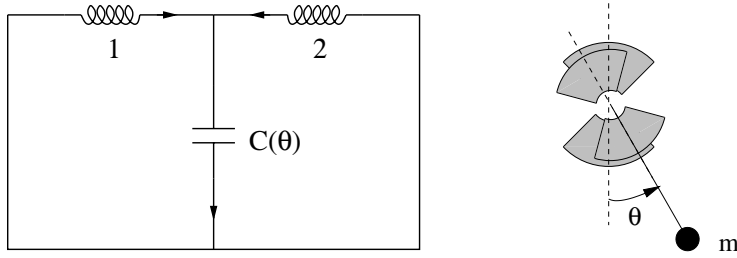
Step 5. Differentiating (4.8) with respect to time and eliminating the  $\xi$  variables by means of (4.7) yields

$$(4.9) \quad \begin{aligned} L_1 \frac{di_1}{dt} - L_3 \frac{di_3}{dt} &= \frac{q_6}{C_6}, \\ L_2 \frac{di_2}{dt} &= -\frac{q_6}{C_6}, \\ -L_3 \frac{di_3}{dt} + L_4 \frac{di_4}{dt} &= \frac{q_7}{C_7}, \\ -\frac{q_5}{C_5} + \frac{q_6}{C_6} - \frac{q_7}{C_7} &= 0. \end{aligned}$$

Equations (4.5)–(4.6) and (4.9) constitute a mathematical model for the circuit.

**5. Application to Electromechanical Systems.** In the previous sections, we have introduced a variational approach for the modeling of electric inductor-capacitor circuits. This approach generalizes the classical variational principle from theoretical mechanics. Some distinctive features of the present approach are that (i) it is applicable to circuits with arbitrary interconnection structure, and (ii) the resulting equations are expressed in terms of physically meaningful variables. This is in contrast to, for example, the variational approaches of [7, 4] where nonphysical inductor charges or capacitor fluxes have to be introduced and where the associated Euler–Lagrange equations are expressed in terms of these nonphysical variables.

We are aware that, in general, the variational formalism is seldom used for the routine analysis of linear electric circuits. However, variational methods are of great utility and are widely employed in the study of electromechanical systems [7]. As the theory of the present paper encompasses the classical variational principle of theoretical mechanics (sections 2.5 and 3.5.2), it is natural to expect that the current approach can be applied to the study of electromechanical systems. A rigorous proof of this statement would involve a systematic analysis of the dynamics of general



**Fig. 4** Electromechanical system: electrical part (left) and butterfly capacitor with pendulum (right).

electromechanical systems, which lies outside the scope of the present paper. Instead, we illustrate the application of the present approach to electromechanical systems by means of a particular example. For brevity of the exposition, we restrict our attention to the derivation of the generalized Euler–Lagrange equations (thus omitting the transformation of these equations into a Dirac Hamiltonian system by means of a Legendre transformation).

**EXAMPLE 7.** Consider the electromechanical system of Figure 4. The circuit has no excess elements: the inductor currents are independent and determine the capacitor current. For simplicity of the exposition, we assume that the inductors are linear; they are characterized by their inductance ( $L_1$  and  $L_2$ ). The capacitor is a butterfly capacitor that is assumed to be electrically linear and whose capacitance  $C(\theta)$  depends on the orientation  $\theta$  of the rotor. The angular velocity  $d\theta/dt$  of the rotor is denoted by  $\omega$ . For simplicity of the exposition, we neglect the axial moment of inertia of the rotor. Attached to the rotor of the butterfly capacitor, there is an idealized pendulum with length  $l$  and mass  $m$ .

A variational description of this electromechanical system starts with the formulation of the optimal control problem seeking the minimization of the time integral of magnetic and kinetic coenergy minus electric and potential energy

$$(5.1) \quad \int_{t_0}^{t_1} \left\{ \frac{1}{2} L_1 i_1^2 + \frac{1}{2} L_2 i_2^2 + \frac{1}{2} m l^2 \omega^2 - \frac{q^2}{2C(\theta)} + mgl \cos(\theta) \right\} dt$$

subject to the dynamics

$$(5.2) \quad \dot{q} = i_1 + i_2, \quad \dot{\theta} = \omega$$

on a fixed time interval with fixed terminal points and integral constraints on the inputs.

The control Hamiltonian with abnormal multiplier equal to  $-1$  is given by<sup>9</sup>

$$(5.3) \quad \begin{aligned} \mathcal{H}^{\zeta_1, \zeta_2, \zeta_\omega}(q, \theta, \xi_q, \xi_\theta, i_1, i_2, \omega) = & -\frac{1}{2} L_1 i_1^2 - \frac{1}{2} L_2 i_2^2 - \frac{1}{2} m l^2 \omega^2 + \frac{1}{2C(\theta)} q^2 \\ & - mgl \cos(\theta) + \xi_q(i_1 + i_2) + \xi_\theta \omega + \zeta_1 i_1 + \zeta_2 i_2 + \zeta_\omega \omega. \end{aligned}$$

<sup>9</sup>Following the lines of the previous section, we may omit the last three terms in the control Hamiltonian (containing the parameters  $\zeta_1$ ,  $\zeta_2$ , and  $\zeta_\omega$ ) if the only purpose of the control Hamiltonian equations is to serve as an intermediate step in the derivation of the generalized Euler–Lagrange equations.

The corresponding control Hamiltonian equations are

$$(5.4) \quad \begin{aligned} \dot{q} &= i_1 + i_2, \\ \dot{\theta} &= \omega, \end{aligned}$$

$$(5.5) \quad \begin{aligned} \dot{\xi}_q &= -\frac{q}{C(\theta)}, \\ \dot{\xi}_\theta &= \frac{q^2}{2C(\theta)^2} \frac{dC}{d\theta}(\theta) - mgl \sin(\theta), \end{aligned}$$

$$(5.6) \quad \begin{aligned} L_1 i_1 &= \xi_q + \zeta_1, \\ L_2 i_2 &= \xi_q + \zeta_2, \\ ml^2 \omega &= \xi_\theta + \zeta_\omega. \end{aligned}$$

Differentiating (5.6) with respect to time and eliminating the  $\xi$  variables by means of (5.5) yields

$$(5.7) \quad \begin{aligned} L_1 \frac{di_1}{dt} &= -\frac{q}{C(\theta)}, \\ L_2 \frac{di_2}{dt} &= -\frac{q}{C(\theta)}, \\ ml^2 \dot{\omega} &= \frac{q^2}{2C(\theta)^2} \frac{dC}{d\theta}(\theta) - mgl \sin(\theta). \end{aligned}$$

Equations (5.4) and (5.7) constitute a mathematical model for the electromechanical system.

**6. Conclusion.** Classically, a Lagrangian model is obtained from a variational principle by means of the calculus of variations. The approach that we propose still starts from variational principles, but we use the maximum principle of optimal control theory to derive the dynamic equations. With this approach we are able to consider variational principles that are characterized not only by a cost function (the Lagrangian) but also by additional underlying dynamics that may be nontrivial. This additional degree of freedom enables us to describe a broader class of systems within a generalized Lagrangian framework. In the present paper we have applied this approach to the study of electric inductor-capacitor circuits. This study has yielded a new, variational approach for deriving the dynamic equations of an inductor-capacitor circuit. Compared with the classical variational approaches in [7, 4, 19], the present approach has the advantage that it does not impose constraints on the circuit topology and that it yields a system of equations expressed directly in terms of physically meaningful variables. The present framework may provide a useful modeling tool, especially in those areas where electric circuits interact dynamically with other types of systems and where variational methods are of great interest. This has been illustrated in the article by means of an electromechanical example.

**Appendix. The Maximum Principle.** We present a version of the maximum principle that applies to a particular class of smooth control problems on vector spaces. Let  $X$ ,  $U$ , and  $\Lambda$  be real, finite-dimensional vector spaces and consider smooth

functions

$$(A.1) \quad f : X \times U \rightarrow X : (x, u) \rightarrow f(x, u),$$

$$(A.2) \quad \lambda : X \times U \rightarrow \Lambda : (x, u) \rightarrow \lambda(x, u),$$

$$(A.3) \quad L : X \times U \rightarrow \mathbb{R} : (x, u) \rightarrow L(x, u).$$

We are interested in the following optimal control problem. *Consider arbitrary but fixed  $t_0 < t_1 \in \mathbb{R}$ ,  $x_0, x_1 \in X$ , and  $\lambda_1 \in \Lambda$ . Among all piecewise continuous controls  $u : [t_0, t_1] \rightarrow U$  that steer  $x(t_0) = x_0$  to  $x(t_1) = x_1$  according to the dynamics*

$$(A.4) \quad \dot{x}(t) = f(x(t), u(t))$$

*while satisfying the constraint*

$$(A.5) \quad \int_{t_0}^{t_1} \lambda(x(t), u(t)) \, dt = \lambda_1,$$

*find the particular control for which the cost functional*

$$(A.6) \quad \int_{t_0}^{t_1} L(x(t), u(t)) \, dt$$

*is minimized.*

DEFINITION 4. *A Pontryagin extremal on the interval  $[t_0, t_1]$  is a 5-tuple  $(\psi, \zeta, x, \xi, u)$  with  $\psi \in \{-1, 0\}$ ,  $\zeta \in \Lambda^*$ ,  $x : [t_0, t_1] \rightarrow X : t \mapsto x(t)$  absolutely continuous,  $\xi : [t_0, t_1] \rightarrow X^* : t \mapsto \xi(t)$  absolutely continuous, and  $u : [t_0, t_1] \rightarrow U : t \mapsto u(t)$  piecewise continuous, with not all  $\psi, \zeta, \xi$  vanishing, and satisfying*

$$(A.7) \quad \dot{x}(t) = \frac{\partial \mathcal{H}^{\psi, \zeta}}{\partial \xi}(x(t), \xi(t), u(t)),$$

$$(A.8) \quad \dot{\xi}(t) = -\frac{\partial \mathcal{H}^{\psi, \zeta}}{\partial x}(x(t), \xi(t), u(t)),$$

$$(A.9) \quad 0 = \frac{\partial \mathcal{H}^{\psi, \zeta}}{\partial u}(x(t), \xi(t), u(t))$$

*almost everywhere on the time interval  $[t_0, t_1]$  with*

$$(A.10) \quad \mathcal{H}^{\psi, \zeta} : X \times X^* \times U : (x, \xi, u) \mapsto \psi L(x, u) + \langle \xi, f(x, u) \rangle + \langle \zeta, \lambda(x, u) \rangle.$$

*The function  $\mathcal{H}^{\psi, \zeta}$  is called the control Hamiltonian, and (A.7)–(A.9) the associated control Hamiltonian equations. If  $\psi = -1$ , then the Pontryagin extremal is said to be normal and the control Hamiltonian is denoted by  $\mathcal{H}^\zeta$ .*

The following result is a straightforward application of the theory of optimal processes developed by Pontryagin et al. in [14].

THEOREM 5. *A piecewise continuous control  $u : [t_0, t_1] \rightarrow U$  solves the given optimal control problem only if there exists a corresponding Pontryagin extremal  $(\psi, \zeta, x, \xi, u)$  on the interval  $[t_0, t_1]$  satisfying the terminal point constraints  $x(t_0) = x_0$  and  $x(t_1) = x_1$  and the integral constraint (A.5).*

*Proof.* We reformulate the constraint (A.5) by augmenting the dimension of the state space: If the equation

$$(A.11) \quad \dot{y}(t) = \lambda(x(t), u(t))$$



is added to the dynamics, then (A.5) is equivalent to the terminal point constraints  $y(t_0) = 0$  and  $y(t_1) = \lambda_1$  and a standard optimal control problem is obtained on a fixed time interval with fixed terminal points. The conclusion of the theorem now follows from the discussion in [14, pp. 66–68].  $\square$

The following remarks are important from the perspective of variational modeling of dynamical systems: (i) It is a fundamental and perhaps remarkable fact that the Pontryagin extremals satisfy a set of equations (A.7)–(A.9) that is independent of the data  $t_0$ ,  $t_1$ ,  $x_0$ ,  $x_1$ , and  $\lambda_1$ . (ii) The theory of optimal control actually leads to a stronger result than that presented here. According to [14, pp. 66–68], the conclusion of the theorem is still valid if (A.9) is replaced by the stronger condition

$$(A.12) \quad \mathcal{H}^{\psi, \zeta}(x(t), \xi(t), u(t)) = \max_{v \in U} \mathcal{H}^{\psi, \zeta}(x(t), \xi(t), v).$$

With an abuse of terminology, we refer to Theorem 5 as a version of the Pontryagin maximum principle, although we have replaced the maximum condition (A.12) by its first-order necessary condition (A.9).

#### REFERENCES

- [1] V. I. ARNOLD, *Mathematical Methods of Classical Mechanics*, 2nd ed., Graduate Texts in Math., Springer-Verlag, New York, 1989.
- [2] A. M. BLOCH AND P. E. CROUCH, *Representations of Dirac structures on vector spaces and nonlinear L-C circuits*, in *Differential Geometry and Control*, Proc. Sympos. Pure Math. 64, AMS, Providence, RI, 1999, pp. 103–117.
- [3] W. M. BOOTHBY, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, 2nd ed., Pure Appl. Math. 120, Academic Press, New York, 1986.
- [4] L. O. CHUA AND J. D. MCPHERSON, *Explicit topological formulation of Lagrangian and Hamiltonian equations for nonlinear networks*, IEEE Trans. Circuits Systems, 21 (1974), pp. 277–286.
- [5] J. CLEMENTE-GALLARDO AND J. M. A. SCHERPEN, *Relating Lagrangian and Hamiltonian frameworks for LC circuits*, IEEE Trans. Circuits Systems I Fund. Theory Appl., 50 (2003), pp. 1359–1363.
- [6] T. J. COURANT, *Dirac manifolds*, Trans. Amer. Math. Soc., 319 (1990), pp. 631–661.
- [7] S. H. CRANDALL, ED., *Dynamics of Mechanical and Electromechanical Systems*, McGraw-Hill, New York, 1968.
- [8] I. DORFMAN, *Dirac Structures and Integrability of Nonlinear Evolution Equations*, Wiley, Chichester, UK, 1993.
- [9] V. JURDJEVIC, *Geometric Control Theory*, Cambridge University Press, Cambridge, UK, 1997.
- [10] H. G. KWATNY, F. M. MASSIMO, AND L. Y. BAHAR, *The generalized Lagrange formulation for nonlinear RLC networks*, IEEE Trans. Circuits Systems, 29 (4) (1982), pp. 220–233.
- [11] B. M. MASCHKE, A. J. VAN DER SCHAFT, AND P. C. BREEDVELD, *An intrinsic Hamiltonian formulation of the dynamics of LC-circuits*, IEEE Trans. Circuits Systems I Fund. Theory Appl., 42 (2) (1995), pp. 73–82.
- [12] L. MOREAU AND D. AEYELS, *A variational principle for nonlinear LC circuits with arbitrary interconnection structure*, in *Preprints of the 5th IFAC Symposium on Nonlinear Control Systems (NOLCOS'01)*, Saint Petersburg, Russia, 2001, pp. 54–59.
- [13] B. NOBLE AND M. J. SEWELL, *On dual extremum principles in applied mathematics*, J. Inst. Math. Appl., 9 (1972), pp. 123–193.
- [14] L. S. PONTRYAGIN, V. G. BOLTYANSKII, R. V. GAMKRELIDZE, AND E. F. MISHCHENKO, *The Mathematical Theory of Optimal Processes*, Interscience, New York, 1965.
- [15] K. J. REINSCHKE, *On network models and the symbolic solution of network equations*, Int. J. Appl. Math. Comput. Sci., 11 (2001), pp. 237–270.
- [16] J. M. A. SCHERPEN, D. JELTSEMA, AND J. B. KLAASSENS, *Lagrangian modeling of switching electrical networks*, Systems Control Lett., 48 (5) (2003), pp. 365–374.
- [17] S. SESHU AND M. B. REED, *Linear Graphs and Electrical Networks*, Addison-Wesley, Reading, MA, 1961.
- [18] H. J. SUSSMANN AND J. C. WILLEMS, *300 years of optimal control: From the brachystochrone to the maximum principle*, IEEE Control Systems, 17 (3) (1997), pp. 32–44.

- [19] A. SZATKOWSKI, *Remark on “Explicit topological formulation of Lagrangian and Hamiltonian equations for nonlinear networks,”* IEEE Trans. Circuits Systems, 26 (5) (1979), pp. 358–360.
- [20] A. VAN DER SCHAFT AND B. MASCHKE, *The Hamiltonian formulation of energy conserving physical systems with external ports,* AEU-Int. J. Electron. Commun., 49 (5/6) (1995), pp. 362–371.