IMPLICIT PORT-HAMILTONIAN SYSTEMS, CONTINUOUS AND SAMPLED-DATA MODELS

FERNANDO CASTAÑOS[†], HANNAH MICHALSKA[‡], DMITRY GROMOV[‡], AND VINCENT HAYWARD[§]

Abstract. We consider smooth finite-dimensional port-Hamiltonian (PH) systems described explicitly (as an ordinary differential equation) and implicitly (as a differential-algebraic equation). We make a precise connection between the implicit and explicit representations by embedding the configuration manifold in Euclidean space. We use the implicit representation to construct a simple sampled-data model that preserves the PH structure under sample and hold.

Key words. port-Hamiltonian systems, nonlinear implicit systems, symplectic integration, sampled-data systems, modeling of physical systems

AMS subject classifications. 65P10, 93C10, 93C57, 34A09, 93A30

1. Introduction.

Hamiltonian systems constitute an important class of conservative systems that appear in many domains like: mechanics, circuit theory, astronomy, molecular dynamics and ecology. The class of Hamiltonian systems was extended in [10] to include open systems, systems that interact with the environment via a set of inputs and outputs (called *ports*), giving rise to port-Hamiltonian (PH) systems. These extended models immediately reveal the passive properties of the corresponding system, making them particularly well suited for designing passivity-based control (PBC) laws. In this paper we derive the connection between the implicit and explicit representations of PH systems and propose a splitting method for constructing discrete models for sampled-data PH systems.

1.1. Implicit and explicit representations.

When dealing with complex dynamic systems, it is common in science and engineering practice to decompose a model into several, simpler submodels. These submodels, which typically consist of ordinary differential equations (ODEs), are then interconnected to construct a model for the original aggregate system. Interconnecting the submodels amounts to imposing a set of algebraic constraints, so the overall resulting model is typically a differential-algebraic equation (DAE).

Under smoothness assumptions on the solutions, it is possible to write a given DAE (an implicit representation of the system) as an equivalent ODE on an abstract manifold (an explicit representation) or, in other words, as a vector field [43, 44, 40, 41]. This approach is usually preferred for general analysis (see [2, 13, 27] in the context of analytical mechanics, or [49, 29] in the context of circuit theory), while working directly with DAEs is typically preferred for numerical computations.

For Hamiltonian systems, both the DAE and the ODE models have been extensively studied in the literature, but little has been said about the formal relation between the DAE and the ODE approaches, about how the coordinate systems of each framework relate to each other. In § 2 we show that such a formal relation can

[†]Automatic Control Department, CINVESTAV del IPN, Av. IPN 2508, Col San Pedro Zacatenco, C.P. 07360, México D.F., México (castanos@ieee.org)

[‡]McGill Centre for Intelligent Machines, 3480 University Street, Montreal Quebec, Canada H3A 2A7 (hannah.michalska@mcgill.ca, gromov@cim.mcgill.ca)

[§]**Please provide**

be established if one regards the system's configuration space as a manifold embedded in n-dimensional Euclidean space.

1.2. Passivity of sampled-data systems. We consider the problem of constructing sampled-data (discrete-time) models out of PH continuous-time models. Sampled-data models arise, e.g., in digital control problems, where the system output is being sampled at regular intervals, a control is computed and then held constant during the sampling period. In principle, the discrete models here derived could be used to design PBC laws directly in discrete time, instead of taking the usual approach of emulating a previously-computed continuous-time control law.

Suppose, for concreteness, that an explicit, affine in the controls, continuous-time system is given in local coordinates by the ODE

$$\dot{x}(t) = f_0(x(t)) + f_{\mathrm{u}}(x(t))u(t) \tag{1.1a}$$

$$y(t) = l(x(t)) \,, \tag{1.1b}$$

where $u(t) \in \mathbb{R}^m$ is the system input at time t and $x(t) \in \mathbb{R}^n$ and $y(t) \in \mathbb{R}^m$ are the corresponding state and output, respectively (later on, we will refine the problem statement). Recall that the system is said to be passive if there exists a continuous function $H : \mathbb{R}^n \to \mathbb{R}$, which is bounded from below and satisfies the dissipation inequality

$$H(x(t)) - H(x(0)) \leq \int_0^t y(\tau)^\top u(\tau) \mathrm{d}\tau \tag{1.2}$$

along the trajectories of the system. Suppose now that the input is held constant during regular sampling intervals of period h, that is,

$$u(t) \equiv u_{\mathrm{d}}(\alpha) \,, \quad t \in [h\alpha, h\alpha + h) \,, \quad \alpha \in \mathbb{N} := \{0, 1, 2, \dots\} \,,$$

where $u_{\mathrm{d}}(\alpha)$ is a new, discrete-time input. A difference equation of the form

$$x_{\mathrm{d}}(\alpha + 1) = f_{\mathrm{d}}(x_{\mathrm{d}}(\alpha), u_{\mathrm{d}}(\alpha)) \,, \quad x_{\mathrm{d}}(\alpha) \in \mathbb{R}^n \,, \tag{1.3}$$

is a sampled-data model of (1.1) if, whenever $x_{\mathrm{d}}(0) = x(0)$, $x_{\mathrm{d}}(\alpha) = x(h\alpha)$ for all $\alpha \in \mathbb{N}$ and every locally bounded sequence $u_{\mathrm{d}}(\alpha)$. Notice that we have written (1.3) in a more general form than (1.1a). This is becsude affinity of the state derivative with respect to the controls u is typically lost under sampling [33].

Except for linear systems, it is typically impossible to construct an *exact* sampled-data model (for it requires the exact solution of the corresponding ODE or DAE), so one has to settle for an approximate solution. For general dynamic systems, the choice of an integration method for generating an approximate solution (like Euler or Runge-Kutta) usually depends on the compromise between complexity and order of approximation. But for structured dynamic systems like Hamiltonian systems, structure preservation is usually the main criterion for choosing a numerical method (see, e.g., [16] for more details on structure-preserving numerical schemes).

Structure preservation is not only important for devising numerical schemes, it is also important for control techniques like PBC. In fact, it could be argued that structure preservation is *the* design principle for PBC: the design cycle begins with a system having some structure (passivity) and the objective is to find a control law such that the closed-loop systems retains the same structure, i.e., such that the closed-loop system is also passive [35], but with a different, desired storage function.

If, moreover, the open-loop system is PH (a more refined structure than passivity), then the objective is to construct a closed-loop system which is also PH [52, 37, 38, 36].

Several definitions of discrete-time passivity appear in the literature. A definition of discrete-time positive realness along with the discrete-time analog of Kalman-Yakubovich-Popov (KYP) Lemma[1] can be found in [18]. A reasonable definition for discrete-time nonlinear systems is as follows. A discrete-time system is said to be passive if there exists a function $H_{\rm d} : \mathbb{R}^n \to \mathbb{R}$ which is bounded from below and

$$H_{\rm d}(x_{\rm d}(\alpha+1)) - H_{\rm d}(x_{\rm d}(\alpha)) \leq y_{\rm d}(\alpha)^\top u_{\rm d}(\alpha)\ , \tag{1.4}$$

where $y_{\rm d}(\alpha) \in \mathbb{R}^m$ is the system output [26]. A discrete version of Hill-Moylan Theorem[2] can be found in [7] (necessary conditions only), where it is also shown that, if a system is discrete-time passive, then the output must be of zero relative degree, i.e., it must have a direct dependence on the input[3].

If a continuous system is passive, it is natural to expect the sampled-data system to be passive as well and, moreover, with the same storage function ($H_{\rm d} = H$). It is clear from (1.2) and (1.4), that this requires

$$y_{\rm d}(\alpha) = \int_{\alpha h}^{\alpha+h} y(\tau){\rm d}\tau \tag{1.5}$$

Thus, even if the exact discrete model (1.3) is available, it is necessary to integrate the continuous output y. Again, for linear systems the passive output can be integrated explicitly [9]. For mechanical systems with passive output equal to the velocity, the integral (1.5) can be written as a position difference [50]. For more general systems, the passivity of sampled-data systems can be formulated in terms of a pair of coupled difference and differential equations [31]. The case of PH systems was worked out in [32]. Another problem that appears while constructing a discrete-time model of a passive system is the loss of affinity in the controls, which stems the use of a simple algebraic condition (like the one given in Hill-Moylan Theorem) to verify passivity.

In this paper we show that, using an implicit representation, it is possible to construct a simple approximate discrete-time model which not only preserves passivity, but also the PH structure of the continuous system. The passive structure is not preserved in the discrete-time sense (1.4), it is rather preserved in the sense that the discrete model is the *exact* representation of another continuous-time PH system which, up to an approximation error of order two with respect to the sampling interval h, has the same storage function H and same output function y. The method is an extension of a method devised for pure Hamiltonian systems (i.e., without ports).

Pure Hamiltonian systems conserve two quantities: the Hamiltonian function H (i.e., the energy or storage function) and a certain two-form ω, called the symplectic form. While building a discrete model, one can build a method that either preserves H or ω, but not both (see § 2 for details). The discrete model presented here is based on a method that preserves the symplectic form [42]. For a method that preserves H see, e.g., [21, 22]. The reason why preservation of ω is usually preferred over preservation of H is that preservation of ω precisely characterizes Hamiltonian systems (see Theorem 2.7 and Remark 4).

[1] Recall that, for linear systems, positive realness is equivalent to passivity [55].

[2] Hill-Moylan Theorem gives an algebraic characterization of nonlinear affine passive systems. It can be interpreted as the nonlinear version of KYP Lemma [17].

[3] In continuous time, the relative degree of a passive system can be zero or one [6].

2. Implicit Hamiltonian systems. In this section we use the notion of embedding in order to establish a formal relation between the implicit and explicit representations of Hamiltonian systems. We also recall, in both frameworks, the main properties of Hamiltonian systems: energy conservation and symplecticity.

2.1. Phase space, intrinsic and extrinsic representations. We begin by introducing two alternative representations for the configuration space and its tangent and cotangent bundles. Let $g : \mathbb{R}^n \to \mathbb{R}^k$, with $k \leq n$, be a smooth function and let

$$\mathbb{M} := \{a \in \mathbb{R}^n \mid g(a) = 0\} \tag{2.1}$$

be the *configuration space* of a constrained system.

ASSUMPTION 1. *The vector* $\mathbf{0} \in \mathbb{R}^k$ *is a* regular value *of* g*, that is to say, at every point in* $\mathbb{M}$*, the Jacobian matrix* $G := \left\{\frac{\partial g_i}{\partial r_j}\right\}_{ij}$ *has full rank, i.e.,* $\operatorname{rank}(G) = k$.

This assumption ensures that $\mathbb{M}$ is a manifold, an o-dimensional differentiable manifold defined by $k = n - o$ functionally independent equations. Notice that without this assumption, $\mathbb{M}$ is not necessarily a manifold, even if g is smooth. We say that the system is constrained by the relations $g^1 = 0, \ldots, g^k = 0$ and that the constraints are *holonomic*.

We have defined $\mathbb{M}$ in (2.1) as an embedded submanifold of $\mathbb{R}^n$, but $\mathbb{M}$ can of course be defined intrinsically (without invoking an ambient space) if one resorts to local coordinate charts. In what follows, we will write $\mathcal{M}$ when using the abstract or intrinsic representation of the manifold and we will write $\mathbb{M}$ when regarding the manifold as a subset of $\mathbb{R}^n$. The Constant-Rank Level Set Theorem [23] together with Assumption 1 guarantees that $\mathbb{M}$ is an embedded submanifold, which by definition means that there exists an injective smooth map (i.e., an embedding) $\imath : \mathcal{M} \hookrightarrow \mathbb{R}^n$ taking a point in the abstract set $\mathcal{M}$ and mapping it to the same point, but treated as an element of $\mathbb{R}^n$:

$$\imath(\mathcal{M}) = \mathbb{M} \subset \mathbb{R}^n \, . \tag{2.2}$$

In other words, $\imath$ satisfies $g \circ \imath \equiv 0$ (as can be readily seen from (2.2) and (2.1)) and can be used to relate both representations.

Let $T_a\mathbb{R}^n$ and $T_a^*\mathbb{R}^n$ be, respectively, the tangent and cotangent spaces of $\mathbb{R}^n$ at $a \in \mathbb{M}$. Also, let $T_\alpha\mathcal{M}$ and $T_\alpha^*\mathcal{M}$ be the tangent and cotangent spaces of $\mathcal{M}$ at α. Besides $\mathbb{M}$ itself, the constraint $g = 0$ defines a tangent subbundle of the tangent bundle $T\mathbb{R}^n$, a subbundle that we wish to identify with $T\mathcal{M}$. In order to write this subbundle down, let us define first a particular subspace of $T_a\mathbb{R}^n$. In terms of derivations, it is given by

$$T_a\mathbb{M} := \left\{X \in T_a\mathbb{R}^n \mid X(g^i) = 0\right\} \, , \tag{2.3}$$

where $i = 1, \ldots, k$ (this is simply the tangent space of $\mathbb{M}$ at $a \in \mathbb{M}$). The tangent subbundle is then

$$T\mathbb{M} := \coprod_{a \in \mathbb{M}} T_a\mathbb{M} \subset T\mathbb{R}^n \, .$$

The following proposition allows us to identify $T\mathcal{M}$ with $T\mathbb{M}$.

PROPOSITION 2.1. $T_\alpha\mathcal{M}$ *is isomorphic to* $T_{\imath(\alpha)}\mathbb{M}$.

Proof. The linear transformation $\imath_* : T_\alpha\mathcal{M} \to T_{\imath(\alpha)}\mathbb{R}^n$ between two linear spaces is an isomorphism if $\imath_*$ is a one-to-one mapping of $T_\alpha\mathcal{M}$ onto $T_{\imath(\alpha)}\mathbb{R}^n$. Note that $\imath_*$ is an injection (and so an isomorphism onto its image) thus, it suffices to show that

$$\imath_*(T_\alpha\mathcal{M}) = T_{\imath(\alpha)}\mathbb{M} \,. \tag{2.4}$$

Suppose that $X \in \imath_*(T_\alpha\mathcal{M})$, i.e., $X = \imath_* Y$ for some $Y \in T_\alpha\mathcal{M}$. Then,

$$X(g^i) = \imath_* Y(g^i) = Y(g^i \circ \imath) = 0 \,,$$

which implies that $X \in T_{\imath(\alpha)}\mathbb{M}$. This proves that $\imath_*(T_\alpha\mathcal{M}) \subset T_{\imath(\alpha)}\mathbb{M}$. The equality (2.4) follows from linearity and the fact that the dimensions of $\imath_*(T_\alpha\mathcal{M})$ and $T_{\imath(\alpha)}\mathbb{M}$ are both equal to o ($\dim T_{\imath(\alpha)}\mathbb{M} = o$ because of Assumption 1). □

At a purely geometric level, g does not define a canonical subbundle of $T^*\mathbb{R}^n$, as it does with $T\mathbb{R}^n$ (recall that there is no canonical isomorphism between a linear space and its dual), but it its possible to construct a particular submanifold $\mathcal{L}\mathbb{M}$ of $T^*\mathbb{R}^n$ by incorporating physical knowledge of the system.

Let $L : T\mathbb{R}^n \to \mathbb{R}$ be the Lagrangian function of the system under study. For mechanical systems, L is the difference between the kinetic and the potential energy.

Definition 2.2 ([27]). *The fiber derivative of L is a map $\mathcal{F}L : T\mathbb{R}^n \to T^*\mathbb{R}^n$ that satisfies the following equality:*

$$\langle \mathcal{F}L(X), Y\rangle = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} L(a, X + sY) \tag{2.5}$$

for all $X, Y \in T_a\mathbb{R}^n$, where $\langle\,\cdot\,,\,\cdot\,\rangle$ is the standard pairing of a vector and a covector.

Intuitively, $\langle \mathcal{F}L(X), Y\rangle$ gives the variation in L at X, when one varies along the fiber $T_a\mathbb{R}^n$ in the direction Y. Notice that $\mathcal{F}L$ is fiber-preserving in the sense that it maps the fiber $T_a\mathbb{R}^n$ into the fiber $T^*_a\mathbb{R}^n$, but in general it is not a bundle mapping [23], since linearity might fail. In the context of classical mechanics, the covector $\mathcal{F}L(X)$ is simply the *momentum*.

Definition 2.3. *The Legendre bundle $\mathcal{L}\mathbb{M}$ is a subbundle of $T^*\mathbb{R}^n$ given by*

$$\mathcal{L}_a\mathbb{M} := \mathcal{F}L(T_a\mathbb{M}) \quad \textit{and} \quad \mathcal{L}\mathbb{M} = \coprod_{a\in\mathbb{M}} \mathcal{L}_a\mathbb{M} \subset T^*\mathbb{R}^n \,. \tag{2.6}$$

Hereafter, we will use $\{r^i, \dot r^i\}$ and $\{r^i, p_i\}$ to denote (global) coordinates on $T\mathbb{R}^n$ and $T^*\mathbb{R}^n$, respectively, and we will use $\{q^i, \dot q^i\}$ and $\{q^i, \hat p_i\}$ to denote local coordinates on $T\mathcal{M}$ and $T^*\mathcal{M}$. In coordinates, the fiber derivative takes the form

$$\mathcal{F}L : (r, \dot r) \mapsto (r, p) : p_i = \frac{\partial L(r, \dot r)}{\partial \dot r^i} \,. \tag{2.7}$$

The following assumption ensures that $\mathcal{F}L$ is a $\mathcal{C}^{s-1}$-diffeomorphism of $T\mathbb{M}$ onto $\mathcal{L}\mathbb{M}$ [12, Lemma 1].

Assumption 2. *For each $r \in \mathbb{R}^n$, the restricted mapping $L|_{T_r\mathbb{R}^n}$ is of class $\mathcal{C}^s(T_r\mathbb{R}^n)$ with $s \geq 2$. Moreover, the Hessian matrix*

$$\left\{ \frac{\partial^2 L(r, \dot r)}{\partial \dot r^i \partial \dot r^j} \right\}_{ij}$$

is positive definite for all $(r, \dot r) \in T\mathbb{R}^n$ and thus, $L(r, \dot r)$ is convex in $\dot r$.

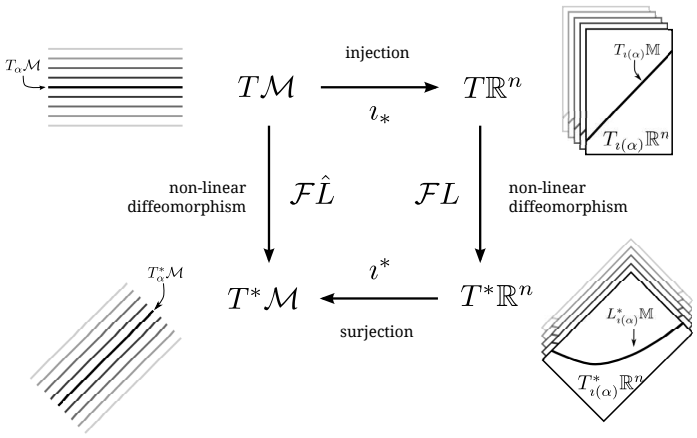

FIG. 2.1. *Commutative diagram showing the relationship between the phase spaces $T^*\mathcal{M}$ and $\mathcal{L}\mathbb{M}$. Notice that $\mathcal{L}\mathbb{M}$ is not a fiber bundle if $\mathcal{F}L$ is nonlinear (i.e., if L is not quadratic in $\dot{r}$).*

In practice, this assumption is not restrictive and is satisfied by most mechanical systems.

Now we depict the relation between the four manifolds introduced so far.

PROPOSITION 2.4. *Consider the Lagrangian function $\hat{L} : T\mathcal{M} \to \mathbb{R}$, defined as*

$$\hat{L} := L \circ (\imath, \imath_*) \,, \tag{2.8}$$

and consider its corresponding fiber derivative $\mathcal{F}\hat{L}$. Let $\imath^*$ be the pullback of $\imath$. Then, the diagram of Fig. 2.1 commutes. Moreover, $\mathcal{F}\hat{L}$ is a $\mathcal{C}^{s-1}$-diffeomorphism of $T\mathcal{M}$ onto $T^*\mathcal{M}$.

Proof. First we will show that $\mathcal{F}\hat{L} = \imath^* \circ \mathcal{F}L \circ \imath_*$. Applying the definition of the fiber derivative (2.5) to $\hat{L}$ gives

$$\langle \mathcal{F}\hat{L}(\xi), \eta \rangle = \left. \frac{\mathrm{d}}{\mathrm{d}s} \right|_{s=0} \hat{L}(\alpha, \xi + s\eta) \,, \tag{2.9}$$

where $\xi, \eta \in T_\alpha\mathcal{M}$. By substituting (2.8) in (2.9) we can see that

$$\langle \mathcal{F}\hat{L}(\xi), \eta \rangle = \left. \frac{\mathrm{d}}{\mathrm{d}s} \right|_{s=0} L(\imath(\alpha), \imath_*(\xi) + s \cdot \imath_*(\eta)) = \langle \mathcal{F}L \circ \imath_*(\xi), \imath_*(\eta) \rangle \tag{2.10}$$

for $\xi, \eta \in T_\alpha\mathcal{M}$. According to the definition of the pullback, Eq. (2.10) can be written as

$$\langle \mathcal{F}\hat{L}(\xi), \eta \rangle = \langle \imath^* \circ \mathcal{F}L \circ \imath_*(\xi), \eta \rangle \,. \tag{2.11}$$

Eq. (2.11) must hold for all ξ and η, so $\mathcal{F}\hat{L} = \imath^* \circ \mathcal{F}L \circ \imath_*$.

In view of the smoothness of $\imath_*$, the restricted mapping $\hat{L}|_{T_q\mathcal{M}}$ is of class $\mathcal{C}^s(T_q\mathcal{M})$. To show that $\mathcal{F}\hat{L}$ is a diffeomorphism, we will write $\hat{L}$ in local coordinates and verify that its Hessian is positive definite [12, Lemma 1]. Applying the chain rule to $L \circ (\imath, \imath_*)$ gives

$$\frac{\partial \hat{L}}{\partial \dot{q}^j} = \frac{\partial \imath^l_*}{\partial \dot{q}^j} \frac{\partial L}{\partial \dot{r}^l} \circ (\imath, \imath_*) \,,$$

where we have used Einstein's summation convention to alleviate the notation[4]. By applying the same rule, we obtain the second derivatives

$$\frac{\partial^2 \hat{L}}{\partial \dot{q}^i \partial \dot{q}^j} = \frac{\partial \imath_*^k}{\partial \dot{q}^i} \frac{\partial \imath_*^l}{\partial \dot{q}^j} \frac{\partial^2 L}{\partial \dot{r}^k \partial \dot{r}^l} \circ (\imath, \imath_*)$$

(recall that $\imath_*$ is linear in $\dot{q}$, so $\partial^2 \imath_*^l / (\partial \dot{q}^i \partial \dot{q}^j) = 0$). Since the Jacobian $\left\{ \partial \imath_*^l / \partial \dot{q}^j \right\}_{lj}$ is full-rank, the Hessian $\left\{ \partial^2 \hat{L} / \partial \dot{q}^i \partial \dot{q}^j \right\}_{ij}$ is positive definite for all $(q, \dot{q})$ belonging to the same coordinate neighborhood. Each fiber $T_q\mathcal{M}$ can be covered using a single coordinate chart and $\mathcal{F}\hat{L}$ is fiber preserving, so $\mathcal{F}\hat{L}$ is a $\mathcal{C}^{s-1}$-diffeomorphism of $T\mathcal{M}$ onto $T^*\mathcal{M}$. □

2.2. The Legendre transform. Once the invertibility of $\mathcal{F}L$ has been established (i.e., by Assumption 2), we can define the Hamiltonian function $H : T^*\mathbb{R}^n \to \mathbb{R}$ as the Legendre transformation [12, 56] of L, with L viewed as function of $\dot{r}$ only, i.e., as

$$H(r, p) = \left(\dot{r}^i p_i - L(r, \dot{r}) \right) \Big|_{\dot{r} = \mathcal{F}L^{-1}(p)} \ . \tag{2.12}$$

(The function H is equal to the total energy, expressed using the redundant coordinates $\{r^i, p_i\}$.) An interesting property of the Legendre transform is that it is its own inverse if L is convex [12, 56]. More precisely, the fiber derivative of H, $\mathcal{F}H : T^*\mathbb{R}^n \to T\mathbb{R}^n$, defined as

$$\langle W, \mathcal{F}H(Z) \rangle = \left. \frac{\mathrm{d}}{\mathrm{d}s} \right|_{s=0} H(a, Z + sW) \ ,$$

where $Z, W \in T_a^*\mathbb{R}^n$, is the inverse of $\mathcal{F}L$. In coordinates,

$$\mathcal{F}H : (r, p) \mapsto (r, \dot{r}) : \dot{r}^i = \frac{\partial H(r, p)}{\partial p_i} \ . \tag{2.13}$$

Also, the Lagrangian can be recovered by setting

$$L(r, \dot{r}) = \left(\dot{r}^i p_i - H(r, p) \right) \Big|_{p = \mathcal{F}H^{-1}(\dot{r})} \ .$$

Likewise, the Hamiltonian function $\hat{H} : T^*\mathcal{M} \to \mathbb{R}$ is defined as

$$\hat{H}(q, \hat{p}) = \left(\dot{q}^i \hat{p}_i - \hat{L}(q, \dot{q}) \right) \Big|_{\dot{q} = \mathcal{F}\hat{L}^{-1}(\hat{p})} \ . \tag{2.14}$$

The invertibility of $\mathcal{F}L$ also implies, in view of (2.3) and (2.6), that

$$\mathcal{L}_a\mathbb{M} = \left\{ Z \in T_a^*\mathbb{R}^n \mid \mathcal{F}L^{-1}(Z)(g^j) = 0 \ , \ j = 1, \ldots, k \right\} \ .$$

Since $\mathcal{F}H$ is the inverse of $\mathcal{F}L$, we have, in coordinates,

$$\mathcal{L}_r\mathbb{M} = \left\{ p \in T_r^*\mathbb{R}^n \ \middle| \ \frac{\partial H(r, p)}{\partial p_j} \frac{\partial g^i(r)}{\partial r^j} = 0 \ , \ i = 1, \ldots, k \right\} \ . \tag{2.15}$$

[4] If the same index appears twice in any monomial term, once as an upper index and once as a lower index, that term is understood to be summed over all possible values of that index [23].

The commutative diagram of Fig. 2.1 suggests

$$\jmath^* := \mathcal{F}L \circ \imath_* \circ \mathcal{F}\hat{L}^{-1} \tag{2.16}$$

as a right inverse for $\imath^*$. Indeed, we have

$$\imath^* \circ \jmath^* = \imath^* \circ \mathcal{F}L \circ \imath_* \circ \mathcal{F}\hat{L}^{-1} = \mathcal{F}\hat{L} \circ \mathcal{F}\hat{L}^{-1} = \mathrm{Id}$$

and, additionally, $\jmath^*\left(T^*\mathcal{M}\right) = \mathcal{L}\mathbb{M}$.

PROPOSITION 2.5. *The Hamiltonians are related by*

$$\hat{H} = H \circ (\imath, \jmath^*) \ .$$

Thus, $\hat{H}$ corresponds to the total energy too, but expressed using the local coordinates $\{q^i, \hat{p}_i\}$.

Proof. Substitution of (2.16) and (2.12) into $H \circ (\imath, \jmath^*)$ shows that

$$H \circ (\imath(q), \jmath^*(\hat{p})) = \left(\dot{r}^i \jmath_i^*(\hat{p}) - L(\imath(q), \dot{r})\right)\big|_{\dot{r}=\mathcal{F}L^{-1}\circ\mathcal{F}L\circ\imath_*\circ\mathcal{F}\hat{L}^{-1}(\hat{p})}$$

or, equivalently,

$$H \circ (\imath(q), \jmath^*(\hat{p})) = \left(\imath_*^i(\dot{q})\jmath_i^*(\hat{p}) - L(\imath(q), \imath_*(\dot{q}))\right)\big|_{\dot{q}=\mathcal{F}\hat{L}^{-1}(\hat{p})} \ .$$

Finally, the Hamiltonian (2.14) is recovered by recalling (2.8) and noting that $\imath_*^i(\dot{q})\jmath_i^*(\hat{p}) = \dot{q}^i(\imath^* \circ \jmath^*(\hat{p}))_i = \dot{q}^i\hat{p}_i$. ∎

2.3. Hamilton's equations. Consider again the local coordinates $\{q^i, \hat{p}_i\}$ on the abstract manifold $T^*\mathcal{M}$. We say that a system is *Hamiltonian* if its trajectories are integral curves of the vector field $D_{\hat{H}} : T^*\mathcal{M} \to T(T^*\mathcal{M})$, where

$$D_{\hat{H}} = \frac{\partial \hat{H}}{\partial \hat{p}_i}\frac{\partial}{\partial q^i} - \frac{\partial \hat{H}}{\partial q^i}\frac{\partial}{\partial \hat{p}_i} \ . \tag{2.17}$$

This vector field unfolds into the more familiar ODE

$$\dot{q} = +\nabla_{\hat{p}}\hat{H}(q, \hat{p}) \tag{2.18a}$$

$$\dot{\hat{p}} = -\nabla_q\hat{H}(q, \hat{p}) \tag{2.18b}$$

(see [2] for more details and a coordinate-free definition of the Hamiltonian vector field).

Let us now describe the implicit model for Hamiltonian systems. In global coordinates $\{r^i, p_i\}$ on $T^*\mathbb{R}^n$, the *implicit* Hamiltonian vector field $X_{H,g} : \mathcal{L}\mathbb{M} \to T(\mathcal{L}\mathbb{M})$, takes the form

$$X_{H,g} = D_H - \lambda_j\frac{\partial g^j}{\partial r^i}\frac{\partial}{\partial p_i} \ , \quad g = 0 \ , \tag{2.19}$$

with

$$D_H = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial r^i} - \frac{\partial H}{\partial r^i}\frac{\partial}{\partial p_i} \tag{2.20}$$

(see [16, 42, 4] for details on the derivation of this equation). The Lagrange multipliers λ_j are defined implicitly by (2.19) and the restriction that the integral curve must lie

on $\mathcal{LM}$. More concretely, by applying $X_{H,g}$ to both sides of the constraints $g^j = 0$, we obtain the *hidden constraints*

$$f^j := X_{H,g}(g^j) = \frac{\partial H}{\partial p_i}\frac{\partial g^j}{\partial r^i} = 0 \ . \tag{2.21}$$

If we apply $X_{H,g}$ to the hidden constraints, then the λ^j appear:

$$X_{H,g}(f^l) = D_H(f^l) - \lambda_j \frac{\partial g^j}{\partial r^i}\frac{\partial f^l}{\partial p_i} = 0 \ . \tag{2.22}$$

Thus, if the matrix

$$\left\{ \frac{\partial g^j}{\partial r^i}\frac{\partial f^l}{\partial p_i} \right\}_{jl} = \left\{ \frac{\partial g^j}{\partial r^i}\frac{\partial^2 H}{\partial p_i \partial p_m}\frac{\partial g^l}{\partial r^m} \right\}_{jl} \tag{2.23}$$

is non-singular on $\mathcal{LM}$, then there are unique λ_j satisfying (2.22) and ensuring that the integral curve stays on $\mathcal{LM}$ —In mechanical systems, λ is the covector of constraint forces that ensure that the constraints are being enforced.

Remark 1. *Since the Hessian of L is positive definite, the Hessian of H is positive definite too [12] (convexity is preserved by the Legendre transform). The matrix* (2.23) *is thus positive definite (hence invertible) on account of Assumptions 1 and 2.*

Remark 2. *The statements $(r,p) \in \mathcal{LM}$ and $g(r) = 0$, $f(r,p) = 0$ are equivalent (cf.* (2.15)*).*

The implicit vector field develops into the semi-explicit DAE[5] of index 2:

$$\dot{r} = +\nabla_p H(r,p) \tag{2.24a}$$
$$\dot{p} = -\nabla_r H(r,p) - G(r)^\top \lambda \tag{2.24b}$$
$$0 = g(r) \ , \tag{2.24c}$$

where $G(r) := \{\partial g^i(r)/\partial r^j\}_{ij}$ is the Jacobian of $g(r)$. A general discussion on DAEs as differential equations on manifolds can be found in [43, 40].

Example: A double planar pendulum. Consider the mechanical system shown in Fig. 2.2. We model the double pendulum as a pair of point masses a and b with coordinates $r^a = (r^{a_x}, r^{a_y})$, $r^b = (r^{b_x}, r^{b_y})$, and masses m_a and m_b, respectively. We set $r := (r^a, r^b) \in \mathbb{R}^4$ (i.e., $n = 4$). The first mass is attached to the origin of the reference frame by a massless bar of length l_a, while the second mass is attached to the first one by a massless bar of length l_b. This is expressed by the pair of constraints ($k = 2$)

$$g^1(r) = \|r^a\|^2 - l_a^2 = 0 \tag{2.25a}$$
$$g^2(r) = \|r^\delta\|^2 - l_b^2 = 0 \tag{2.25b}$$

with $r^\delta := r^b - r^a$. We have

$$\operatorname{rank} G(r) = \operatorname{rank} \begin{pmatrix} r^{a_x} & r^{a_y} & 0 & 0 \\ -r^{\delta_x} & -r^{\delta_y} & r^{\delta_x} & r^{\delta_y} \end{pmatrix} = k \tag{2.26}$$

[5] See [5, p. 16] for a definition of semi-explicit DAEs.

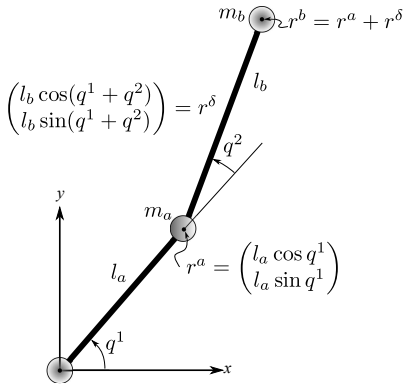

FIG. 2.2. *A double planar pendulum, a simple Hamiltonian system.*

for all $r \in \mathbb{M}$. Therefore, 0 is a regular value of g and $\mathbb{M}$ is an embedded submanifold of $\mathbb{R}^4$.

To obtain the Lagrangian, we write down the kinetic energy of the system, which is given by the simple quadratic form

$$K(\dot{r}) = \frac{1}{2}\dot{r}^\top M \dot{r} \ , \quad M := \begin{pmatrix} m_a\mathbf{I} & \mathbf{0} \\ \mathbf{0} & m_b\mathbf{I} \end{pmatrix} \ , \tag{2.27}$$

where $\mathbf{0}$ and $\mathbf{I}$ are the null and identity elements in $\mathbb{R}^{2\times 2}$. The potential energy is the *linear* function

$$V(r) = \bar{g}(m_a r^{a_y} + m_b r^{b_y}) \ , \tag{2.28}$$

where $\bar{g}$ is the standard gravity. The Lagrangian is $L = K - V$, so the vector of momenta is, according to (2.7), $p = M\dot{r}$. It follows then from (2.27) and (2.28) that the total energy $K + V$ is

$$H(r,p) = \frac{1}{2}p^\top M^{-1} p + \bar{g}(m_a r^{a_y} + m_a r^{a_y}) \ . \tag{2.29}$$

Substituting (2.25) and (2.29) in (2.24) gives

$$\dot{r}^a = m_a^{-1} p_a \tag{2.30a}$$

$$\dot{r}^b = m_b^{-1} p_b \tag{2.30b}$$

$$\begin{pmatrix} \dot{p}_{a_x} \\ \dot{p}_{a_y} \\ \dot{p}_{b_x} \\ \dot{p}_{b_y} \end{pmatrix} = - \begin{pmatrix} 0 \\ \bar{g}m_a \\ 0 \\ \bar{g}m_b \end{pmatrix} - 2 \begin{pmatrix} r^{a_x} & -r^{\delta_x} \\ r^{a_y} & -r^{\delta_y} \\ 0 & r^{\delta_x} \\ 0 & r^{\delta_y} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \tag{2.30c}$$

which, together with (2.25), constitutes a set of DAEs describing the motion of the double pendulum. We note in passing that λ_1 and λ_2 are the magnitudes of the tension along the two bars.

Now we model the double pendulum as an ODE. The dimension of $\mathcal{M}$ is $o = n - k = 2$. Motivated by Fig. 2.2, we choose $q^1 \in (-\pi, \pi)$ and $q^2 \in (-\pi, \pi)$ as local coordinates for $\mathcal{M}$. Eq. (2.26) ensures the existence of an embedding $r = \imath(q)$ such

that $g \circ \imath \equiv 0$. It can be readily verified that such embedding is

$$\begin{pmatrix} r^{a_x} \\ r^{a_y} \\ r^{b_x} \\ r^{b_y} \end{pmatrix} = \begin{pmatrix} l_a \cos q^1 \\ l_a \sin q^1 \\ l_a \cos q^1 + l_b \cos q^{\mathrm{t}} \\ l_a \sin q^1 + l_b \sin q^{\mathrm{t}} \end{pmatrix} , \quad q^{\mathrm{t}} := q^1 + q^2 . \tag{2.31}$$

By differentiating both sides of (2.31) we obtain the following mapping between velocities:

$$\dot{r} = \begin{pmatrix} -l_a \sin q^1 & 0 \\ l_a \cos q^1 & 0 \\ -l_a \sin q^1 - l_b \sin q^{\mathrm{t}} & -l_b \sin q^{\mathrm{t}} \\ l_a \cos q^1 + l_b \cos q^{\mathrm{t}} & l_b \cos q^{\mathrm{t}} \end{pmatrix} \dot{q} , \quad q = \begin{pmatrix} q^1 \\ q^2 \end{pmatrix} . \tag{2.32}$$

To obtain an expression for the kinetic energy in terms of the generalized positions and velocities, we substitute (2.32) in (2.27), which gives

$$K \circ \imath_*(\dot{q}) = \frac{1}{2} \dot{q}^\top \hat{M}(q) \dot{q} \tag{2.33}$$

with

$$\hat{M}(q) = \begin{pmatrix} m_{\mathrm{t}} l_a^2 + m_b l_b^2 + 2 m_b l_a l_b \cos q^2 & m_b l_b^2 + m_b l_a l_b \cos q^2 \\ m_b l_b^2 + m_b l_a l_b \cos q^2 & m_b l_b^2 \end{pmatrix} \tag{2.34}$$

and $m_{\mathrm{t}} = m_a + m_b$. The expression for the potential energy

$$V \circ \imath(q) = \bar{g} m_a l_a \sin q^1 + \bar{g} m_b (l_a \sin q^1 + l_b \sin q^{\mathrm{t}}) \tag{2.35}$$

is obtained by substituting (2.31) in (2.28). The vector of momenta is $\hat{p} = \hat{M}(q)\dot{q}$ and the total energy is

$$\hat{H}(q, \hat{p}) = \frac{1}{2} \hat{p}^\top \hat{M}(q)^{-1} \hat{p} + \bar{g} \left(m_{\mathrm{t}} l_a \sin q^1 + m_b l_b \sin q^{\mathrm{t}} \right) . \tag{2.36}$$

Finally, Eq. (2.18) states that the motion of the system is described by

$$\dot{q} = \hat{M}(q)^{-1} \hat{p} \tag{2.37a}$$

$$\dot{\hat{p}} = -\nabla_q V(q) - \nabla_q \left(\frac{1}{2} \hat{p}^\top \hat{M}(q)^{-1} \hat{p} \right) . \tag{2.37b}$$

The main point of this example is summarized in the following the remark.

REMARK 3. *We have written two representations for the same system, one has the form of an ODE* (2.37) *and the other is a DAE* (2.30)*. For the DAE the Hamiltonian function* (2.29) *is separable, i.e., the kinetic energy does not depend on* r*: The inertia matrix* M *is a* constant *diagonal matrix. Moreover, the potential energy* (2.28) *is* linear*, which results in a constant gradient. On the other hand, the Hamiltonian* (2.36) *that appears in the ODE is not separable (the inertia matrix depends on* q*) and it is composed of transcendental functions.*

There is for sure a cost associated to the simpler Hamiltonian function: The dimensions of q and $\hat{p}$ that describe the system using the vector field (2.37) are lower than the dimensions of r and p in (2.30). The price is small, however, when it comes to constructing Hamiltonian sample-data models.

2.4. Energy conservation and symplecticity. It is well known that the flow generated by (2.17) preserves the Hamiltonian function [52, 37, 38]. In other words, Hamiltonian systems conserve energy, a property that is easily shown by computing $\mathcal{L}_{D_{\hat{H}}}\hat{H}$, the Lie derivative of $\hat{H}$ along the flow generated by $D_{\hat{H}}$. The implicit vector field (2.19), on the other hand, represents the same system as (2.17), so the flow generated by it must surely preserve the Hamiltonian function too [16]. Indeed,

$$\mathcal{L}_{X_{H,g}}H = X_{H,g}(H) = D_H(H) - \lambda_j \frac{\partial g^j}{\partial r^i}\frac{\partial H}{\partial p_i} = \lambda_j f^j \ , \tag{2.38}$$

where the third equality holds because

$$D_H(H) = \frac{\partial H}{\partial p_i}\frac{\partial H}{\partial r^i} - \frac{\partial H}{\partial r^i}\frac{\partial H}{\partial p_i} = 0 \tag{2.39}$$

and because of the definition of f given in (2.21). Eq. (2.38) shows that

$$(\mathcal{L}_{X_{H,g}}H)(r,p) = 0$$

for all $(r,p) \in \mathcal{LM}$ (cf. Remark 2), so H remains constant along the system trajectories.

It is also well known that, besides the 0-form $\hat{H}$, Hamiltonian flows preserve the 2-form[6]

$$\hat{\omega} := \mathrm{d}q^i \wedge \mathrm{d}\hat{p}_i$$

(Einstein's summation convention implied, with a slight abuse of notation), which acts on vectors of $T(T^*\mathcal{M})$. In standard coordinates $\{\dot{q}^i, \hat{\dot{p}}_i\}$ on $T(T^*\mathcal{M})$, $\hat{\omega}$ takes the form

$$\hat{\omega}(\xi,\eta) = \xi^\top \hat{J}\eta \ , \quad \hat{J} = \begin{pmatrix} \mathbf{0}_o & \mathbf{I}_o \\ -\mathbf{I}_o & \mathbf{0}_o \end{pmatrix} \ , \quad \xi,\eta \in T(T^*\mathcal{M}) \ .$$

The invariance of $\hat{\omega}$ with respect to $D_{\hat{H}}$ can be established by showing that the Lie derivative $\mathcal{L}_{D_{\hat{H}}}\hat{\omega}$ is equal to zero, the case being analogous to the one on the conservation of $\hat{H}$.

The 2-form $\hat{\omega}$ can be interpreted as the sum of the oriented areas of the parallelograms formed by the projections of ξ and η onto the coordinate planes $(\dot{q}^i, \hat{\dot{p}}_i)$. Thus, a mapping that preserves $\hat{\omega}$ is simply an area preserving mapping. Let us elaborate on this notion: let $A : \mathbb{R}^{2o} \to \mathbb{R}^{2o}$ be a linear mapping. The mapping is called *symplectic* if $(Aa)^\top \hat{J} Ab = a^\top \hat{J} b$ for all $a, b \in \mathbb{R}^{2o}$ (see Fig. 2.3) or, equivalently, if

$$A^\top \hat{J} A = \hat{J} \ .$$

This definition can be extended to general differentiable nonlinear mappings on $T^*\mathcal{M}$.

Definition 2.6. *A differentiable mapping $\hat{\phi} : T^*\mathcal{M} \to T^*\mathcal{M}$ is called* symplectic *if*

$$\hat{\phi}^*\hat{\omega} = \hat{\omega} \ . \tag{2.40}$$

Spelling out the definition of the pullback, Eq. (2.40) can be alternatively written as $\hat{\omega}(\hat{\phi}_*\xi, \hat{\phi}_*\eta) = \hat{\omega}(\xi,\eta)$ for all $\xi, \eta \in T_x(T^*\mathcal{M})$. In standard coordinates for $T(T^*\mathcal{M})$, $\hat{\phi}$ is symplectic if its Jacobian matrix $\hat{\Phi}$ is everywhere symplectic, i.e., if

$$\hat{\Phi}(q,\hat{p})^\top \hat{J} \hat{\Phi}(q,\hat{p}) = \hat{J} \qquad \text{for all } (q,\hat{p}) \in T^*\mathcal{M} \ .$$

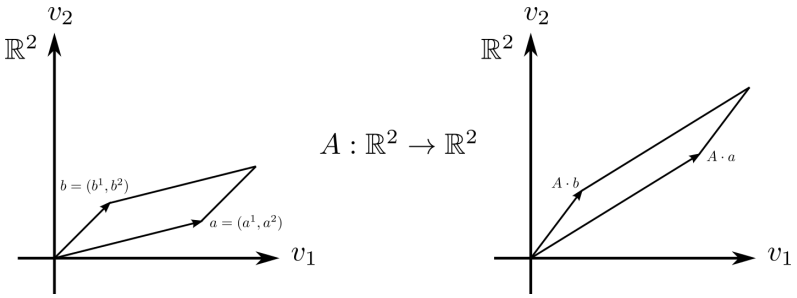

FIG. 2.3. *Symplecticity (area preservation) of a linear mapping.*

Remarkably, the flow $\hat{\phi}_t$ that is generated by the Hamiltonian vector field (2.17) is symplectic. Again, since the implicit Hamiltonian vector field $X_{H,g}$ refers to the same physical system as $D_{\hat{H}}$, one might reasonably expect $X_{H,g}$ to generate a symplectic flow. Similarly to $\hat{\omega}$, we define

$$\omega := \mathrm{d}r^i \wedge \mathrm{d}p_i \ , \tag{2.41}$$

which acts on vectors of $T(T^*\mathbb{R}^n)$. In standard coordinates $\{\dot{r}^i, \dot{p}_i\}$ on $T(T^*\mathbb{R}^n)$, ω takes the form

$$\omega(\xi, \eta) = \xi^\top J \eta \ , \quad J = \begin{pmatrix} \mathbf{0}_n & \mathbf{I}_n \\ -\mathbf{I}_n & \mathbf{0}_n \end{pmatrix} \ , \quad \xi, \eta \in T(T^*\mathbb{R}^n) \ .$$

THEOREM 2.7. *[16, 24] Let H be twice continuously differentiable. The flow $\phi_t : \mathcal{LM} \to \mathcal{LM}$ of $X_{H,g}$ (2.19) is a symplectic transformation on $\mathcal{LM}$, i.e.,*

$$\phi_t^* \omega = \omega$$

for every t for which ϕ_t is defined.

REMARK 4. *The converse statement, that* every symplectic flow ϕ_t solves Hamilton's equations for some H, *is also true, so symplecticity is a characteristic property of Hamiltonian systems [2]. This does not translate to the case of energy conservation, i.e., while every Hamiltonian system conserves energy, not every energy-conserving system is Hamiltonian.*

3. Adding external ports. With the purpose of taking external forces and dissipation into consideration, a system of the form (2.17) can be fit up with a pair of port variables (u, y), giving rise to a *PH system* [28, 52, 37] that is described by the vector field $X_{\hat{H},u} : T^*\mathcal{M} \times (\mathbb{R}^m)^* \to T(T^*\mathcal{M})$:

$$X_{\hat{H},u} = D_{\hat{H}} + u_l \hat{U}_i{}^l \frac{\partial}{\partial \hat{p}_i} \ , \tag{3.1}$$

where $u \in (\mathbb{R}^m)^*$ is the controlled or input variable,

$$y^l = \hat{U}_i{}^l \frac{\partial \hat{H}}{\partial \hat{p}_i} \tag{3.2}$$

[6] See [2] for a coordinate-free definition.

is the dependent or output variable ($y \in \mathbb{R}^m$) and $\hat{U}_i{}^l$ are maps from $\mathcal{M}$ to $\mathbb{R}$. Carrying forward this idea over an implicit Hamiltonian system (2.19) gives the implicit control vector field $X_{H,u,g} : \mathcal{LM} \times (\mathbb{R}^m)^* \to T(\mathcal{LM})$:

$$X_{H,u,g} = D_H + \left(u_l U_i{}^l - \lambda_j \frac{\partial g^j}{\partial r^i} \right) \frac{\partial}{\partial p_i} , \quad g = 0 , \tag{3.3}$$

with the output given by

$$y^l = U_i{}^l \frac{\partial H}{\partial p_i} . \tag{3.4}$$

The vector field (3.3) and the output unravel as

$$\begin{aligned} \dot{r} &= +\nabla_p H(r,p) \\ \dot{p} &= -\nabla_r H(r,p) - G(r)^\top \lambda + U(r)u \\ y &= U(r)^\top \nabla_p H(r,p) \\ 0 &= g(r) . \end{aligned}$$

By analogy with the results described in Sec. 2.3, one can determine the Lagrange multipliers λ explicitly. The constraints $f^a = 0$ imply that

$$X_{H,u,g}(f^a) = D_H(f^a) + u_l U_i{}^l \frac{\partial f^a}{\partial p_i} - \lambda_j \frac{\partial g^j}{\partial r^i} \frac{\partial f^a}{\partial p_i} = 0 , \tag{3.5}$$

from where it follows that, as long as (2.23) is non-singular, there are unique λ_j (in general dependent on u as well as on r and p) such that $X_{H,u,g}(f^a) = 0$ and the integral curve stays on $\mathcal{LM}$.

REMARK 5. *It follows from* (3.2) *and* (3.4) *that* $\hat{U}_i{}^l$ *and* $U_j{}^l$ *must satisfy*

$$\hat{U}_i{}^l = \imath_*{}^j{}_i \cdot (U \circ \imath)_j{}^l . \tag{3.6}$$

A system described by (3.3) is called an *implicit port-Hamiltonian system* [52], and can also be understood as a particular instance of the generalized Dirac structures introduced in [10], or as an example of the behavioral systems studied in [39].

3.1. Passivity. It can be readily seen that an implicit PH system described by (3.3) no longer preserves H. The Lie derivative of H is now

$$\mathcal{L}_{X_{H,u,g}}(H) = X_{H,u,g}(H) = D_H(H) + u_l U_i{}^l \frac{\partial H}{\partial p_i} - \lambda_j \frac{\partial g^j}{\partial r^i} \frac{\partial H}{\partial p_i} = u_l y^l - \lambda_j f^j ,$$

which establishes the power balance

$$\mathcal{L}_{X_{H,u,g}} H = u_l y^l \qquad \text{for all } (r,p) \in \mathcal{LM} . \tag{3.7}$$

Since the product $u_l y^l$ is equal to the rate of change in energy, we say that (u, y) is a power-conjugated pair of port variables. If, in addition, the restriction of H to $\mathcal{LM}$ is bounded from below, i.e., if the image of $\mathcal{LM}$ under H is bounded from below, we say that (3.3) is passive, or more precisely, *lossless*. Boundedness of H can be easily assessed using the following proposition.

PROPOSITION 3.1. *Suppose that the potential energy V is lower semi-continuous and $\mathbb{M}$ is compact. Then, $H(\mathcal{L}\mathbb{M})$ is bounded from below (hence, the vector field* (3.3) *describes a loss-less system).*

Proof. Since a lower semi-continuous function attains its minimum on a compact set, $V(\mathbb{M})$ is bounded from below [45]. The function K is positive definite (Assumption 2), hence also bounded from below. The sum of two bounded functions is again bounded, so H is bounded from below. □

For many systems (like the double pendulum), establishing compactness of $\mathbb{M}$ is straightforward.

Example: A double planar pendulum (continued). Suppose that the double pendulum (Fig. 2.2) is actuated, so that we can apply torques u_1 and u_2 to the joints that correspond to the angles q^1 and q^2, respectively. The resulting linear forces are then $U^1 u_1$ and $U^2 u_2$ with

$$U^1 := \{U_i{}^1\}_i = \begin{pmatrix} -r^{a_y} \\ r^{a_x} \\ 0 \\ 0 \end{pmatrix} \frac{1}{l_a^2} \tag{3.8}$$

and

$$U^2 := \{U_i{}^2\}_i = \begin{pmatrix} r^{\delta_y} \\ -r^{\delta_x} \\ -r^{\delta_y} \\ r^{\delta_x} \end{pmatrix} \frac{1}{l_b^2} - U^1 \ . \tag{3.9}$$

The manifold defined by (2.25) is compact and the potential energy (2.28) is continuous, which confirms that the double pendulum is passive with passive outputs

$$y^l = U_i{}^l \frac{\partial H}{\partial p_i} = U_i{}^l \dot{r}^i \ .$$

Upon substitution of (2.31) and (2.32) in (3.6), we have,

$$\begin{aligned}
\hat{U}_1{}^1 &= \imath_*{}^j{}_1 \cdot (U \circ \imath)_j{}^1 = -l_a \sin q^1 \cdot (-\sin q^1/l_a) + l_a \cos q^1 \cdot \cos q^1/l_a = 1 \\
\hat{U}_2{}^1 &= \imath_*{}^j{}_2 \cdot (U \circ \imath)_j{}^1 = 0 \\
\hat{U}_1{}^2 &= \imath_*{}^j{}_1 \cdot (U \circ \imath)_j{}^2 = -l_a \sin q^1 \cdot \sin q^{\mathrm{t}}/l_b + l_a \cos q^1 \cdot (-\cos q^{\mathrm{t}}/l_b) + \\
&\quad (-l_a \sin q^1 - l_b \sin q^{\mathrm{t}}) \cdot (-\sin q^{\mathrm{t}}/l_b) + (l_a \cos q^1 + l_b \cos q^{\mathrm{t}}) \cdot \cos q^{\mathrm{t}}/l_b - \hat{U}_1{}^1 \\
&\quad = 0 \\
\hat{U}_2{}^2 &= \imath_*{}^j{}_2 \cdot (U \circ \imath)_j{}^2 = l_a \cos q^1 \cdot \sin q^{\mathrm{t}}/l_b + l_a \sin q^1 \cdot (-\cos q^{\mathrm{t}}/l_b) + \\
&\quad (l_a \cos q^1 + l_b \cos q^{\mathrm{t}}) \cdot (-\sin q^{\mathrm{t}}/l_b) + (l_a \sin q^1 + l_b \sin q^{\mathrm{t}}) \cos q^{\mathrm{t}}/l_b - \hat{U}_2{}^1 \\
&\quad = 1 \ ,
\end{aligned}$$

so in local coordinates, the passive outputs correspond to the angular velocities

$$y^1 = \hat{U}_1{}^1 \dot{q}^1 + \hat{U}_2{}^1 \dot{q}^2 = \dot{q}^1 \quad \text{and} \quad y^2 = \hat{U}_1{}^2 \dot{q}^1 + \hat{U}_2{}^2 \dot{q}^2 = \dot{q}^2 \ .$$

3.2. The evolution of the symplectic form. With the inclusion of the control variable u, we can no longer expect the flow of (3.3) to be symplectic either, as we will show by evaluating the Lie derivative of ω along $X_{H,u,g}$.

PROPOSITION 3.2. *The Lie derivative of* ω (2.41) *restricted to* $\mathcal{L}\mathbb{M}$ *satisfies*

$$\mathcal{L}_{X_{H,u,g}}\omega\big|_{\mathcal{L}\mathbb{M}} = \mathrm{d}r^i \wedge \mathrm{d}(u_l U_i{}^l) \; . \tag{3.10}$$

Proof. Recall first that the Lie derivative of a general l-form α along a vector field X can be computed using Cartan's formula [1, Ch. 11] as

$$\mathcal{L}_X\alpha = \mathrm{d}\left(i_X\alpha\right) + i_X \mathrm{d}\alpha \; . \tag{3.11}$$

That is, the l-form $\mathcal{L}_X\alpha$ is the sum of two l-forms. The first one, $\mathrm{d}\left(i_X\alpha\right)$, is the exterior derivative of $i_X\alpha$, which is the contraction [46, 23] of α on X. The second l-form is the contraction of $\mathrm{d}\alpha$ on X.

According to (3.11), we have

$$\begin{aligned}\mathcal{L}_{X_{H,u,g}}\omega &= \mathrm{d}\left(i_{X_{H,u,g}}(\mathrm{d}r^i \wedge \mathrm{d}p_i)\right) + i_{X_{H,u,g}}\mathrm{d}(\mathrm{d}r^i \wedge \mathrm{d}p_i) \\ &= \mathrm{d}\Big((i_{X_{H,u,g}}\mathrm{d}r^i) \wedge \mathrm{d}p_i - \mathrm{d}r^i \wedge (i_{X_{H,u,g}}\mathrm{d}p_i)\Big) + 0 \; ,\end{aligned}$$

where the second equality follows from the anti-derivation property of the contraction and the wedge product [1, p. 152] and the fact that $\mathrm{d}^2\beta = 0$ for any differential form β. Performing the contraction of $\mathrm{d}r^i$ and $\mathrm{d}p_i$ on $X_{H,u,g}$ gives

$$\mathcal{L}_{X_{H,u,g}}\omega = \mathrm{d}\left(\frac{\partial H}{\partial p_i}\mathrm{d}p_i + \mathrm{d}r^i\left(\frac{\partial H}{\partial r^i} - u_l U_i{}^l + \lambda_j \frac{\partial g^j}{\partial r^i}\right)\right)$$

or

$$\mathcal{L}_{X_{H,u,g}}\omega = \mathrm{d}\left(\mathrm{d}H - u_l U_i{}^l \mathrm{d}r^i + \lambda_j \mathrm{d}g^j\right) = \mathrm{d}r^i \wedge \mathrm{d}(u_l U_i{}^l) + \mathrm{d}\lambda_j \wedge \mathrm{d}g^j \tag{3.12}$$

(see [46, Ch. 4], [1, Ch. 11] or [24] for a list of rules for the exterior derivative, contraction and wedge product). The constraint $g = 0$ implies the restrictions $\mathrm{d}g^j = 0$ on the differentials of g^j, i.e., $\mathrm{d}g^j(\xi) = 0$ for all $\xi \in T_x(\mathcal{L}\mathbb{M})$, so (3.12) reduces to (3.10). ∎

REMARK 6. *Equation* (3.10) *shows that, when* $u \equiv 0$*, the Lie derivative of* ω *is equal to zero, so* ω *remains constant and in consequence the flow is symplectic (this in fact proves Theorem 2.7).*

There is another particular case in which the flow can be made symplectic: Suppose that u_l are functions on $\mathbb{M}$ such that

$$\frac{\partial V_\mathrm{a}}{\partial r^i} = u_l U_i{}^l \tag{3.13}$$

for some artificial potential function $V_\mathrm{a} : \mathbb{M} \to \mathbb{R}$. Then,

$$\mathcal{L}_{X_{H,u,g}}\omega\big|_{\mathcal{L}\mathbb{M}} = \mathrm{d}r^i \wedge \mathrm{d}(u_l U_i{}^l) = \mathrm{d}r^i \wedge \frac{\partial^2 V_\mathrm{a}}{\partial r^j \partial r^i}\mathrm{d}r^j = 0 \; .$$

Equation (3.13) appears, e.g., in the context of passivity based control, when performing potential energy shaping [51, 35]. In general, however, u destroys symplecticity.

	Explicit model	**Implicit model**	**Corresponding map**
Positions	q	r	$r = \imath(q)$
Velocities	$\dot{q}$	$\dot{r}$	$\dot{r} = \imath_*(\dot{q})$
Momenta	$\hat{p}$	p	$p = \jmath^*(\hat{p})$
Control vector fields	$\hat{U}_i{}^l \frac{\partial}{\partial \hat{p}_i}$	$U_i{}^l \frac{\partial}{\partial p_i}$	$\hat{U}_i{}^l(q) = \imath_*{}^j{}_i(q) \cdot U_j{}^l(\imath(q))$
Lagrangian	$\hat{L}(q, \dot{q})$	$L(r, \dot{r})$	$\hat{L}(q, \dot{q}) = L(\imath(q), \imath_*(q))$
Hamiltonian	$\hat{H}(q, \hat{p})$	$H(r, p)$	$\hat{H}(q, \hat{p}) = H(\imath(q), \jmath^*(\hat{p}))$
Symplectic form	$\hat{\omega} = \mathrm{d}q^i \wedge \mathrm{d}\hat{p}_i$	$\omega = \mathrm{d}r^i \wedge \mathrm{d}p_i$	

TABLE 3.1
Relationship between the implicit and explicit representations of a PH system.

3.3. Non-affine port-Hamiltonian systems. While writing both, the explicit (3.1) and implicit (3.3) representations of a PH system, it was assumed that the control variables enter the vector field affinely. Many physical systems exhibit this property so, from a practical point of view, this assumption is not too restrictive. However, for the purpose of performing backward error analysis, we will need to consider PH systems for which the control might enter in a nonaffine way. Motivated by Remark 6, we propose the following extended definition of a PH system.

DEFINITION 3.3. *A controlled vector field $X : \mathfrak{M} \times (\mathbb{R}^m)^* \to T\mathfrak{M}$, with $\mathfrak{M} = T^*\mathcal{M}$ if given explicitly (i.e., by an ODE) and $\mathfrak{M} = \mathcal{L}\mathbb{M}$ if given implicitly (i.e. by a semi-explicit DAE), is said to be* port-Hamiltonian *if $u \equiv 0$ implies that the generated flow is symplectic.*

LEMMA 3.4. *A (not necessarily affine) smooth PH system described by a vector field X can always be decomposed as*

$$X = X_0 + u_l Z^l \ ,$$

where $X_0 : \mathfrak{M} \to T\mathfrak{M}$ is a Hamiltonian vector field and $Z^l : \mathfrak{M} \times (\mathbb{R}^m)^* \to T\mathfrak{M}$ are the input vector fields. Hence, X satisfies the power balance

$$X(H) = u_l y^l$$

for some real-valued function H and real-valued output functions $y^l = Z^l(H)$. (The output functions might depend directly on u as well as on x.)

Proof. Following [25], we first show that a smooth control vector field can be split into a drift and a set of vector fields having u factored out. Let us define the drift $X_0 : \mathfrak{M} \to T\mathfrak{M}(x)$ as

$$X_0(x) = X(x, 0) \quad \text{for all } x \in \mathfrak{M}$$

and let us define the vector fields W^l by the equations[7]

$$W^l(\alpha) = \frac{\partial}{\partial u_l}\Big(X(\alpha)\Big) \quad \text{for all } \alpha \in \mathcal{C}^\infty(\mathfrak{M}, \mathbb{R}) \ .$$

It follows from the chain rule that

$$u_l W^l(x, \theta u)(\alpha) = \frac{\mathrm{d}}{\mathrm{d}\theta}\Big(X(x, \theta u)\Big)(\alpha) \ .$$

[7] If given explicitly, the i'th element W^{il} of the l'th vector field W^l is obtained by differentiating the i'th component X^i of X with respect to u_l, $W^{il} = \frac{\partial X^i}{\partial u_l}$.

Upon integration on both sides of the equation we arrive at

$$u_l \int_0^1 W^l(x,\theta u)(\alpha)\mathrm{d}\theta = X(x,\theta u)(\alpha)\Big|_0^1 = X(x,u)(\alpha) - X_0(x)(\alpha)\ .$$

Therefore, we have

$$X = X_0 + u_l Z^l\ , \tag{3.14}$$

where the input vector fields Z^l are defined by

$$Z^l(x,u)(\alpha) = \int_0^1 W^l(x,\theta u)(\alpha)\mathrm{d}\theta \quad \text{for all } \alpha \in \mathcal{C}^\infty(\mathfrak{M},\mathbb{R})\ .$$

It follows from the hypothesis that X_0 generates a symplectic flow, so it is a Hamiltonian vector field and satisfies $X_0(H) = 0$ for some real-valued function H. Applying X to H shows that

$$X(H) = X_0(H) + u_l Z^l(H) = u_l y^l\ ,$$

with $y^l = Z^l(H)$. □

Notice that, for an affine PH system, the formulae of this lemma recover the output functions (3.4) (or (3.2), depending on the setting) with $Z^l = {U_j}^l \frac{\partial}{\partial p_j}$, that is,

$$y^l = Z^l(H) = {U_i}^l \frac{\partial H}{\partial p_i}\ .$$

4. Sampled-data port-Hamiltonian systems. In this section we derive discrete-time (or sampled-data) models for PH systems given by implicit vector fields. The PH vector field is split in two components: the vector field describing an unconstrained system with state-space equal to the whole $T^*\mathbb{R}^n$, and a vector field containing the Lagrange multipliers, the one that maintains the trajectories on the submanifold $\mathcal{LM}$. Splitting the vector field simplifies the computation of the sampled-data model by decomposing the problem into two simpler sub-problems.

Computing a sampled-data model basically amounts to computing an approximate solution of the differential equations during a small interval of time. This problem has been studied extensively in the literature of numerical analysis, from which we borrow some results and terminology. In numerical analysis, a *sampled-data model* is the central component of an *integration method* or a *numerical integrator*, up to the point that these terms are used interchangeably.

Before stating the main result of this section, we define sampled-data models and discuss some important properties of numerical integrators, like order and symmetry.

Suppose that there is a sequence of commands $\{u_{l\alpha}\}_{\alpha\in\mathbb{N}}$. Each command in the sequence arrives at the discrete instants of time $\alpha = 0, h, 2h, \ldots$ where h is a positive real number —These commands could be generated, e.g., by a computer program— Suppose further that a zero-order hold transforms this sequence into piecewise constant controls

$$u_l(t) \equiv u_{l\alpha} \quad \text{for } t \in [\alpha h, \alpha h + h)\ , \tag{4.1}$$

which are fed into a PH system. Let $\phi_t(x_0, u(\,\cdot\,))$ be the integral curve of the non-autonomous vector field (3.3) with control (4.1) and passing through $x_0 \in \mathcal{LM}$ at $t = 0$. Let

$$y^l(t) = \left({U_i}^l \frac{\partial H}{\partial p_i}\right) \circ \phi_t(x_0, u(\,\cdot\,))$$

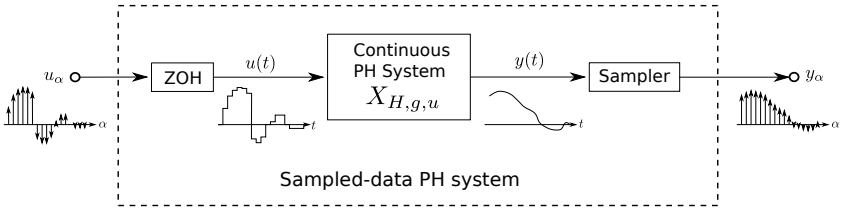

FIG. 4.1. *Sampled-data PH system with sampling period h. The zero-order hold produces piecewise constant inputs $u_l(t) \equiv u_{l\alpha}$ for $t \in [\alpha h, \alpha h + h)$, which is fed to the continuous-time PH system. The output is then sampled to generate the discrete-time output sequences $\{y^l{}_\alpha\}$.*

be the corresponding outputs and let $\{y^l{}_\alpha\}$ with $y^l{}_\alpha := y^l(\alpha h)$ be the sequence obtained by sampling them at discrete instants of time αh (see Fig. 4.1). We call the resulting system a *sampled-data port-Hamiltonian system*.

Mathematical models for sampled-data systems arise in diverse circumstances. In the direct approach to digital control, i.e., as opposed to the emulation of continuous control laws, the design of the controller is performed in discrete time, the designer working directly over a sampled-data model. When designing directly in discrete time, the controller can be directly implemented on a digital device. Also, it is possible to exploit the advantages of switched controls or, e.g., the multi-rate control techniques presented in [30].

Computing the sampled-data model for a given nonlinear system basically amounts to computing the solution ϕ_t of the corresponding ODE or DAE, which is in general impossible to do analytically, so one has to settle for an approximate solution.

When simulating the behavior of dynamic systems, a discrete-time model of the continuous system is also used for computing a numerical solution to the initial-value problem. Needless to say, many different integration methods (or *methods* for short) can be found in the literature of numerical analysis. Let us focus on integration methods for autonomous systems first and let us further restrict our attention to one-step methods defined by a transformation

$$\psi_h : x_\alpha \mapsto x_{\alpha+1} \ ,$$

where the constant step-size h is regarded as a parameter of the method[8]. For a given initial condition x_0 in the phase space, ψ_h is applied recursively to generate a discrete flow $x_1, x_2, x_3, \ldots$ that approximates the true flow $\phi_h(x_0), \phi_{2h}(x_0), \phi_{3h}(x_0), \ldots$ of a given vector field X at time instants $h, 2h, 3h, \ldots$ In this sense, the map ψ_h is a discrete-time approximation of ϕ_h (or a sampled-data model of ϕ_h).

DEFINITION 4.1. *A one-step method has* order *s if the local error satisfies*[9]

$$\psi_h(x_0) - \phi_h(x_0) = \mathcal{O}(h^{s+1}) \quad \textit{as} \quad h \to 0 \tag{4.2}$$

uniformly in x_0. A one-step method is said to be consistent if $s \geq 1$.

[8]For a more general method, the value of the $x_{\alpha+1}$ need not depend only on x_α, but may also depend on the previous values $x_{\alpha-1}, x_{\alpha-2}, \ldots$ (a multi-step method). Also, the value of h need not be constant in general.

[9]We use big-O notation when quantifying approximation errors, i.e., for a given pair of functions $e_1(h)$, $e_2(h)$, we write $e_1(h) = \mathcal{O}(e_2(h))$ as $h \to a$ as shorthand for

$$\limsup_{h \to a} \frac{\|e_1(h)\|}{\|e_2(h)\|} < \infty \ .$$

4.1. Symplectic methods. If the sampled-data model approximates the discrete-time behavior of a Hamiltonian system, one could hope for ψ_h to inherit its fundamental qualitative properties: energy conservation and symplecticity. Unfortunately, it is not possible to preserve H and ω simultaneously, unless ψ_h agrees with the exact flow ϕ_h up to a reparametrization of time [11]. For this reason, one has to choose either in favor of one or the other invariant[10]. Energy conserving methods have received some attention [14, 48, 15, 19, 20], but in light of Remark 4, most of the literature focuses on symplectic integration algorithms (see [16, 24, 8] and references therein). A comparison between both approaches is done in [47] for the rigid body.

A theoretical advantage of constructing a symplectic one-step method is that, even though ψ_h only approximates ϕ_h up to the s'th order, it coincides *exactly* (if one disregards convergence issues) with the flow of another Hamiltonian system, a modified Hamiltonian system described by a *modified differential equation*.

THEOREM 4.2. *[16, p. 352] A symplectic method $\psi_h : \mathcal{LM} \to \mathcal{LM}$ for the constrained Hamiltonian system* (2.24) *has a modified equation that is locally of the form*

$$\dot{r} = +\nabla_p \tilde{H}(r,p) \tag{4.3a}$$

$$\dot{p} = -\nabla_r \tilde{H}(r,p) - G(r)^\top \tilde{\lambda} \tag{4.3b}$$

$$0 = g(r) \tag{4.3c}$$

with $\tilde{H} = H + hH_2 + h^2 H_3 + \dots$ Furthermore,

$$\frac{\partial H_j(r,p)}{\partial p_i} \frac{\partial g^l(r)}{\partial r^i} = 0 \,, \quad \textit{for all } (r,p) \in \mathcal{LM} \,,$$

all $l = 1, \dots, k$ and all j. Note that the actual value of the Legendre multipliers $\tilde{\lambda}$ differ in general from those obtained for the original system (2.24).

In other words, for an initial condition x_0, $\psi_h(x_0)$ is equal to the solution of (4.3) at time $t = h$.

Notice that (4.2) provides information about the difference between the actual flow ϕ_h of X and the approximate discrete flow ψ_h. This is the kind of information that *forward* error analysis aims at. While certainly useful as an indicator of the quality of the approximation, Eq. (4.2) only evaluates the behavior of the approximate flow on the first iteration, but says nothing about its long time behavior. From (4.2) alone we cannot infer anything about the error $x_\alpha - \phi_{\alpha h}(x_0)$ when α is large, so we do not know if errors accumulate or if they average out to zero.

On the other hand, Theorem 4.2 tells us that if ψ_h is symplectic, then there exists a modified continuous system whose flow coincides exactly with the discrete flow generated by ψ_h. The modified system (4.3) preserves the Hamiltonian structure of the original system (2.24) and it is 'close' to it in the sense that $\tilde{H} = H + \mathcal{O}(h^s)$ for a method of order s. In other words, a symplectic integration method preserves the original 2-form ω and a different (but close) Hamiltonian function. This property guaranties that the good behavior of the integration scheme is maintained during many iterations, giving a global nature to the local property (4.2). This observation is at the center of *backward error analysis* [16].

[10] For particular Hamiltonians there might be other invariants, such as momentum or angular momentum, but in general there need not be.

4.2. Splitting methods. A practical advantage of symplectic schemes is their suitability for using splitting methods. To illustrate the idea, consider again the Hamiltonian vector field $D_{\hat{H}}$ on the abstract manifold $T^*\mathcal{M}$ (2.17). If the Hamiltonian function is separable, i.e., if it can be written as $\hat{H}(q,\hat{p}) = \hat{H}_{\mathrm{a}}(q) + \hat{H}_{\mathrm{b}}(\hat{p})$, then the vector field can be spilt into two Hamiltonian vector fields

$$D_{\hat{H}_{\mathrm{a}}} = -\frac{\partial \hat{H}_{\mathrm{a}}}{\partial q^i}\frac{\partial}{\partial \hat{p}_i} \quad \text{and} \quad D_{\hat{H}_{\mathrm{b}}} = \frac{\partial \hat{H}_{\mathrm{b}}}{\partial \hat{p}_i}\frac{\partial}{\partial q^i}$$

with $D_{\hat{H}} = D_{\hat{H}_{\mathrm{a}}} + D_{\hat{H}_{\mathrm{b}}}$. Notice that, taken separately, each vector field can be integrated trivially and exactly. For $(q_\alpha, \hat{p}_\alpha) \in T^*\mathcal{M}$ (we use the sub-index α to refer to an element in the sequence, not a particular coordinate) we have

$$\begin{pmatrix}\dot{q}\\ \dot{\hat{p}}\end{pmatrix} = \begin{pmatrix}0\\ -\nabla_q \hat{H}_{\mathrm{a}}(q)\end{pmatrix} \quad \Longrightarrow \quad \begin{pmatrix}q_{\alpha+1}\\ \hat{p}_{\alpha+1}\end{pmatrix} = \begin{pmatrix}q_\alpha\\ \hat{p}_\alpha - h\cdot\nabla_q \hat{H}_{\mathrm{a}}(q_\alpha)\end{pmatrix} = \hat{\phi}_{\mathrm{a},h}\begin{pmatrix}q_\alpha\\ \hat{p}_\alpha\end{pmatrix}$$

and

$$\begin{pmatrix}\dot{q}\\ \dot{\hat{p}}\end{pmatrix} = \begin{pmatrix}+\nabla_{\hat{p}} \hat{H}_{\mathrm{b}}(\hat{p})\\ 0\end{pmatrix} \quad \Longrightarrow \quad \begin{pmatrix}q_{\alpha+1}\\ \hat{p}_{\alpha+1}\end{pmatrix} = \begin{pmatrix}q_\alpha + h\cdot\nabla_{\hat{p}} \hat{H}_{\mathrm{b}}(p_\alpha)\\ p_\alpha\end{pmatrix} = \hat{\phi}_{\mathrm{b},h}\begin{pmatrix}q_\alpha\\ \hat{p}_\alpha\end{pmatrix} .$$

A first-order symplectic method can be easily constructed by performing the composition

$$\hat{\psi}_h = \hat{\phi}_{\mathrm{b},h} \circ \hat{\phi}_{\mathrm{a},h} . \tag{4.4}$$

Indeed, the maps $\hat{\phi}_{\mathrm{b},h}$ and $\hat{\phi}_{\mathrm{a},h}$ are symplectic because they are the exact flows of Hamiltonian vector fields. Since the composition of two symplectic maps is again symplectic, $\hat{\psi}_h$ is symplectic.

Many simple Hamiltonian systems with phase space $T^*\mathcal{M}$ are *not* governed by separable Hamiltonians, so splitting methods cannot be applied directly. However, the Hamiltonian function of many mechanical systems becomes separable if the phase space is embedded in $T^*\mathbb{R}^n$ (see, e.g., Remark 3). An interesting symplectic method that is particularly well suited for this class of systems was proposed in [42]. Roughly speaking, the idea is to compute a symplectic method for the *unconstrained* Hamiltonian vector field $D_H : T^*\mathbb{R}^n \to T(T^*\mathbb{R}^n)$, i.e., a symplectic map $\psi_{H,h}$ approximating the solution of the ODE (notice the absence of the constraint equations)

$$\begin{aligned}\dot{r} &= +\nabla_p H(r,p)\\ \dot{p} &= -\nabla_r H(r,p)\end{aligned}$$

at time $t = h$.

If H is separable, $\psi_{H,h}$ can be readily found. The method $\Psi_{H,g,h}$ for the original constrained Hamiltonian vector field $X_{H,g}$ is then constructed by taking the image of $\psi_{H,h}$ and applying a correction term that ensures that the value of $\Psi_{H,g,h}$ belongs to $\mathcal{L}\mathbb{M}$, so that the constraints are satisfied. The correction is done in a careful way so that the resulting map is still symplectic (see also [24]). Depending on the accuracy of $\psi_{H,h}$, the resulting $\Psi_{H,g,h}$ can be of first or second order (see § 4.5 for details).

4.3. Symmetry. For each t for which the solution is defined, the flow $\phi_t(x_0)$ of an autonomous differential equation defines a transformation on the phase space. It follows from the group property of the flow [3] that the inverse of the transformation

can be obtained simply by reversing time, that is, $\phi_t^{-1}(x_1) = \phi_{-t}(x_1) = x_0$. Needless to say, this property does not hold in general for a discrete approximation ψ_h, which motivates the following definition.

DEFINITION 4.3. *The* adjoint method ψ_h^* *of a method* ψ_h *is the inverse map of the original method with reversed time step* $-h$*, i.e.,*

$$\psi_h^* := (\psi_{-h})^{-1} .$$

In other words, $x_1 = \psi_h^*(x_0)$ *is implicitly defined by* $\psi_{-h}(x_1) = x_0$*. A method for which* $\psi_h^* = \psi_h$ *is called* symmetric*.*

From a theoretical point of view, an approximate discrete-time flow should be symmetric because actual continuous flows are. But symmetry is important from a practical point of view too. It has been proved in [54] that all symmetric methods are of even order, a fact that can be exploited to construct high-order methods from simple lower-order methods. For example, one can take a first-order non-symmetric method, compute its adjoint and construct a symmetric method

$$\Psi_h = \psi_{\frac{h}{2}} \circ \psi^*_{\frac{h}{2}} . \tag{4.5}$$

We know that Ψ_h is at least first order, but since Ψ_h is symmetric, we also know that the order has to be even, so we conclude that the method is actually of second order.

The scheme (4.5) works particularly well for splitting methods. Take, e.g., the integration scheme (4.4). The maps $\hat{\phi}_{\mathrm{b},h}$ and $\hat{\phi}_{\mathrm{a},h}$ are symmetric (because they are exact solutions of a differential equation), but their composition is not symmetric in general. To remedy this, one can compute the adjoint method

$$\hat{\psi}_h^* = \left(\hat{\phi}_{\mathrm{b},-h} \circ \hat{\phi}_{\mathrm{a},-h}\right)^{-1} = \hat{\phi}^{-1}_{\mathrm{a},-h} \circ \hat{\phi}^{-1}_{\mathrm{b},-h} = \hat{\phi}_{\mathrm{a},h} \circ \hat{\phi}_{\mathrm{b},h} \tag{4.6}$$

and, using (4.4), (4.6) and (4.5), construct

$$\hat{\Psi}_h = \hat{\phi}_{\mathrm{b},\frac{h}{2}} \circ \hat{\phi}_{\mathrm{a},\frac{h}{2}} \circ \hat{\phi}_{\mathrm{a},\frac{h}{2}} \circ \hat{\phi}_{\mathrm{b},\frac{h}{2}} = \hat{\phi}_{\mathrm{b},\frac{h}{2}} \circ \hat{\phi}_{\mathrm{a},h} \circ \hat{\phi}_{\mathrm{b},\frac{h}{2}} ,$$

which is a second-order symmetric method.

4.4. Modified vector fields and exponential representations. Consider a vector-valued function F and a vector field X, both on $\mathcal{LM}$. If F and X are analytic, then the composition of F and the generated flow $\phi_t(x_0)$ can be expanded in a Taylor series around $t = 0$,

$$F \circ \phi_t(x_0) = \exp(tX)F(x_0) := \sum_{i=0}^{\infty} \frac{t^i}{i!} X^i(F)(x_0) ,$$

where $X^0(F) = F$, $X^2(F) = X(X(F))$, $X^3(F) = X(X^2(F))$ etc (see [34, 53] for details). In particular, if F is taken as the identity function Id, one obtains the flow

$$\phi_t(x_0) = \exp(tX)\,\mathrm{Id}(x_0)$$

Since an s-order method ψ_h for X coincides with the flow of a modified vector field $\tilde{X} = X + \mathcal{O}(h^s)$ [16, p. 340], it is also possible to expand ψ_h in a Taylor series

$$\psi_h(x_0) = \exp(h\tilde{X})\,\mathrm{Id}(x_0) .$$

This exponential notation is a convenient way to express the relationship between a vector field and the flow generated by it —and as it will be shown in the next section, to analyze the composition of flows.

4.5. A splitting method for implicit port-Hamiltonian systems. In this section we extend the results of [42] to the PH case. We show that, with a straightforward modification, the method presented in [42] (originally intended as an integration scheme for pure Hamiltonian systems) can be used to compute sampled-data models that preserve the main properties of a PH system.

Consider again the implicit vector field

$$X_{H,u,g} = D_H + \left(u_l U_i{}^l - \lambda_j \frac{\partial g^j}{\partial r^i} \right) \frac{\partial}{\partial p_i} , \quad g = 0 , \quad y^l = U_i{}^l \frac{\partial H}{\partial p_i} , \tag{4.7}$$

defined on $\mathcal{L}\mathbb{M}$, with D_H as in (2.20) and with piece-wise constant controls (4.1). Suppose that a method $\psi_{H,u,h} : T^*\mathbb{R}^n \to T^*\mathbb{R}^n$ of order $s \geq 1$ for the unconstrained vector field

$$X_{H,u} = D_H + u_l U_i{}^l \frac{\partial}{\partial p_i}$$

has been computed. Again, in many cases H is separable so a high-order and symmetric method with a PH modified vector field can be easily found. The controls u_l are constant during each sampling interval, which also simplifies the task of finding $\psi_{H,u,h}$.

Let us define the map

$$\Phi_{\Lambda,h} \begin{pmatrix} r_\alpha \\ p_\alpha \end{pmatrix} := \begin{pmatrix} r_\alpha \\ p_\alpha - hG(r_\alpha)^\top \Lambda \end{pmatrix} . \tag{4.8}$$

Loosely, this is an approximation of the integral curve of the remnant vector field $-\lambda_j \frac{\partial g^j}{\partial r^i} \frac{\partial}{\partial p_i}$, evaluated at $t = h$ and subject to $g = 0$. More precisely, for arbitrary functions λ_j of r and p, we have that

$$D_{\lambda_j g^j} := \frac{\partial(\lambda_j g^j)}{\partial p_i} \frac{\partial}{\partial r^i} - \frac{\partial(\lambda_j g^j)}{\partial r_i} \frac{\partial}{\partial p^i} = g^j \frac{\partial \lambda_j}{\partial p_i} \frac{\partial}{\partial r^i} - \left(\lambda_j \frac{\partial g^j}{\partial r_i} + g^j \frac{\partial \lambda_j}{\partial r_i} \right) \frac{\partial}{\partial p^i}$$

and, for $r \in \mathbb{M}$, the vector field reduces to

$$D_{\lambda_j g^j} = \lambda_j D_{g^j} = -\lambda_j \frac{\partial g^j}{\partial r^i} \frac{\partial}{\partial p_i} . \tag{4.9}$$

In other words, when restricted to $\mathbb{M}$, the vector field $-\lambda_j \frac{\partial g^j}{\partial r^i} \frac{\partial}{\partial p_i}$ is Hamiltonian (hence it generates a symplectic flow).

LEMMA 4.4. *[42] Let $g(r_\alpha) = 0$. Then, the map* (4.8) *is a first-order symplectic method for $D_{\Lambda_j g^j}$. That is, for $r_\alpha \in \mathbb{M}$,*

$$\Phi_{\Lambda,h} \begin{pmatrix} r_\alpha \\ p_\alpha \end{pmatrix} = \exp(h \tilde{D}_{\Lambda_j g^j}) \operatorname{Id} \begin{pmatrix} r_\alpha \\ p_\alpha \end{pmatrix} , \quad \tilde{D}_{\Lambda_j g^j} = D_{\tilde{\Lambda}_j g^j}$$

with $\tilde{\Lambda}$ a modified or perturbed version of Λ.

A method for (4.7) can be obtained from the symmetric composition

$$\Psi_{H,u,g,h} = \Phi_{\mu,\frac{h}{2}} \circ \psi_{H,u,h} \circ \Phi_{\nu,\frac{h}{2}} . \tag{4.10}$$

For each $(r_\alpha, p_\alpha) \in \mathcal{L}\mathbb{M}$, the values of μ and ν are determined *implicitly* by the constraints $g(r_{\alpha+1}) = 0$ and $f(r_{\alpha+1}, p_{\alpha+1}) = 0$, i.e., by the constraint $(r_{\alpha+1}, p_{\alpha+1}) \in$

$\mathcal{LM}$ (recall that f^j are the hidden constraints (2.21)). In this way, $\Psi_{H,u,g,h}$ defines a transformation on $\mathcal{LM}$.

The transformation $\Psi_{H,u,g,h}$ produces an approximate discrete flow for a given command sequence $\{u_{l\alpha}\}$. From this flow, an approximate output sequence can be obtained by evaluating the output function $y^l = U_i{}^l \frac{\partial H}{\partial p_i}$ at each discrete time αh.

THEOREM 4.5. *Consider the implicit method $\Psi_{H,u,g,h}$ (4.10) and let $\tilde{X}_{H,u}$ be the modified vector field of $\psi_{H,u,h}$.*

1. *The method preserves the constraints $g^j = 0$, $f^j = 0$ and is of order $\bar{s} = \min(s, 2)$, where s is the order of $\psi_{H,u,h}$.*

2. *The method is symmetric if $\psi_{H,u,h}$ is symmetric.*

3. *If $\psi_{H,u,h}$ is symplectic for $u_{l\alpha} \equiv 0$ (i.e., if $\tilde{X}_{H,u}$ is port-Hamiltonian), then the modified vector field $\tilde{X}_{H,u,g} : \mathcal{LM} \to T(\mathcal{LM})$ is port-Hamiltonian with Hamiltonian and output functions*

$$\tilde{H} = H + \mathcal{O}(h^{\bar{s}}) \quad \textit{and} \quad \tilde{y}^l = y^l + \mathcal{O}(h^{\bar{s}}) \,. \tag{4.11}$$

Proof. The method preserves the constraints by construction. The proof about the order of the method follows the same lines as the one given in [42] except that, since we are dealing with non Hamiltonian vector fields, Lie brackets have to be used instead of Poisson brackets. We will compute $\tilde{X}_{H,u,g}$, the modified vector field generating $\Psi_{H,u,g,h}$, and show that it agrees with $X_{H,u,g}$ up to the first or second order, depending on whether $\Psi_{H,u,g,h}$ is, respectively, first or second order.

Let us consider the case $s = 1$. Using the exponential notation and Lemma 4.4, the composition (4.10) takes the form

$$\Psi_{H,u,g,h} = \exp\left(\frac{h}{2} D_{\tilde{\nu}_j g^j}\right) \exp\left(h \tilde{X}_{H,u}\right) \exp\left(\frac{h}{2} D_{\tilde{\mu}_j g^j}\right) \mathrm{Id} \,,$$

where $\tilde{X}_{H,u} = X_{H,u} + \mathcal{O}(h)$. Applying the Baker-Campbell-Hausdorff (BCH) formula [53] to the product of the first two factors and truncating after the first term gives

$$\exp\left(\frac{h}{2} D_{\tilde{\nu}_j g^j}\right) \exp\left(h \tilde{X}_{H,u}\right) = \exp\left(h \tilde{X}'\right) \tag{4.12}$$

with $\tilde{X}' = \tilde{X}_{H,u} + \frac{1}{2} D_{\tilde{\nu}_j g^j} + \mathcal{O}(h)$. Applying BCH again to include the third factor gives

$$\exp\left(h \tilde{X}'\right) \exp\left(\frac{h}{2} D_{\tilde{\mu}_j g^j}\right) = \exp\left(h \tilde{X}_{H,u,g}\right) \tag{4.13}$$

with the modified vector field $\tilde{X}_{H,u,g} = \tilde{X}_{H,u} + \frac{1}{2}\left(D_{\tilde{\nu}_j g^j} + D_{\tilde{\mu}_j g^j}\right) + \mathcal{O}(h)$. Using (4.9) and $\tilde{X}_{H,u} = X_{H,u} + \mathcal{O}(h)$, we can write the modified vector field as

$$\tilde{X}_{H,u,g} = X_{H,u} + \frac{\tilde{\nu}_j + \tilde{\mu}_j}{2} D_{g^j} + \mathcal{O}(h) \,. \tag{4.14}$$

The hidden constraints $f^l = 0$ imply that

$$\tilde{X}_{H,u,g}(f^l) = X_{H,u}(f^l) + \frac{\tilde{\nu}_j + \tilde{\mu}_j}{2} D_{g^j}(f^l) + \mathcal{O}(h) = 0 \,. \tag{4.15}$$

It follows from (4.15), (4.9) and (3.5), that the Lagrange multipliers λ_j and the 'modified Lagrange multipliers' $\tilde{\nu}_j$ and $\tilde{\mu}_j$ are related by the equation

$$\frac{\tilde{\nu}_j + \tilde{\mu}_j}{2} = \lambda_j + \mathcal{O}(h) \,,$$

which when substituted back in (4.14) gives the desired result:

$$\tilde{X}_{H,u,g} = X_{H,u} + \lambda_j D_{g^j} + \mathcal{O}(h) = X_{H,u,g} + \mathcal{O}(h) \,.$$

For $s = 2$ we follow the same procedure, but we truncate the BCH formula after the second term. For the expression (4.12), the intermediate vector field is

$$\tilde{X}' = \tilde{X}_{H,u} + \frac{1}{2} D_{\tilde{\nu}_j g^j} + \frac{h}{4} \left[D_{\tilde{\nu}_j g^j}, \tilde{X}_{H,u} \right] + \mathcal{O}(h^2)$$

where $[\,\cdot\,,\,\cdot\,]$ is the standard Lie bracket. Using the initial assumption $\tilde{X}_{H,u} = X_{H,u} + \mathcal{O}(h^2)$, we can write $\tilde{X}'$ as

$$\tilde{X}' = X_{H,u} + \frac{1}{2} D_{\tilde{\nu}_j g^j} + \frac{h}{4} \left[D_{\tilde{\nu}_j g^j}, X_{H,u} \right] + \mathcal{O}(h^2) \,.$$

Regarding (4.13), the modified vector field for the complete scheme is

$$\begin{aligned} \tilde{X}_{H,u,g} = X_{H,u} &+ \frac{1}{2} \left(D_{\tilde{\nu}_j g^j} + D_{\tilde{\mu}_j g^j} \right) \\ &+ \frac{h}{4} \left[D_{\tilde{\nu}_j g^j}, X_{H,u} \right] + \frac{h}{4} \left[X_{H,u} + \frac{1}{2} D_{\tilde{\nu}_j g^j}, D_{\tilde{\mu}_j g^j} \right] + \mathcal{O}(h^2) \,. \end{aligned}$$

Using (4.9) and the skew symmetry and bilinearity of the Lie bracket, the vector field can be equivalently written as

$$\begin{aligned} \tilde{X}_{H,u,g} = X_{H,u} + \frac{\tilde{\nu}_j + \tilde{\mu}_j}{2} D_{g^j} &+ \frac{h}{4} \left[(\tilde{\nu}_j - \tilde{\mu}_j) D_{g^j}, X_{H,u} \right] \\ &+ \frac{h}{8} \left[\tilde{\nu}_j D_{g^j}, \tilde{\mu}_j D_{g^j} \right] + \mathcal{O}(h^2) \,. \end{aligned} \tag{4.16}$$

In order to extract information from the equation $\tilde{X}_{H,u,g}(g^l) = 0$, let us first open the brackets in (4.16) and write

$$\begin{aligned} \tilde{X}_{H,u,g} = X_{H,u} &+ \frac{\tilde{\nu}_j + \tilde{\mu}_j}{2} D_{g^j} + \frac{h}{4} (\tilde{\nu}_j - \tilde{\mu}_j) D_{g^j} X_{H,u} \\ &- \frac{h}{4} X_{H,u} \left((\tilde{\nu}_j - \tilde{\mu}_j) D_{g^j} \right) + \frac{h}{8} \tilde{\nu}_j D_{g^j} \left(\tilde{\mu}_j D_{g^j} \right) - \frac{h}{8} \tilde{\mu}_j D_{g^j} \left(\tilde{\nu}_j D_{g^j} \right) + \mathcal{O}(h^2) \,. \end{aligned}$$

Taking into account that $f^l = X_{H,u}(g^l) = 0$ and $D_{g^j}(g^l) \equiv 0$, we have that

$$\tilde{X}_{H,u,g}(g^l) = \frac{h}{4} (\tilde{\nu}_j - \tilde{\mu}_j) D_{g^j} f^l + \mathcal{O}(h^2) = 0 \,,$$

from which we can see that modified Lagrange multipliers satisfy the order relation

$$\tilde{\nu}_j - \tilde{\mu}_j = \mathcal{O}(h) \,. \tag{4.17}$$

By substituting (4.17) back in (4.16) we can verify that the commutators are actually second order, that is,

$$\begin{aligned}\tilde{X}_{H,u,g} &= X_{H,u} + \frac{\tilde{\nu}_j + \tilde{\mu}_j}{2} D_{g^j} + \frac{h}{8}\left[\tilde{\nu}_j D_{g^j}, (\tilde{\nu}_j + \mathcal{O}(h))\, D_{g^j}\right] + \mathcal{O}(h^2) \\ &= X_{H,u} + \frac{\tilde{\nu}_j + \tilde{\mu}_j}{2} D_{g^j} + \mathcal{O}(h^2)\ . \end{aligned} \tag{4.18}$$

From $\tilde{X}_{H,u,g}(f^l) = 0$ and (3.5) we conclude that, when $s = 2$,

$$\frac{\tilde{\nu}_j + \tilde{\mu}_j}{2} = \lambda_j + \mathcal{O}(h^2)\ ,$$

so the desired result follows:

$$\tilde{X}_{H,u,g} = X_{H,u} + \lambda_j D_{g^j} + \mathcal{O}(h^2) = X_{H,u,g} + \mathcal{O}(h^2)\ .$$

For statement (2), notice that, when restricted to $\mathcal{LM}$, the integration method (4.10) can be described by the implicit equations

$$\begin{pmatrix} r_{\alpha+1} \\ p_{\alpha+1} + hG(r_{\alpha+1})^\top \mu \end{pmatrix} = \psi_{H,u,h} \begin{pmatrix} r_\alpha \\ p_\alpha - hG(r_\alpha)^\top \nu \end{pmatrix} \tag{4.19a}$$

$$g(r_{\alpha+1}) = g(r_\alpha) \tag{4.19b}$$

$$f(r_{\alpha+1}, p_{\alpha+1}) = f(r_\alpha, p_\alpha)\ , \tag{4.19c}$$

where r_α, p_α are the independent variables and $r_{\alpha+1}, p_{\alpha+1}$ are the dependent variables. The vectors ν, μ are (also dependent) dummy variables that can be discarded after $r_{\alpha+1}, p_{\alpha+1}$ have been found.

After reversing time (that is, after substituting h by $-h$), equation (4.19a) becomes

$$\begin{pmatrix} r_{\alpha+1} \\ p_{\alpha+1} - hG(r_{\alpha+1})^\top \mu \end{pmatrix} = \psi_{H,u,-h} \begin{pmatrix} r_\alpha \\ p_\alpha + hG(r_\alpha)^\top \nu \end{pmatrix}\ .$$

Recall that $\psi_{H,u,-h} = \psi_{H,u,h}^{-1}$ if $\psi_{H,u,h}$ is symmetric. Therefore, when restricted to $\mathcal{LM}$, the reverse-time method is

$$\begin{pmatrix} r_\alpha \\ p_\alpha + hG(r_\alpha)^\top \nu \end{pmatrix} = \psi_{H,u,h} \begin{pmatrix} r_{\alpha+1} \\ p_{\alpha+1} - hG(r_\alpha + 1)^\top \mu \end{pmatrix} \tag{4.20a}$$

$$g(r_{\alpha+1}) = g(r_\alpha) \tag{4.20b}$$

$$f(r_{\alpha+1}, p_{\alpha+1}) = f(r_\alpha, p_\alpha)\ , \tag{4.20c}$$

which is the same as (4.19), but with r_α, p_α and ν interchanged with $r_{\alpha+1}, p_{\alpha+1}$ and μ, respectively. This implies that, if we input $r_{\alpha+1}, p_{\alpha+1}$ as independent variables, we recover r_α, p_α as the dependent variables, that is: $\Psi_{H,u,g,-h}$ is the inverse mapping of $\Psi_{H,u,g,h}$. (In general, the vectors ν, μ obtained using (4.19) will be different from those obtained using (4.20), but this is inconsequential since they are dummy variables.)

In statement (3), the fact that $\tilde{X}_{H,u,g}$ is PH follows directly from Lemma 4.4 and the fact that the composition of symplectic maps is again symplectic. In other words, $\Psi_{H,u,g,h}$ is symplectic when $u_{l\alpha} \equiv 0$, so according to Lemma 3.4,

$$\tilde{X}_{H,u,g} = X_0 + u_l Z^l = D_{\tilde{H}} + u_l Z^l + \tilde{\lambda}^j D_{g^j} \tag{4.21}$$

for some Hamiltonian function $\tilde{H}$ and some input vector fields Z^l. Since the method is of order $\bar{s}$, we have

$$\tilde{X}_{H,u,g} = D_H + u_l {U_i}^l \frac{\partial H}{\partial p_i} + \lambda^j D_{g^j} + \mathcal{O}(h^{\bar{s}}) \ . \tag{4.22}$$

By setting $u_l = 0$ and recalling that $\tilde{\lambda}_j = \lambda_j + \mathcal{O}(h^{\bar{s}})$, it follows from (4.21) and (4.22) that $D_{\tilde{H}} = D_H + \mathcal{O}(h^{\bar{s}})$, which implies that $\tilde{H} = H + \mathcal{O}(h^{\bar{s}})$, and, in turn, that

$$Z^l = {U_i}^l \frac{\partial}{\partial p_i} + \mathcal{O}(h^{\bar{s}}) \ .$$

The modified output functions are thus

$$\tilde{y}^l = Z^l(\tilde{H}) = {U_i}^l \frac{\partial}{\partial p_i}\big(H + \mathcal{O}(h^{\bar{s}})\big) + \mathcal{O}(h^{\bar{s}}) = y^l + \mathcal{O}(h^{\bar{s}}) \ .$$

□

4.6. Energy balance under sample and hold. The power balance (3.7) implies that

$$H_{\alpha+1} - H_\alpha = \int_{\alpha h}^{\alpha h + h} u_l(t) y^l(t) \mathrm{d}t \ , \tag{4.23}$$

where we have defined the sampled Hamiltonian $H_\alpha := H(x_\alpha)$. A usual way to improve the transient behavior of the system is to add damping by means of a continuous control law [37]

$$u_l(t) = -K_{lj} y^j(t) \ , \tag{4.24}$$

with $\left\{K_{lj}\right\}$ a symmetric and positive semi-definite matrix. With the control law (4.24), the power balance (4.23) results in the dissipation inequality

$$H_{\alpha+1} - H_\alpha \leq 0 \ ,$$

which guarantees that H_α decreases monotonically and, if the right conditions are met, the system converges to a state of minimal energy.

Suppose that the output is being sampled and that the input is being held at intervals of length h. The control sequence is then given by

$$u_{l\alpha} = -K_{lj} {y^j}_\alpha \tag{4.25}$$

and the power balance (4.23) takes the form

$$H_{\alpha+1} - H_\alpha = u_{l\alpha} \int_{\alpha h}^{\alpha h + h} y^l(t) \mathrm{d}t = u_{l\alpha} \int_0^h y^l(\alpha h + \tau) \mathrm{d}\tau \ .$$

Applying Taylor's theorem to the integral term gives

$$H_{\alpha+1} - H_\alpha = \sum_l u_{l\alpha} \left({y^l}_\alpha h + \mathcal{O}(h^2)\right) = -h K_{lj} {y^l}_\alpha {y^j}_\alpha + \sum_l u_{l\alpha} \mathcal{O}(h^2) \ ,$$

so H_α decreases when h is small enough and the norm of y_α is large enough.

Since the approximate sampled-data model (4.10) is also PH (cf. item (3) of Theorem 4.5)), it satisfies (again, after applying Taylor's theorem)

$$\tilde{H}_{\alpha+1} - \tilde{H}_\alpha = \sum_l u_{l\alpha} \left(\tilde{y}^l{}_\alpha h + \mathcal{O}(h^2) \right) \tag{4.26}$$

for some $\tilde{H}$ and $\tilde{y}$. According to (4.11), the energy balance (4.26) takes the form

$$H_{\alpha+1} - H_\alpha + \mathcal{O}(h^{\bar{s}}) = \sum_l u_{l\alpha} \left(\left(y^l{}_\alpha + \mathcal{O}(h^{\bar{s}}) \right) h + \mathcal{O}(h^2) \right) \ .$$

The same control sequence (4.25) produces

$$H_{\alpha+1} - H_\alpha = -hK_{lj} y^l{}_\alpha y^j{}_\alpha + \sum_l u_{l\alpha} \mathcal{O}(h^2) + \mathcal{O}(h^{\bar{s}}) \ .$$

Thus, for $\bar{s} = 2$, the qualitative behavior of the approximated sampled data model is the same as the exact one: H_α decreases when h is small enough and the norm of y_α is large enough.

Example: A double planar pendulum (final part). Let us compute a sampled-data model for the double pendulum described in the previous examples. The first step is to compute a sample-data model for the simple unconstrained PH system

$$X_{H,u} = D_H + u_l U_i{}^l \frac{\partial}{\partial p_i} \ ,$$

where H is given by (2.29) and $U_i{}^l$ by (3.8) and (3.9). The unconstrained and unactuated Hamiltonian vector field D_H describes a pair of masses with initial positions $r^a{}_0$ and $r^b{}_0$ and initial momenta p_{a0} p_{b0}, simply falling under the influence of gravity. The exact flow generated by D_H, denoted by $(r_{\alpha+1}, p_{\alpha+1}) = \phi_{H,h}(r_\alpha, p_\alpha)$, is then given by

$$\begin{aligned}
r^{a_x}{}_{\alpha+1} &= r^{a_x}{}_\alpha + \frac{h}{m_a} p_{a_x\alpha} \\
r^{a_y}{}_{\alpha+1} &= r^{a_y}{}_\alpha + \frac{h}{m_a} p_{a_y\alpha} - \bar{g}\frac{h^2}{2} \\
r^{b_x}{}_{\alpha+1} &= r^{b_x}{}_\alpha + \frac{h}{m_b} p_{b_x\alpha} \\
r^{b_y}{}_{\alpha+1} &= r^{b_y}{}_\alpha + \frac{h}{m_b} p_{b_y\alpha} - \bar{g}\frac{h^2}{2}
\end{aligned}$$

and

$$\begin{aligned}
p_{a_x\alpha+1} &= p_{a_x\alpha} \\
p_{a_y\alpha+1} &= p_{a_y\alpha} - m_a \bar{g} h \\
p_{b_x\alpha+1} &= p_{b_x\alpha} \\
p_{b_y\alpha+1} &= p_{b_y\alpha} - m_b \bar{g} h \ .
\end{aligned}$$

The exact flow generated by $u_l U_i{}^l \frac{\partial}{\partial p_i}$ (the control vector field without drift) is denoted by $(r_{\alpha+1}, p_{\alpha+1}) = \phi_{u,h}(r_\alpha, p_\alpha)$. It is given by

$$r^i{}_{\alpha+1} = r^i{}_\alpha \ , \quad i \in \{a_x, a_y, b_x, b_y\}$$

Value	Description
$l_a = 0.6$[m]	Length of the first link
$l_b = 0.3$[m]	Length of the second link
$m_a = 0.2$[kg]	Value of the first mass
$m_b = 0.6$[kg]	Value of the second mass
$\bar{g} = 9.81$[m/s]^2	Acceleration due to gravity

TABLE 4.1
Parameters for the double pendulum.

and

$$
\begin{aligned}
p_{a_x\alpha+1} &= p_{a_x\alpha} - \frac{h}{l_a^2}(u_1 - u_2) r^{a_y}{}_\alpha \\
p_{a_y\alpha+1} &= p_{a_y\alpha} + \frac{h}{l_a^2}(u_1 - u_2) r^{a_x}{}_\alpha \\
p_{b_x\alpha+1} &= p_{b_x\alpha} - \frac{h}{l_b^2} u_2 r^{\delta_y}{}_\alpha \\
p_{b_y\alpha+1} &= p_{b_y\alpha} + \frac{h}{l_b^2} u_2 r^{\delta_x}{}_\alpha \ .
\end{aligned}
$$

From § 4.2, we know that a simple symmetric method of order two for $X_{H,u}$ is

$$
\psi_{H,u,h} = \phi_{H,\frac{h}{2}} \circ \phi_{u,h} \circ \phi_{H,\frac{h}{2}} \ . \tag{4.27}
$$

Notice that $\phi_{u,h} = \mathrm{Id}$ when $u \equiv 0$, so $\psi_{H,u,h} = \phi_{H,\frac{h}{2}} \circ \phi_{H,\frac{h}{2}} = \phi_{H,h}$, which is a symplectic map because it is the exact solution of a Hamiltonian system. Therefore, $\tilde{X}_{H,u}$ is PH according to Definition 3.3. From Theorem 4.5 we know that the implicit method (4.10), with $\Psi_{H,u,h}$ as in (4.27) is the exact solution of a PH system $\tilde{X}_{H,u,g}$ with Hamiltonian function $\tilde{H} = H + \mathcal{O}(h^2)$ and output function $\tilde{y} = y + \mathcal{O}(h^2)$ (i.e., $\bar{s} = s = 2$).

The sampled-data model was tested using the parameters shown in Table 4.1. For illustration purposes, we chose a damping control $u_\alpha = -0.3 \cdot y_\alpha$ and simulated the closed-loop system using the sampled-data model (4.10). Figure 4.2 shows the discrete-time series of H for different values of h. It can be seen that the time series converge and, as expected, the value of H_α decreases monotonically when h is small enough (in this case, less or equal to 30 [ms]). For comparison purposes, we have included the evolution of H that is obtained by simulating (with Matlab's Simulink) the explicit model (2.37) in series with a sampler and a zero-order hold.

5. Conclusions. Assumption 1 ensures that the configuration space of a port-Hamiltonian system can be embedded in Euclidean space, which in turn allows to establish the commutative diagram shown in Fig. 2.1 and, with the functions depicted in Table 3.1, relate the implicit and explicit representations of such system. On a practical ground, this can be useful to compare different integration methods or, e.g., to design a controller using one representation and to implement it using the other one.

Using the notion of symplecticity in the absence of a control action, the definition of a port-Hamiltonian was extended to cover the case where the controls do not enter the system's equation affinely (Definition 3.3).

When an implicit representation is available, a sampled-data model of order two can be easily constructed. Such model preserves the Hamiltonian structure (hence

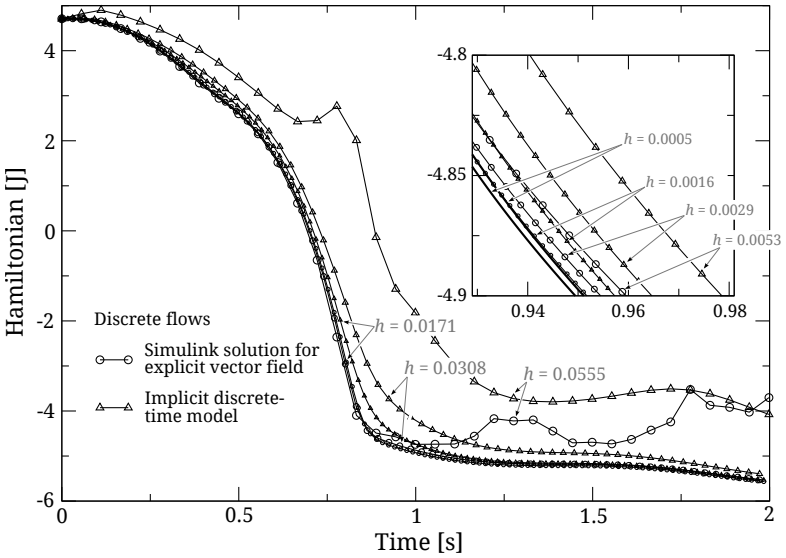

FIG. 4.2. *Results of the numerical experiment. The Hamiltonian function is plotted against time. The explicit model (simulated using Matlab's module Simulink) is compared with the implicit model (simulated using a Matlab script). As expected, H is monotonically decreasing when h is small enough and the time series converges as h goes to zero.*

passivity, or losslessness to be precise) (Theorem 4.5). This provides a starting point for discrete-time passivity-based control.

REFERENCES

[1] ANDREI A. AGRACHEV AND YURI L. SACHKOV, *Control Theory from the Geometric Viewpoint*, Springer-Verlag, Berlin, Germany, 2004.

[2] VLADIMIR I. ARNOLD, *Mathematical methods of classical mechanics*, Springer-Verlag, New York, 1989.

[3] ———, *Ordinary Differential Equations*, Springer-Verlag, Berlin, 1992.

[4] VLADIMIR I. ARNOLD, VALERY KOZLOV, AND ANATOLY I. NEISHTADT, *Mathematical aspects of classical and celestial mechanics*, Springer-Verlag, 2006.

[5] K. E. BRENAN, S. L. CAMPBELL, AND L. R. PETZOLD, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, North-Holland, New York, 1989.

[6] CHRISTOPHER I. BYRNES, ALBERTO ISIDORI, AND JAN. C. WILLEMS, *Passivity, feedback equivalence, and the global stabilization of minimium phase nonlinear systems*, IEEE Trans. Autom. Control, 36 (1991), pp. 1228–1240.

[7] CHRISTOPHER I. BYRNES AND WEI LIN, *Losslessness, feedback equivalence, and the global stabilization of discrete-time nonlinear systems*, IEEE Trans. Autom. Control, 39 (1994), pp. 83 – 98.

[8] P. J. CHANNELL AND F. R. NERI, *Symplectic integrators*, in Integration Algorithms and Classical Mechanics, Jerrold E. Marsden, George W. Patrick, and William F. Shadwick, eds., AMS, Rhode Island, 1996, pp. 45 – 57.

[9] RAMON COSTA-CASTELLÓ AND ENRIC FOSSAS, *On preserving passivity in sampled-data linear systems*, in Proc. of the American Control Conference, Minneapolis, Minnesota, June 2006, pp. 4373 – 4378.

[10] MORTEN DALSMO AND ARJAN J. VAN DER SCHAFT, *On representations and integrability of mathematical structures in energy-conserving physical systems*, SIAM J. Control Optim., 37 (1999), pp. 54 – 91.

[11] ZHONG GE AND JERROLD E. MARSDEN, *Lie-Poisson Hamilton-Jacobi theory and Lie-Poisson integrators*, Physical Review Letters A, 133 (1988), pp. 134 – 139.

[12] MARIANO GIAQUINTA AND STEFAN HILDEBRANDT, *Calculus of Variations II, The Hamiltonian*

Formalism, Springer-Verlag, Berlin, Germany, 1996.

[13] Herbert Goldstein, Charles P. Poole, and John L. Safko, *Classical Mechanics*, Addison Wesley, 2002.

[14] Oscar González, *Design and analysis of conserving integrators for nonlinear Hamiltonian systems with symmetry*, PhD thesis, Stanford University, 1996.

[15] Oscar González, *Mechanical systems subject to holonomic constraints: Differentialalgebraic formulations and conservative integration*, Physica D, 132 (1999), pp. 165 – 174.

[16] Ernst Hairer, Christian Lubich, and Gerhard Wanner, *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*, Springer-Verlag, 2006.

[17] David J. Hill and Peter Moylan, *The stability of nonlinear dissipative systems*, IEEE Trans. Autom. Control, (1976), pp. 708–711.

[18] L. Hitz and Brian D. O. Anderson, *Discrete positive-real functions and their application to system stability*, Proc. IEE, 116 (1969), pp. 153 – 155.

[19] Robert A. LaBudde and Donald Greenspan, *Energy and momentum conserving methods of arbitrary order for the numerical integration of equations of motion. I. motion of a single particle*, Numer. Math., 25 (1976), pp. 323 – 346.

[20] ———, *Energy and momentum conserving methods of arbitrary order for the numerical integration of equations of motion. II. motion of a system of particles*, Numer. Math., 26 (1976), pp. 1 – 16.

[21] Dina Shona Laila and Alessandro Astolfi, *On the construction of discrete-time model for port-controlled Hamiltonian systems with applications*, Systems and Control Letters, 55 (2006), pp. 673 – 680.

[22] ———, *Direct discrete-time design for sampled-data Hamiltonian control systems*, in Lagrangian and Hamiltonian Methods for Nonlinear Control 2006, Francesco Bullo and Kenji Fujimoto, eds., Springer, 2007, pp. 87 – 98.

[23] John M. Lee, *Introduction to Smooth Manifolds*, Springer-Verlag, New York, 2003.

[24] Benedict Leimkuhler and Sebastian Reich, *Simulating Hamiltonian Dynamics*, Cambrige University Press, Cambridge, UK, 2004.

[25] Wei Lin, *Feedback stabilization of general nonlinear control systems: A passive system approach*, Systems and Control Letters, 25 (1995), pp. 41 – 52.

[26] Wei Lin and Christopher I. Byrnes, *Passivity and absolute stabilization of a class of discrete-time nonlinear systems*, Automatica, 31 (1995), pp. 263 – 267.

[27] Jerrold E. Marsden and Tudor S. Ratiu, *Introduction to Mechanics and Symmetry*, Springer-Verlag, New York, 1999.

[28] Bernhard Maschke, Arjan J. van der Schaft, and Peter C. Breedveld, *An intrinsic Hamiltonian formulation of network dynamics: Non-standard Poisson structures and gyrators*, Journal of the Franklin Institute, 329 (1992), pp. 923 – 966.

[29] ———, *An intrinsic Hamiltonian formulation of the dynamics of LC-circuits*, IEEE Trans. Circuits Syst. I, 42 (1995), pp. 73 – 82.

[30] S. Monaco and Dorothée Normand-Cyrot, *Advanced tools for nonlinear sampled-data systems' analysis and control*, European Journal of Control, 13 (2007), pp. 221 – 241.

[31] S. Monaco, Dorothée Normand-Cyrot, and Fernando Tiefensee, *From passivity under sampling to a new discrete-time passivity concept*, in Proc. Conference on Decision and Control, Cancun, Mexico, Dec. 2008, pp. 3157 – 3162.

[32] ———, *Nonlinear port controlled Hamiltonian systems under sampling*, in Proc. Conference on Decision and Control, and Chinese Control Conference, Shanghai, P.R. China, Dec. 2009, pp. 1782 – 1787.

[33] Dragan Nešić and Andrew Teel, *A framework for stabilization of nonlinear sampled-data systems based on their approximate discrete-time models*, IEEE Trans. Autom. Control, 49 (2004), pp. 1103 – 1122.

[34] Peter J. Olver, *Applications of Lie groups to differential equations*, Springer-Verlag, New York, 1993.

[35] Romeo Ortega, Antonio Loría, J. P. Nicklasson, and Hebert Sira-Ramirez, *Passivity-based Control of Euler-Lagrange Systems*, Springer-Verlag, Berlin, 1998.

[36] Romeo Ortega, Arjan J. van der Schaft, Fernando Castaños, and Alessandro Astolfi, *Control by interconnection and standard passivity-based control of port-Hamiltonian systems*, IEEE Trans. Autom. Control, 53 (2008), pp. 2527 – 2542.

[37] Romeo Ortega, Arjan J. van der Schaft, Iven Mareels, and Bernhard Maschke, *Putting energy back in control*, IEEE Control Syst. Mag., (2001), pp. 18–33.

[38] Romeo Ortega, Arjan J. van der Schaft, Bernhard Maschke, and Gerardo Escobar, *Interconnection and damping assignment passivity-based control of port-controlled Hamiltonian systems*, Automatica, 38 (2002), pp. 585 – 596.

[39] Jan Willem Polderman and Jan. C. Willems, *Introduction to Mathematical Systems Theory: A Behavioral Approach*, Springer-Verlag, 1998.
[40] Sebastian Reich, *On a geometrical interpretation of differential-algebraic equations*, Circuits Syst. Signal Process., 9 (1990), pp. 367 – 382.
[41] ———, *On an existence and uniqueness theory for nonlinear differential-algebraic equations*, Circuits Syst. Signal Process., 10 (1991), pp. 343 – 359.
[42] ———, *Symplectic integration of constrained Hamiltonian systems by composition methods*, SIAM J. Numer. Anal., 33 (1996), pp. 475 – 491.
[43] Werner C. Rheinboldt, *Differential-algebraic systems as differential equations on manifolds*, Mathematics of computation, 43 (1984), pp. 473–482.
[44] ———, *On the existence and uniqueness of solutions of nonlinear semi-implicit differential-algebraic equations*, Nonlinear Analysis. Theory, Methods and Applications, 16 (1991), pp. 647 – 661.
[45] R. Tyrrel Rockafellar and Roger J.-B. Wets, *Variational Analysis*, Springer-Verlag, 2009.
[46] Bernard Schutz, *Geometrical methods of mathematical physics*, Cambrige University Press, New York, 1980.
[47] J. C. Simo, N. Tarnow, and K. K. Wong, *Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics*, Int. J. Numerical Methods in Engineering, 100 (1992), pp. 63 – 116.
[48] J. C. Simo and K. K. Wong, *Unconditionally stable algorithms for rigid body dynamics that exactly preserve energy and momentum*, Int. J. Numerical Methods in Engineering, 31 (1991), pp. 19 – 52.
[49] S. Smale, *On the mathematical foundations of electrical circuit theory*, J. Differential Geometry, 7 (1972), pp. 193 – 210.
[50] Stephano Stramigioli, Cristian Secchi, Arjan J. van der Schaft, and Cesare Fantuzzi, *Sampled data systems passivity and discrete port-Hamiltonian systems*, IEEE Trans. Robot., 21 (2005), pp. 574 – 587.
[51] Morikazu Takegaki and Suguru Arimoto, *A new feedback method for dynamic control of manipulators*, ASME Journal of Dynamic Systems, Measurement, and Control, 102 (1981), pp. 119 – 125.
[52] Arjan J. van der Schaft, *$\mathcal{L}_2$-Gain and Passivity Techniques in Nonlinear Control*, Springer-Verlag, London, 2000.
[53] Veeravalli S. Varadarajan, *Lie Groups, Lie Algebras, and Their Representations*, Springer-Verlag, New York, 1984.
[54] Haruo Yoshida, *Construction of higher order symplectic integrators*, Physics Letters A, 150 (1990), pp. 262 – 268.
[55] D. C. Youla, L. J. Castriota, and H. J. Carlin, *Bounded real scattering matrices and the foundations of linear passive network theory*, IRE Trans. Circuit Theory, (1959), pp. 102 – 124.
[56] R.K.P. Zia, Edward F. Redish, and Susan R. McKay, *Making sense of the Legendre transform*, Am. J. Phys., 7 (2009), pp. 614 – 622.