© 2011 Society for Industrial and Applied Mathematics

# DISCRETE HAMILTON–JACOBI THEORY*

TOMOKI OHSAWA†, ANTHONY M. BLOCH‡, AND MELVIN LEOK†

**Abstract.** We develop a discrete analogue of Hamilton–Jacobi theory in the framework of discrete Hamiltonian mechanics. The resulting discrete Hamilton–Jacobi equation is discrete only in time. We describe a discrete analogue of Jacobi's solution and also prove a discrete version of the geometric Hamilton–Jacobi theorem. The theory applied to discrete linear Hamiltonian systems yields the discrete Riccati equation as a special case of the discrete Hamilton–Jacobi equation. We also apply the theory to discrete optimal control problems, and recover some well-known results, such as the Bellman equation (discrete-time HJB equation) of dynamic programming and its relation to the costate variable in the Pontryagin maximum principle. This relationship between the discrete Hamilton–Jacobi equation and Bellman equation is exploited to derive a generalized form of the Bellman equation that has controls at internal stages.

**Key words.** Hamilton–Jacobi equation, Bellman equation, dynamic programming, discrete mechanics, discrete-time optimal control

**AMS subject classifications.** 70H20, 49L20, 93C55, 49J15, 70H05, 70H25

**DOI.** 10.1137/090776822

## 1. Introduction.

**1.1. Discrete mechanics.** Discrete mechanics is a reformulation of Lagrangian and Hamiltonian mechanics with discrete time, as opposed to a discretization of the equations in the continuous-time theory. It not only provides a systematic view of structure-preserving integrators but also has interesting theoretical aspects analogous to continuous-time Lagrangian and Hamiltonian mechanics (see, e.g., Marsden and West [30] and Suris [34, 35]). The main feature of discrete mechanics is its use of discrete versions of variational principles. Namely, discrete mechanics assumes that the dynamics is defined at discrete times from the outset, formulates a discrete variational principle for such dynamics, and then derives a discrete analogue of the Euler–Lagrange or Hamilton's equations from it.

The advantage of this construction is that it naturally gives rise to discrete analogues of the concepts and ideas in continuous time that have the same or similar properties, such as symplectic forms, the Legendre transformation, momentum maps, and Noether's theorem [30]. This in turn provides us with the discrete ingredients that facilitate further theoretical developments, such as discrete analogues of the theories of complete integrability (see, e.g., Moser and Veselov [31] and Suris [34, 35]) and also those of reduction and connections [20, 26, 25]. Whereas the main topic in discrete mechanics is the development of structure-preserving algorithms for Lagrangian and Hamiltonian systems (see, e.g., [30]), the theoretical aspects of it are interesting in their own right, and furthermore provide insight into the numerical aspects as well.

Another notable feature of discrete mechanics, especially on the Hamiltonian side, is that it is a generalization of (regular) discrete optimal control problems. In fact,

---

†Department of Mathematics, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093–0112 (tohsawa@ucsd.edu, mleok@math.ucsd.edu).
‡Department of Mathematics, University of Michigan, 530 Church Street, Ann Arbor, Michigan 48109–1043 (abloch@umich.edu).

as stated in [30], discrete mechanics is inspired by discrete formulations of optimal control problems (see, e.g., Jordan and Polak [21] and Cadzow [9]).

**1.2. Hamilton–Jacobi theory.** In classical mechanics (see, e.g., Lanczos [24], Arnold [3], Marsden and Ratiu [29], and Goldstein, Poole, and Safko [16]), the Hamilton–Jacobi equation is first introduced as a partial differential equation that the action integral satisfies. Specifically, let $Q$ be a configuration space and $T^*Q$ be its cotangent bundle; let $q \in Q$ and $t > 0$ be arbitrary and suppose that $(\hat{q}(s), \hat{p}(s)) \in T^*Q$ is a solution of Hamilton's equations

$$(1.1) \qquad \dot{q} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial q}$$

with the endpoint condition $\hat{q}(t) = q$. Then, calculate the action integral along the solution over the time interval $[0, t]$, i.e.,

$$(1.2) \qquad S(q, t) := \int_0^t \left[ \hat{p}(s) \cdot \dot{\hat{q}}(s) - H(\hat{q}(s), \hat{p}(s)) \right] ds,$$

where we regard the resulting integral as a function of the endpoint $(q, t) \in Q \times \mathbb{R}_+$ with $\mathbb{R}_+$ being the set of positive real numbers. Then, by taking variation of the endpoint $(q, t)$, one obtains a partial differential equation satisfied by $S(q, t)$:

$$(1.3) \qquad \frac{\partial S}{\partial t} + H\left( q, \frac{\partial S}{\partial q} \right) = 0.$$

This is the *Hamilton–Jacobi (H-J) equation.*

Conversely, it is shown that if $S(q, t)$ is a solution of the Hamilton–Jacobi equation, then $S(q, t)$ is a generating function for the family of canonical transformations (or symplectic flows) that describe the dynamics defined by Hamilton's equations. This result is the theoretical basis for the powerful technique of exact integration called separation of variables.

**1.3. Connection with optimal control and the Hamilton–Jacobi–Bellman equation.** The idea of Hamilton–Jacobi theory is also useful in optimal control theory (see, e.g., Jurdjevic [22, Chapter 13] and Bertsekas [6, Section 3.2]). Consider a typical optimal control problem

$$\min_{u(\cdot)} \int_0^T C(q, u) \, dt,$$

subject to the constraints,

$$\dot{q} = f(q, u),$$

and $q(0) = q_0$ and $q(T) = q_T$. We define the augmented cost functional:

$$\hat{S}[u] := \int_0^T \{ C(q, u) + p[\dot{q} - f(q, u)] \} \, dt = \int_0^T \left[ p \cdot \dot{q} - \hat{H}(q, p, u) \right] dt,$$

where we introduced the costate $p$, and also defined the control Hamiltonian,

$$\hat{H}(q, p, u) := p \cdot f(q, u) - C(q, u).$$

Assuming that

$$\frac{\partial \hat{H}}{\partial u}(q, p, u) = 0$$

uniquely defines the optimal control $u = u^*(q, p)$, we set

$$H(q, p) := \max_u \hat{H}(q, p, u) = \hat{H}(q, p, u^*(q, p)).$$

We also define the optimal cost-to-go function

$$J(q, t) := \int_t^T \left\{ C(\hat{q}, u^*) + p\left[\dot{\hat{q}} - f(\hat{q}, u^*)\right] \right\} ds$$

$$= \int_t^T \left[ \hat{p} \cdot \dot{\hat{q}} - H(\hat{q}, \hat{p}) \right] ds = S^* - S(q, t),$$

where $(\hat{q}(s), \hat{p}(s))$ for $s \in [0, T]$ is the solution of Hamilton's equations with the above $H$ such that $\hat{q}(t) = q$; $S^*$ is the optimal cost

$$S^* := \int_0^T \left[ \hat{p} \cdot \dot{\hat{q}} - H(\hat{q}, \hat{p}) \right] ds = \int_0^T \left[ \hat{p} \cdot \dot{\hat{q}} - \hat{H}(\hat{q}, \hat{p}, u^*(\hat{q}, \hat{p})) \right] ds = \hat{S}[u^*],$$

and the function $S(q, t)$ is defined by

$$S(q, t) := \int_0^t \left[ \hat{p} \cdot \dot{\hat{q}} - H(\hat{q}, \hat{p}) \right] ds.$$

Since this definition coincides with (1.2), the function $S(q, t) = S^* - J(q, t)$ satisfies the H-J equation (1.3); this reduces to the Hamilton–Jacobi–Bellman (HJB) equation for the optimal cost-to-go function $J(q, t)$:

(1.4)
$$\frac{\partial J}{\partial t} + \min_u \left[ \frac{\partial J}{\partial q} \cdot f(q, u) + C(q, u) \right] = 0.$$

It can also be shown that the costate $p$ of the optimal solution is related to the solution of the HJB equation.

**1.4. Discrete Hamilton–Jacobi theory.** The main objective of this paper is to present a discrete analogue of Hamilton–Jacobi theory within the framework of discrete Hamiltonian mechanics of Lall and West [23], and also to apply the theory to discrete optimal control problems.

There are some previous works on discrete-time analogues of the Hamilton–Jacobi equation, such as Elnatanov and Schiff [13] and Lall and West [23]. Specifically, Elnatanov and Schiff [13] derived an equation for a generating function of a coordinate transformation that trivializes the dynamics. This derivation is a discrete analogue of the conventional derivation of the continuous-time Hamilton–Jacobi equation (see, e.g., Lanczos [24]). Lall and West [23] formulated a discrete Lagrangian analogue of the Hamilton–Jacobi equation as a separable optimization problem.

**1.5. Main results.** Our work was inspired by the result of Elnatanov and Schiff [13] and starts from a reinterpretation of their result in the language of discrete mechanics. This paper further extends the result by developing discrete analogues

| **Discrete Evolution Equations** | | **Hamilton–Jacobi-type Equations** |
|---|---|---|
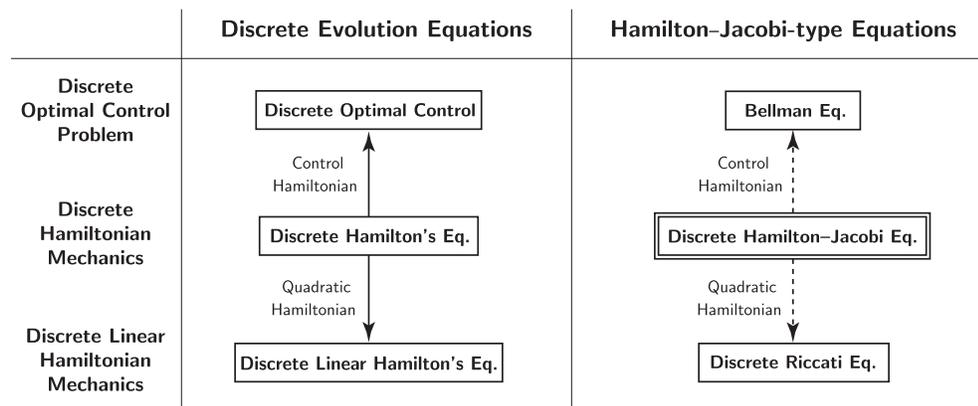


FIG. 1. *Discrete evolution equations (left) and corresponding discrete Hamilton–Jacobi-type equations (right). Dashed lines are the links established in the paper.*

of results in (continuous-time) Hamilton–Jacobi theory. Namely, we formulate a discrete analogue of Jacobi's solution, which relates the discrete action sum to a solution of the discrete Hamilton–Jacobi equation. This also provides a very simple derivation of the discrete Hamilton–Jacobi equation and exhibits a natural correspondence with the continuous-time theory. Another important result in this paper is a discrete analogue of the Hamilton–Jacobi theorem, which relates the solution of the discrete Hamilton–Jacobi equation with the solution of the discrete Hamilton's equations.

We also show that the discrete Hamilton–Jacobi equation is a generalization of the discrete Riccati equation and the Bellman equation (see Figure 1). Specifically, we show that the discrete Hamilton–Jacobi equation applied to linear discrete Hamiltonian systems and discrete optimal control problems reduces to the discrete Riccati and Bellman equations, respectively. This is again a discrete analogue of the well-known results that the Hamilton–Jacobi equation applied to linear Hamiltonian systems and optimal control problems reduces to the Riccati (see, e.g., Jurdjevic [22, p. 421]) and HJB equations (see section 1.3 above), respectively.

The link between the discrete Hamilton–Jacobi equation and the Bellman equation turns out to be useful in deriving a class of generalized Bellman equations that are higher-order approximations of the original continuous-time problem. Specifically, we use the idea of the Galerkin Hamiltonian variational integrator of Leok and Zhang [28] to derive discrete control Hamiltonians that yield higher-order approximations, and then show that the corresponding discrete Hamilton–Jacobi equation gives a class of Bellman equations with controls at internal stages.

**1.6. Outline of the paper.** We first present a brief review of discrete Lagrangian and Hamiltonian mechanics in section 2. In section 3, we describe a reinterpretation of the result of Elnatanov and Schiff [13] in the language of discrete mechanics and a discrete analogue of Jacobi's solution to the discrete Hamilton–Jacobi equation. The remainder of section 3 is devoted to more detailed studies of the discrete Hamilton–Jacobi equation: its left and right variants, more explicit forms of them, and also a digression on the Lagrangian side. In section 4, we prove a discrete version of the Hamilton–Jacobi theorem. In section 5, we apply the theory to linear discrete Hamiltonian systems, and show that the discrete Riccati equation follows from the discrete Hamilton–Jacobi equation. Section 6 establishes the link with

discrete-time optimal control and interprets the results of the preceding sections in this setting. Section 7 further extends this idea to derive a class of Bellman equations with controls at internal stages.

**2. Discrete mechanics.** This section briefly reviews some key results of discrete mechanics following Marsden and West [30] and Lall and West [23].

**2.1. Discrete Lagrangian mechanics.** A discrete Lagrangian flow $\{q_k\}_{k=0}^N$, on an $n$-dimensional differentiable manifold $Q$, can be described by the following discrete variational principle: Let $S_{\mathrm{d}}^N$ be the following action sum of the discrete Lagrangian $L_{\mathrm{d}} : Q \times Q \to \mathbb{R}$:

$$(2.1) \qquad S_{\mathrm{d}}^N(\{q_k\}_{k=0}^N) := \sum_{k=0}^{N-1} L_{\mathrm{d}}(q_k, q_{k+1}) \approx \int_0^{t_N} L(q(t), \dot{q}(t))\, dt,$$

which is an approximation of the action integral as shown above.

Consider discrete variations $q_k \mapsto q_k + \varepsilon\, \delta q_k$, for $k = 0, 1, \ldots, N$, with $\delta q_0 = \delta q_N = 0$. Then, the discrete variational principle $\delta S_{\mathrm{d}}^N = 0$ gives the discrete Euler–Lagrange equations:

$$(2.2) \qquad D_2 L_{\mathrm{d}}(q_{k-1}, q_k) + D_1 L_{\mathrm{d}}(q_k, q_{k+1}) = 0.$$

This determines the discrete flow $F_{L_{\mathrm{d}}} : Q \times Q \to Q \times Q$:

$$(2.3) \qquad F_{L_{\mathrm{d}}} : (q_{k-1}, q_k) \mapsto (q_k, q_{k+1}).$$

Let us define the discrete Lagrangian symplectic one-forms $\Theta_{L_{\mathrm{d}}}^{\pm} : Q \times Q \to T^*(Q \times Q)$ by

$$(2.4a) \qquad \Theta_{L_{\mathrm{d}}}^+ : (q_k, q_{k+1}) \mapsto D_2 L_{\mathrm{d}}(q_k, q_{k+1})\, dq_{k+1},$$

$$(2.4b) \qquad \Theta_{L_{\mathrm{d}}}^- : (q_k, q_{k+1}) \mapsto -D_1 L_{\mathrm{d}}(q_k, q_{k+1})\, dq_k.$$

Then, the discrete flow $F_{L_{\mathrm{d}}}$ preserves the discrete Lagrangian symplectic form

$$(2.5) \qquad \Omega_{L_{\mathrm{d}}}(q_k, q_{k+1}) := -d\Theta_{L_{\mathrm{d}}}^{\pm} = -D_1 D_2 L_{\mathrm{d}}(q_k, q_{k+1})\, dq_k \wedge dq_{k+1}.$$

Specifically, we have

$$(F_{L_{\mathrm{d}}})^* \Omega_{L_{\mathrm{d}}} = \Omega_{L_{\mathrm{d}}}.$$

**2.2. Discrete Hamiltonian mechanics.** Introduce the *right and left discrete Legendre transforms* $\mathbb{F}L_{\mathrm{d}}^{\pm} : Q \times Q \to T^*Q$ by

$$(2.6a) \qquad \mathbb{F}L_{\mathrm{d}}^+ : (q_k, q_{k+1}) \mapsto (q_{k+1}, D_2 L_{\mathrm{d}}(q_k, q_{k+1})),$$

$$(2.6b) \qquad \mathbb{F}L_{\mathrm{d}}^- : (q_k, q_{k+1}) \mapsto (q_k, -D_1 L_{\mathrm{d}}(q_k, q_{k+1})),$$

respectively. Then, we find that the discrete Lagrangian symplectic forms in (2.4) and (2.5) are pullbacks by these maps of the standard symplectic forms on $T^*Q$:

$$\Theta_{L_{\mathrm{d}}}^{\pm} = (\mathbb{F}L_{\mathrm{d}}^{\pm})^* \Theta, \qquad \Omega_{L_{\mathrm{d}}}^{\pm} = (\mathbb{F}L_{\mathrm{d}}^{\pm})^* \Omega.$$

Let us define the momenta

$$p_{k,k+1}^- := -D_1 L_{\mathrm{d}}(q_k, q_{k+1}), \qquad p_{k,k+1}^+ := D_2 L_{\mathrm{d}}(q_k, q_{k+1}).$$

Then, the discrete Euler–Lagrange equations (2.2) become simply $p^+_{k-1,k} = p^-_{k,k+1}$. So, defining

$$p_k := p^+_{k-1,k} = p^-_{k,k+1},$$

one can rewrite the discrete Euler–Lagrange equations (2.2) as follows:

(2.7)
$$\begin{aligned} p_k &= -D_1 L_{\mathrm{d}}(q_k, q_{k+1}), \\ p_{k+1} &= D_2 L_{\mathrm{d}}(q_k, q_{k+1}). \end{aligned}$$

Furthermore, define the *discrete Hamiltonian map* $\tilde{F}_{L_{\mathrm{d}}} : T^*Q \to T^*Q$ by

(2.8)
$$\tilde{F}_{L_{\mathrm{d}}} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1}).$$

Then, one may relate this map with the discrete Legendre transforms in (2.6) as follows:

(2.9)
$$\tilde{F}_{L_{\mathrm{d}}} = \mathbb{F}L^+_{\mathrm{d}} \circ (\mathbb{F}L^-_{\mathrm{d}})^{-1}.$$

Furthermore, one can also show that this map is symplectic, i.e.,

$$(\tilde{F}_{L_{\mathrm{d}}})^* \Omega = \Omega.$$

This is the Hamiltonian description of the dynamics defined by the discrete Euler–Lagrange equation (2.2) introduced by Marsden and West [30]. Notice, however, that no discrete analogue of Hamilton's equations is introduced here, although the flow is now on the cotangent bundle $T^*Q$.

Lall and West [23] pushed this idea further to give discrete analogues of Hamilton's equations: From the point of view that a discrete Lagrangian is essentially a generating function of type one [16], we can apply Legendre transforms to the discrete Lagrangian to find the corresponding generating functions of type two or three [16]. In fact, they turn out to be a natural Hamiltonian counterpart to the discrete Lagrangian mechanics described above. Specifically, with the right discrete Legendre transform

(2.10)
$$p_{k+1} = \mathbb{F}L^+_{\mathrm{d}}(q_k, q_{k+1}) = D_2 L_{\mathrm{d}}(q_k, q_{k+1}),$$

we can define the following *right discrete Hamiltonian*:

(2.11)
$$H^+_{\mathrm{d}}(q_k, p_{k+1}) = p_{k+1} \cdot q_{k+1} - L_{\mathrm{d}}(q_k, q_{k+1}).$$

Then, the discrete Hamiltonian map $\tilde{F}_{L_{\mathrm{d}}} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$ is defined implicitly by the *right discrete Hamilton's equations*

(2.12a)
$$q_{k+1} = D_2 H^+_{\mathrm{d}}(q_k, p_{k+1}),$$
(2.12b)
$$p_k = D_1 H^+_{\mathrm{d}}(q_k, p_{k+1}),$$

which are precisely the characterization of a symplectic map in terms of a generating function, $H^+_{\mathrm{d}}$, of type two. Similarly, with the left discrete Legendre transform

(2.13)
$$p_k = \mathbb{F}L^-_{\mathrm{d}}(q_k, q_{k+1}) = -D_1 L_{\mathrm{d}}(q_k, q_{k+1}),$$

we can define the following *left discrete Hamiltonian*:

(2.14)
$$H^-_{\mathrm{d}}(p_k, q_{k+1}) = -p_k \cdot q_k - L_{\mathrm{d}}(q_k, q_{k+1}).$$

Then, we have the *left discrete Hamilton's equations*

(2.15a) $$q_k = -D_1 H_{\mathrm{d}}^-(p_k, q_{k+1}),$$

(2.15b) $$p_{k+1} = -D_2 H_{\mathrm{d}}^-(p_k, q_{k+1}),$$

which correspond to a symplectic map expressed in terms of a generation function, $H_{\mathrm{d}}^-$, of type three.

On the other hand, Leok and Zhang [28] demonstrate that discrete Hamiltonian mechanics can be obtained as a direct variational discretization of continuous Hamiltonian mechanics, instead of having to go via discrete Lagrangian mechanics.

## 3. Discrete Hamilton–Jacobi equation.

**3.1. Derivation by Elnatanov and Schiff.** Elnatanov and Schiff [13] derived a discrete Hamilton–Jacobi equation based on the idea that the Hamilton–Jacobi equation is an equation for a symplectic change of coordinates under which the dynamics becomes trivial. In this section, we would like to reinterpret their derivation in the framework of discrete Hamiltonian mechanics reviewed above.

THEOREM 3.1. *Suppose that the discrete dynamics $\{(q_k, p_k)\}_{k=0}^N$ is governed by the right discrete Hamilton's equations (2.12). Consider the symplectic coordinate transformation $(q_k, p_k) \mapsto (\hat{q}_k, \hat{p}_k)$ that satisfies the following:*

(i) *The old and new coordinates are related by the type-one generating function[1] $S^k : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$:*

(3.1) $$\hat{p}_k = -D_1 S^k(\hat{q}_k, q_k),$$
$$p_k = D_2 S^k(\hat{q}_k, q_k);$$

(ii) *the dynamics in the new coordinates $\{(\hat{q}_k, \hat{p}_k)\}_{k=0}^N$ is rendered trivial, i.e., $(\hat{q}_{k+1}, \hat{p}_{k+1}) = (\hat{q}_k, \hat{p}_k)$.*

*Then, the set of functions $\{S^k\}_{k=1}^N$ satisfies the* discrete Hamilton–Jacobi equation:

(3.2)
$$S^{k+1}(\hat{q}_0, q_{k+1}) - S^k(\hat{q}_0, q_k) - D_2 S^{k+1}(\hat{q}_0, q_{k+1}) \cdot q_{k+1} + H_{\mathrm{d}}^+\big(q_k, D_2 S^{k+1}(\hat{q}_0, q_{k+1})\big) = 0,$$

*or, with the shorthand notation $S_{\mathrm{d}}^k(q_k) := S^k(\hat{q}_0, q_k)$,*

(3.3) $$S_{\mathrm{d}}^{k+1}(q_{k+1}) - S_{\mathrm{d}}^k(q_k) - DS_{\mathrm{d}}^{k+1}(q_{k+1}) \cdot q_{k+1} + H_{\mathrm{d}}^+\big(q_k, DS_{\mathrm{d}}^{k+1}(q_{k+1})\big) = 0.$$

*Proof.* The key ingredient in the proof is the right discrete Hamiltonian in the new coordinates, i.e., a function $\hat{H}_{\mathrm{d}}^+(\hat{q}_k, \hat{p}_{k+1})$ that satisfies

(3.4) $$\hat{q}_{k+1} = D_2 \hat{H}_{\mathrm{d}}^+(\hat{q}_k, \hat{p}_{k+1}),$$
$$\hat{p}_k = D_1 \hat{H}_{\mathrm{d}}^+(\hat{q}_k, \hat{p}_{k+1}),$$

or equivalently,

(3.5) $$\hat{p}_k \, d\hat{q}_k + \hat{q}_{k+1} \, d\hat{p}_{k+1} = d\hat{H}_{\mathrm{d}}^+(\hat{q}_k, \hat{p}_{k+1}).$$

---

[1] This is essentially the same as (2.7) in the sense that they are both transformations defined by generating functions of type one: Replace $(q_k, p_k, q_{k+1}, p_{k+1}, L_{\mathrm{d}})$ by $(\hat{q}_k, \hat{p}_k, q_k, p_k, S^k)$. However, they have different interpretations: (2.7) describes the dynamics or time evolution, whereas (3.1) is a change of coordinates.

Let us first write $\hat{H}_{\mathrm{d}}^{+}$ in terms of the original right discrete Hamiltonian $H_{\mathrm{d}}^{+}$ and the generating function $S^{k}$. For that purpose, first rewrite (2.12) and (3.1) as follows:

$$p_{k}\,dq_{k} = -q_{k+1}\,dp_{k+1} + dH_{\mathrm{d}}^{+}(q_{k}, p_{k+1})$$

and

$$\hat{p}_{k}\,d\hat{q}_{k} = p_{k}\,dq_{k} - dS^{k}(\hat{q}_{k}, q_{k}),$$

respectively. Then, using the above relations, we have

$$
\begin{aligned}
\hat{p}_{k}\,d\hat{q}_{k} + \hat{q}_{k+1}\,d\hat{p}_{k+1} &= \hat{p}_{k}\,d\hat{q}_{k} + d(\hat{p}_{k+1}\cdot\hat{q}_{k+1}) - \hat{p}_{k+1}\,d\hat{q}_{k+1}\\
&= p_{k}\,dq_{k} - dS^{k}(\hat{q}_{k}, q_{k}) + d(\hat{p}_{k+1}\cdot\hat{q}_{k+1})\\
&\quad - p_{k+1}\,dq_{k+1} + dS^{k+1}(\hat{q}_{k+1}, q_{k+1})\\
&= -q_{k+1}\,dp_{k+1} + dH_{\mathrm{d}}^{+}(q_{k}, p_{k+1}) - dS^{k}(\hat{q}_{k}, q_{k})\\
&\quad + d(\hat{p}_{k+1}\hat{q}_{k+1}) - p_{k+1}\,dq_{k+1} + dS^{k+1}(\hat{q}_{k+1}, q_{k+1})\\
&= d\big(H_{\mathrm{d}}^{+}(q_{k}, p_{k+1}) + \hat{p}_{k+1}\cdot\hat{q}_{k+1} - p_{k+1}\cdot q_{k+1}\\
&\quad + S^{k+1}(\hat{q}_{k+1}, q_{k+1}) - S^{k}(\hat{q}_{k}, q_{k})\big).
\end{aligned}
$$

Thus, in view of (3.5), we obtain
(3.6)
$$\hat{H}_{\mathrm{d}}^{+}(\hat{q}_{k}, \hat{p}_{k+1}) = H_{\mathrm{d}}^{+}(q_{k}, p_{k+1}) + \hat{p}_{k+1}\cdot\hat{q}_{k+1} - p_{k+1}\cdot q_{k+1} + S^{k+1}(\hat{q}_{k+1}, q_{k+1}) - S^{k}(\hat{q}_{k}, q_{k}).$$

Now consider the choice of the new right discrete Hamiltonian $\hat{H}_{\mathrm{d}}^{+}$ that renders the dynamics trivial, i.e., $(\hat{q}_{k+1}, \hat{p}_{k+1}) = (\hat{q}_{k}, \hat{p}_{k})$. It is clear from (3.4) that we can set

$$\hat{H}_{\mathrm{d}}^{+}(\hat{q}_{k}, \hat{p}_{k+1}) = \hat{p}_{k+1}\cdot\hat{q}_{k}.$$

Then, (3.6) becomes

$$\hat{p}_{k+1}\cdot\hat{q}_{k} = H_{\mathrm{d}}^{+}(q_{k}, p_{k+1}) + \hat{p}_{k+1}\cdot\hat{q}_{k+1} - p_{k+1}\cdot q_{k+1} + S^{k+1}(\hat{q}_{k+1}, q_{k+1}) - S^{k}(\hat{q}_{k}, q_{k}),$$

and since $\hat{q}_{k+1} = \hat{q}_{k} = \cdots = \hat{q}_{0}$, we have

$$0 = H_{\mathrm{d}}^{+}(q_{k}, p_{k+1}) - p_{k+1}\cdot q_{k+1} + S^{k+1}(\hat{q}_{0}, q_{k+1}) - S^{k}(\hat{q}_{0}, q_{k}).$$

Eliminating $p_{k+1}$ by using (3.1), we obtain (3.2).  □

*Remark* 3.2. What Elnatanov and Schiff [13] refer to as the *Hamilton–Jacobi difference equation* is the following:
(3.7)
$$S^{k+1}(\hat{q}_{0}, q_{k+1}) - S^{k}(\hat{q}_{0}, q_{k}) - D_{2}S^{k+1}(\hat{q}_{0}, q_{k+1})\cdot D_{2}H_{\mathrm{d}}^{+}(q_{k}, p_{k+1}) + H_{\mathrm{d}}^{+}(q_{k}, p_{k+1}) = 0.$$

It is clear that this is equivalent to (3.2) in view of (2.12)

**3.2. Discrete analogue of Jacobi's solution.** This section presents a discrete analogue of Jacobi's solution. This also gives an alternative derivation of the discrete Hamilton–Jacobi equation that is much simpler than the one shown above.

THEOREM 3.3. *Consider the action sums, (2.1), written in terms of the right discrete Hamiltonian, (2.11):*

(3.8)
$$S_{\mathrm{d}}^{k}(q_{k}) := \sum_{l=0}^{k-1}\big[p_{l+1}\cdot q_{l+1} - H_{\mathrm{d}}^{+}(q_{l}, p_{l+1})\big]$$

*evaluated along a solution of the right discrete Hamilton's equations (2.12); each $S_d^k(q_k)$ is seen as a function of the endpoint coordinates $q_k$ and the discrete end time $k$. Then, these action sums satisfy the discrete Hamilton–Jacobi equation (3.3).*

*Proof.* From (3.8), we have

$$(3.9) \qquad S_d^{k+1}(q_{k+1}) - S_d^k(q_k) = p_{k+1} \cdot q_{k+1} - H_d^+(q_k, p_{k+1}),$$

where $p_{k+1}$ is considered to be a function of $q_k$ and $q_{k+1}$, i.e., $p_{k+1} = p_{k+1}(q_k, q_{k+1})$. Taking the derivative of both sides with respect to $q_{k+1}$, we have

$$DS_d^{k+1}(q_{k+1}) = p_{k+1} + \frac{\partial p_{k+1}}{\partial q_{k+1}} \cdot \left[ q_{k+1} - D_2 H_d^+(q_k, p_{k+1}) \right].$$

However, the terms in the brackets vanish because the right discrete Hamilton's equations (2.12) are assumed to be satisfied. Thus, we have

$$(3.10) \qquad p_{k+1} = DS_d^{k+1}(q_{k+1}).$$

Substituting this into (3.9) gives (3.3). $\qquad \square$

*Remark* 3.4. Recall that, in the derivation of the continuous Hamilton–Jacobi equation (see, e.g., [15, Section 23]), we consider the variation of the action integral, (1.2), with respect to the endpoint $(q, t)$ and find

$$(3.11) \qquad dS = p \, dq - H(q, p) \, dt.$$

This gives

$$\frac{\partial S}{\partial t} = -H(q, p), \qquad p = \frac{\partial S}{\partial q},$$

and hence the Hamilton–Jacobi equation

$$\frac{\partial S}{\partial t} + H\left( q, \frac{\partial S}{\partial q} \right) = 0.$$

In the above derivation of the discrete Hamilton–Jacobi equation (3.3), the difference of two action sums, (3.9), is a natural discrete analogue of the variation $dS$ in (3.11). Notice also that (3.9) plays the same essential role as (3.11) does in deriving the Hamilton–Jacobi equation. Table 1 summarizes the correspondence between the ingredients in the continuous and discrete theories.

**3.3. The right and left discrete Hamilton–Jacobi equations.** Recall that, in (3.8), we wrote the action sum, (2.1), in terms of the right discrete Hamiltonian, (2.11). We can also write it in terms of the left discrete Hamiltonian, (2.14), as follows:

$$(3.12) \qquad S_d^k(q_k) = \sum_{l=0}^{k-1} \left[ -p_l \cdot q_l - H_d^-(p_l, q_{l+1}) \right].$$

Then, we can proceed as in the proof of Theorem 3.3: First, we have

$$(3.13) \qquad S_d^{k+1}(q_{k+1}) - S_d^k(q_k) = -p_k \cdot q_k - H_d^-(p_k, q_{k+1}),$$

where $p_k$ is considered to be a function of $q_k$ and $q_{k+1}$, i.e., $p_k = p_k(q_k, q_{k+1})$. Taking the derivative of both sides with respect to $q_k$, we have

$$-DS_d^k(q_k) = -p_k - \frac{\partial p_k}{\partial q_k} \cdot \left[ q_k + D_1 H_d^-(p_k, q_{k+1}) \right].$$

TABLE 1
*Correspondence between ingredients in continuous and discrete theories; $\mathbb{R}_{\geq 0}$ is the set of non-negative real numbers and $\mathbb{N}_0$ is the set of nonnegative integers (see also Remark 3.4).*

| Continuous | Discrete |
|---|---|
| $(q, t) \in Q \times \mathbb{R}_{\geq 0}$ | $(q_k, k) \in Q \times \mathbb{N}_0$ |
| $\dot{q} = \partial H / \partial p,$ <br> $\dot{p} = -\partial H / \partial q$ | $q_{k+1} = D_2 H_{\mathrm{d}}^+(q_k, p_{k+1}),$ <br> $p_k = D_1 H_{\mathrm{d}}^+(q_k, p_{k+1})$ |
| $S(q,t) := \int_0^t [p(s) \cdot \dot{q}(s) - H(q(s), p(s))] \, ds$ | $S_{\mathrm{d}}^k(q_k) := \sum_{l=0}^{k-1} \left[ p_{l+1} \cdot q_{l+1} - H_{\mathrm{d}}^+(q_l, p_{l+1}) \right]$ |
| $dS = \dfrac{\partial S}{\partial q} \, dq + \dfrac{\partial S}{\partial t} \, dt$ | $S_{\mathrm{d}}^{k+1}(q_{k+1}) - S_{\mathrm{d}}^k(q_k)$ |
| $p \, dq - H(q,p) \, dt$ | $p_{k+1} \cdot q_{k+1} - H_{\mathrm{d}}^+(q_k, p_{k+1})$ |
| $\dfrac{\partial S}{\partial t} + H\left(q, \dfrac{\partial S}{\partial q}\right) = 0$ | $S_{\mathrm{d}}^{k+1}(q_{k+1}) - S_{\mathrm{d}}^k(q_k) - DS_{\mathrm{d}}^{k+1}(q_{k+1}) \cdot q_{k+1}$ <br> $+ H_{\mathrm{d}}^+\left(q_k, DS_{\mathrm{d}}^{k+1}(q_{k+1})\right) = 0$ |

However, the terms in the brackets vanish because the left discrete Hamilton's equations (2.15) are assumed to be satisfied. Thus, we have

$$(3.14) \qquad\qquad p_k = DS_{\mathrm{d}}^k(q_k).$$

Substituting this into (3.13) gives the discrete Hamilton–Jacobi equation with the left discrete Hamiltonian:

$$(3.15) \qquad S_{\mathrm{d}}^{k+1}(q_{k+1}) - S_{\mathrm{d}}^k(q_k) + DS_{\mathrm{d}}^k(q_k) \cdot q_k + H_{\mathrm{d}}^-\left(DS_{\mathrm{d}}^k(q_k), q_{k+1}\right) = 0.$$

We refer to (3.3) and (3.15) as the *right and left discrete Hamilton–Jacobi equations*, respectively.

As mentioned above, (3.8) and (3.12) are the same action sum, (2.1), expressed in different ways. Therefore, we may summarize the above argument as follows.

PROPOSITION 3.5. *The action sums, (3.8) or equivalently (3.12), satisfy both the right and left discrete Hamilton–Jacobi equations, (3.3) and (3.15), respectively.*

**3.4. Explicit forms of the discrete Hamilton–Jacobi equations.** The expressions for the right and left discrete Hamilton–Jacobi equations in (3.3) and (3.15) are implicit in the sense that they contain two spatial variables $q_k$ and $q_{k+1}$; Theorem 3.3 suggests that one may consider $q_k$ and $q_{k+1}$ to be related by the discrete Hamiltonian dynamics defined by either the right or left discrete Hamilton's equations (2.12) or (2.15), or equivalently, the discrete Hamiltonian map $\tilde{F}_{L_{\mathrm{d}}} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$ defined in (2.8). More specifically, we may write $q_{k+1}$ in terms of $q_k$. This results in explicit forms of the discrete Hamilton–Jacobi equations, and we shall *define* the discrete Hamilton–Jacobi equations by the resulting explicit forms. We will see later in section 6 that the explicit form is compatible with the formulation of the Bellman equation.

For the right discrete Hamilton–Jacobi equation (3.3), we first define the map $f_k^+ : Q \to Q$ as follows: Replace $p_{k+1}$ in (2.12a) by $DS_{\mathrm{d}}^{k+1}(q_{k+1})$ as suggested by (3.10):

$$q_{k+1} = D_2 H_{\mathrm{d}}^+\left(q_k, DS_{\mathrm{d}}^{k+1}(q_{k+1})\right).$$

Assuming this equation is solvable for $q_{k+1}$, we define $f_k^+ : Q \to Q$ by $f_k(q_k) = q_{k+1}$; i.e., $f_k^+$ is implicitly defined by

(3.16)
$$f_k^+(q_k) = D_2 H_d^+ \left( q_k, DS_d^{k+1}(f_k^+(q_k)) \right).$$

We may now identify $q_{k+1}$ with $f_k^+(q_k)$ in the implicit form of the right Hamilton–Jacobi equation (3.3):

(3.17) $\quad S_d^{k+1}(f_k^+(q)) - S_d^k(q) - DS_d^{k+1}(f_k^+(q)) \cdot f_k^+(q) + H_d^+ \left( q, DS_d^{k+1}(f_k^+(q)) \right) = 0,$

where we suppressed the subscript $k$ of $q_k$ since it is now clear that $q_k$ is an independent variable as opposed to a function of the discrete time $k$. We *define* (3.17) to be the *right discrete Hamilton–Jacobi equation*. Notice that these are differential-difference-functional equations defined on $Q \times \mathbb{N}$, with the spatial variable $q$ and the discrete time $k$.

For the left discrete Hamilton–Jacobi equation (3.15), we define the map $f_k^- : Q \to Q$ as follows:

(3.18)
$$f_k^-(q_k) := \pi_Q \circ \tilde{F}_{L_d} \left( dS_d^k(q_k) \right),$$

where $\pi_Q : T^*Q \to Q$ is the cotangent bundle projection; equivalently, $f_k^-$ is defined so that the diagram below commutes.

$$
\begin{array}{ccc}
T^*Q & \xrightarrow{\tilde{F}_{L_d}} & T^*Q \\
{\scriptstyle dS_d^k} \big\uparrow & & \big\downarrow {\scriptstyle \pi_Q} \\
Q & \dashrightarrow{\;\; f_k^- \;\;} & Q
\end{array}
\qquad
\begin{array}{ccc}
dS_d^k(q_k) & \longmapsto & \tilde{F}_{L_d}\left( dS_d^k(q_k) \right) \\
\big\uparrow & & \big\downarrow \\
q_k & \dashrightarrow & f_k^-(q_k)
\end{array}
$$

Notice also that, since the map $\tilde{F}_{L_d} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$ is defined by (2.15), $f_k^-$ is defined implicitly by

(3.19)
$$q_k = -D_1 H_d^- \left( DS_d^k(q_k), f_k^-(q_k) \right).$$

In other words, replace $p_k$ in (2.15a) by $DS_d^k(q_k)$ as suggested by (3.14), and define $f_k^-(q_k)$ as the $q_{k+1}$ in the resulting equation.

We may now identify $q_{k+1}$ with $f_k^-(q_k)$ in (3.15):

(3.20) $\quad S_d^{k+1}(f_k^-(q)) - S_d^k(q) + DS_d^k(q) \cdot q + H_d^- \left( DS_d^k(q), f_k^-(q) \right) = 0,$

where we again suppressed the subscript $k$ of $q_k$. We *define* (3.17) and (3.20) to be the *right and left discrete Hamilton–Jacobi equations*, respectively.

*Remark* 3.6. Notice that the right discrete Hamilton–Jacobi equation (3.17) is more complicated than the left one (3.20), particularly because the map $f_k^+$ appears more often than $f_k^-$ does in the latter; notice here that, as shown in (3.18), the maps $f_k^{\pm}$ in the discrete Hamilton–Jacobi equations (3.17) and (3.20) depend on the function $S_d^k$, which is the unknown one has to solve for. However, it is possible to define an equally simple variant of the right discrete Hamilton–Jacobi equation by writing $q_{k-1}$ in terms of $q_k$: Let us first define $g_k : Q \to Q$ by

$$g_k(q_k) := \pi_Q \circ \tilde{F}_{L_d}^{-1} \left( dS_d^k(q_k) \right),$$

or so that the diagram below commutes.

$$
\begin{array}{ccc}
T^*Q & \xleftarrow{\ \tilde{F}_{L_\mathrm{d}}^{-1}\ } & T^*Q \\
\pi_Q \downarrow & & \downarrow dS_\mathrm{d}^k \\
Q & \xleftarrow{\ \ g_k\ \ } & Q
\end{array}
\qquad\qquad
\begin{array}{ccc}
\tilde{F}_{L_\mathrm{d}}^{-1}\big(dS_\mathrm{d}^k(q_k)\big) & \longleftarrow & dS_\mathrm{d}^k(q_k) \\
\downarrow & & \uparrow \\
g_k(q_k) & \longleftarrow & q_k
\end{array}
$$

Now, in (3.3), change the indices from $(k, k+1)$ to $(k-1, k)$ and identify $q_{k-1}$ with $g_k(q_k)$ to obtain

$$
S_\mathrm{d}^k(q) - S_\mathrm{d}^{k-1}(g_k(q)) - DS_\mathrm{d}^k(q) \cdot q + H_\mathrm{d}^+\big(g_k(q), DS_\mathrm{d}^k(q)\big) = 0,
$$

where we again suppressed the subscript $k$ of $q_k$. This is as simple as the left discrete Hamilton–Jacobi equation (3.20). However, the map $g_k$ is, being backward in time, rather unnatural compared to $f_k^\pm$. Furthermore, as we shall see in section 6, in the discrete optimal control setting, the map $f_k^\pm$ is defined by a given function and thus the formulation with $f_k^\pm$ will turn out to be more convenient.

**3.5. The discrete Hamilton–Jacobi equation on the Lagrangian side.** First, notice that (2.1) gives

(3.21) $$ S_\mathrm{d}^{k+1}(q_{k+1}) - S_\mathrm{d}^k(q_k) = L_\mathrm{d}(q_k, q_{k+1}). $$

This is essentially the Lagrangian equivalent of the discrete Hamilton–Jacobi equation (3.17) as Lall and West [23] suggest. Let us apply the same argument as above to obtain the explicit form for (3.21). Taking the derivative of the above equation with respect to $q_k$, we have

$$
-D_1 L_\mathrm{d}(q_k, q_{k+1})\, dq_k = dS_\mathrm{d}^k(q_k),
$$

and hence from the definition of the left discrete Legendre transform, (2.6b),

$$
\mathbb{F}L_\mathrm{d}^-(q_k, q_{k+1}) = dS_\mathrm{d}^k(q_k).
$$

Assuming that $\mathbb{F}L_\mathrm{d}^-$ is invertible, we have

$$
(q_k, q_{k+1}) = (\mathbb{F}L_\mathrm{d}^-)^{-1}\big(dS_\mathrm{d}^k(q_k)\big) =: (q_k, f_k^L(q_k)),
$$

where we defined the map $f_k^L : Q \to Q$ as follows (see the commutative diagram below):

(3.22) $$ f_k^L(q_k) := pr_2 \circ (\mathbb{F}L_\mathrm{d}^-)^{-1}\big(dS_\mathrm{d}^k(q_k)\big), $$

where $pr_2 : Q \times Q \to Q$ is the projection to the second factor, i.e., $pr_2(q_1, q_2) = q_2$. Thus, eliminating $q_{k+1}$ from (3.21) and then replacing $q_k$ by $q$, we obtain the discrete Hamilton–Jacobi equation on the Lagrangian side:

(3.23) $$ S_\mathrm{d}^{k+1}(f_k^L(q)) - S_\mathrm{d}^k(q) = L_\mathrm{d}\big(q, f_k^L(q)\big). $$

The map $f_k^L$ defined in (3.22) is identical to $f_k^-$ defined above in (3.18) as the commutative diagram below demonstrates.
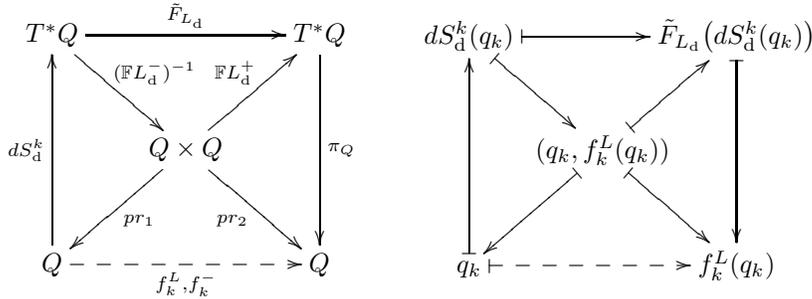


The commutativity of the square in the diagram defines the $f_k^-$ as we saw earlier, whereas that of the right-angled triangle on the lower left defines the $f_k^L$ in (3.22); note the relation $\tilde{F}_{L_d} = \mathbb{F}L_d^+ \circ (\mathbb{F}L_d^-)^{-1}$ from (2.9).

The map $f_k^L$ being identical to $f_k^-$ implies that the discrete Hamilton–Jacobi equations on the Hamiltonian and Lagrangian sides, (3.20) and (3.23), are equivalent.

**4. Discrete Hamilton–Jacobi theorem.** The following gives a discrete analogue of the geometric Hamilton–Jacobi theorem by Abraham and Marsden [1, Theorem 5.2.4].

THEOREM 4.1 (discrete Hamilton–Jacobi). *Suppose that $S_d^k$ satisfies the right discrete Hamilton–Jacobi equation* (3.17), *and let $\{c_k\}_{k=0}^N \subset Q$ be a set of points such that*

$$(4.1) \qquad c_{k+1} = f_k^+(c_k) \quad for \quad k = 0, 1, \ldots, N - 1.$$

*Then, the set of points $\{(c_k, p_k)\}_{k=0}^N \subset T^*Q$ with*

$$(4.2) \qquad p_k := DS_d^k(c_k)$$

*is a solution of the right discrete Hamilton's equations* (2.12).

*Similarly, suppose that $S_d^k$ satisfies the left discrete Hamilton–Jacobi equation* (3.20), *and let $\{c_k\}_{k=0}^N \subset Q$ be a set of points that satisfy*

$$(4.3) \qquad c_{k+1} = f_k^-(c_k) \quad for \quad k = 0, 1, \ldots, N - 1.$$

*Furthermore, assume that the Jacobian $Df_k^-$ is invertible at each point $c_k$. Then, the set of points $\{(c_k, p_k)\}_{k=0}^N \subset T^*Q$ with*

$$(4.4) \qquad p_k := DS_d^k(c_k)$$

*is a solution of the left discrete Hamilton's equations* (2.15).

*Proof.* To prove the first assertion, first recall the implicit definition of $f_k^+$ in (3.16):

$$(4.5) \qquad f_k^+(q) = D_2 H_d^+\left(q, DS_d^{k+1}(f_k^+(q))\right).$$

In particular, for $q = c_k$, we have

$$(4.6) \qquad c_{k+1} = D_2 H_d^+(c_k, p_k),$$

where we used (4.1) and (4.2). On the other hand, taking the derivative of (3.17) with respect to $q$,

$$
DS_{\mathrm{d}}^{k+1}(f_k^+(q)) \cdot Df_k^+(q) - DS_{\mathrm{d}}^k(q) - Df_k^+(q) \cdot D^2 S_{\mathrm{d}}^{k+1}(f_k^+(q)) \cdot f_k^+(q)
$$
$$
- DS_{\mathrm{d}}^{k+1}(f_k^+(q)) \cdot Df_k^+(q) + D_1 H_{\mathrm{d}}^+\big(q, DS_{\mathrm{d}}^{k+1}(f_k^+(q))\big)
$$
$$
+ D_2 H_{\mathrm{d}}^+\big(q, DS_{\mathrm{d}}^{k+1}(f_k^+(q))\big) \cdot D^2 S_{\mathrm{d}}^{k+1}(f_k^+(q)) \cdot Df_k^+(q) = 0,
$$

which reduces to

$$
-DS_{\mathrm{d}}^k(q) + D_1 H_{\mathrm{d}}^+\big(q, DS_{\mathrm{d}}^{k+1}(f_k^+(q))\big) = 0,
$$

due to (4.5). Then, substituting $q = c_k$ gives

$$
-DS_{\mathrm{d}}^k(c_k) + D_1 H_{\mathrm{d}}^+\big(c_k, DS_{\mathrm{d}}^{k+1}(f_k^+(c_k))\big) = 0.
$$

Using (4.1) and (4.2), we obtain

(4.7) 
$$
p_k = D_1 H_{\mathrm{d}}^+(c_k, p_{k+1}).
$$

Equations (4.6) and (4.7) show that the sequence $\{(c_k, p_k)\}$ satisfies the right discrete Hamilton's equations (2.12).

Now, let us prove the latter assertion. First, recall the implicit definition of $f_k^-$ in (3.19):

(4.8) 
$$
q = -D_1 H_{\mathrm{d}}^-\big(DS_{\mathrm{d}}^k(q), f_k^-(q)\big).
$$

In particular, for $q = c_k$, we have

(4.9) 
$$
c_k = -D_1 H_{\mathrm{d}}^-(p_k, c_{k+1}),
$$

where we used (4.3) and (4.4). On the other hand, taking the derivative of (3.17) with respect to $q$ yields

$$
DS_{\mathrm{d}}^{k+1}(f_k^-(q)) \cdot Df_k^-(q) - DS_{\mathrm{d}}^k(q) + D^2 S_{\mathrm{d}}^k(q) \cdot q + DS_{\mathrm{d}}^k(q)
$$
$$
+ D_1 H_{\mathrm{d}}^-\big(DS_{\mathrm{d}}^k(q), f_k^-(q)\big) \cdot D^2 S_{\mathrm{d}}^k(q) + D_2 H_{\mathrm{d}}^-\big(DS_{\mathrm{d}}^k(q), f_k^-(q)\big) \cdot Df_k^-(q) = 0,
$$

which reduces to

$$
\big[DS_{\mathrm{d}}^{k+1}(f_k^-(q)) + D_2 H_{\mathrm{d}}^-\big(DS_{\mathrm{d}}^k(q), f_k^-(q)\big)\big] \cdot Df_k^-(q) = 0,
$$

due to (4.8). Then, substituting $q = c_k$ gives

$$
DS_{\mathrm{d}}^{k+1}(f_k^-(c_k)) = -D_2 H_{\mathrm{d}}^-\big(DS_{\mathrm{d}}^k(c_k), f_k^-(c_k)\big),
$$

since $Df_k^-(c_k)$ is invertible by assumption. Then, using (4.3) and (4.4), we obtain

(4.10) 
$$
p_{k+1} = -D_2 H_{\mathrm{d}}^-(p_k, c_{k+1}).
$$

Equations (4.9) and (4.10) show that the sequence $\{(c_k, p_k)\}$ satisfies the left discrete Hamilton's equations (2.15).    ☐

## 5. Application to discrete linear Hamiltonian systems.

### 5.1. Discrete linear Hamiltonian systems and matrix Riccati equation.

*Example* 5.1 (quadratic discrete Hamiltonian—discrete linear Hamiltonian systems). Consider a discrete Hamiltonian system on $T^*\mathbb{R}^n \cong \mathbb{R}^n \times \mathbb{R}^n$ (the configuration space is $Q = \mathbb{R}^n$) defined by the quadratic left discrete Hamiltonian

$$(5.1) \qquad H_{\mathrm{d}}^-(p_k, q_{k+1}) = \frac{1}{2}p_k^T M^{-1} p_k + p_k^T L q_{k+1} + \frac{1}{2}q_{k+1}^T K q_{k+1},$$

where $M$, $K$, and $L$ are real $n \times n$ matrices; we assume that $M$ and $L$ are invertible and also that $M$ and $K$ are symmetric. The left discrete Hamilton's equations (2.15) are

$$q_k = -(M^{-1}p_k + Lq_{k+1}),$$
$$p_{k+1} = -(L^T p_k + K q_{k+1}),$$

or

$$(5.2) \qquad \begin{pmatrix} q_{k+1} \\ p_{k+1} \end{pmatrix} = \begin{pmatrix} -L^{-1} & -L^{-1}M^{-1} \\ KL^{-1} & KL^{-1}M^{-1} - L^T \end{pmatrix} \begin{pmatrix} q_k \\ p_k \end{pmatrix},$$

and hence are a discrete linear Hamiltonian system (see section A.1).

Now, let us solve the left discrete Hamilton–Jacobi equation (3.20) for this system. For that purpose, we first generalize the problem to that with a set of initial points instead of a single initial point $(q_0, p_0)$. More specifically, consider the set of initial points that is a Lagrangian affine space $\tilde{\mathcal{L}}_{z_0}$ (see Definition A.2), which contains the point $z_0 := (q_0, p_0)$. Then, the dynamics is formally written as, for any discrete time $k \in \mathbb{N}$,

$$\tilde{\mathcal{L}}^{(k)} := (\tilde{F}_{L_{\mathrm{d}}})^k\left(\tilde{\mathcal{L}}_{z_0}\right) = \underbrace{\tilde{F}_{L_{\mathrm{d}}} \circ \cdots \circ \tilde{F}_{L_{\mathrm{d}}}}_{k}\left(\tilde{\mathcal{L}}_{z_0}\right),$$

where $\tilde{F}_{L_{\mathrm{d}}} : T^*Q \to T^*Q$ is the discrete Hamiltonian map defined in (2.8). Since $\tilde{F}_{L_{\mathrm{d}}}$ is a symplectic map, Proposition A.4 implies that $\tilde{\mathcal{L}}^{(k)}$ is a Lagrangian affine space. Then, assuming that $\tilde{\mathcal{L}}^{(k)}$ is transversal to $\{0\} \oplus Q^*$, Corollary A.6 implies that there exists a set of functions $S_{\mathrm{d}}^k$ of the form

$$(5.3) \qquad S_{\mathrm{d}}^k(q) = \frac{1}{2}q^T A_k q + b_k^T q + c_k,$$

such that $\tilde{\mathcal{L}}^{(k)} = \mathrm{graph}\, dS_{\mathrm{d}}^k$; here $A_k$ are symmetric $n \times n$ matrices, $b_k$ are elements in $\mathbb{R}^n$, and $c_k$ are in $\mathbb{R}$.

Now that we know the form of the solution, we substitute the above expression into the discrete Hamilton–Jacobi equation to find the equations for $A_k$, $b_k$, and $c_k$. Notice first that the map $f_k^-$ is given by the first half of (5.2) with $p_k$ replaced by $DS_{\mathrm{d}}^k(q)$:

$$(5.4) \qquad \begin{aligned} f_k^-(q) &= -L^{-1}\left(q + M^{-1}DS_{\mathrm{d}}^k(q)\right) \\ &= -L^{-1}(I + M^{-1}A_k)q - L^{-1}M^{-1}b_k. \end{aligned}$$

Then, substituting (5.3) into the left-hand side of the left discrete Hamilton–Jacobi equation (3.20) yields the following recurrence relations for $A_k$, $b_k$, and $c_k$:

(5.5a) $$A_{k+1} = L^T (I + A_k M^{-1})^{-1} A_k L - K,$$

(5.5b) $$b_{k+1} = -L^T (I + A_k M^{-1})^{-1} b_k,$$

(5.5c) $$c_{k+1} = c_k - \frac{1}{2} b_k^T (M + A_k)^{-1} b_k,$$

where we assumed that $I + A_k M^{-1}$ is invertible.

*Remark* 5.2. For the $A_{k+1}$ defined by (5.5a) to be symmetric, it is sufficient that $A_k$ is invertible; for if it is, then (5.5a) becomes

$$A_{k+1} = L^T (A_k^{-1} + M^{-1})^{-1} L - K,$$

and so $A_k$, $M$, and $K$ being symmetric implies that $A_{k+1}$ is as well.

*Remark* 5.3. We can rewrite (5.5a) as follows:

(5.6) $$A_{k+1} = \left[ KL^{-1} + (KL^{-1}M^{-1} - L^T)A_k \right] \left( -L^{-1} - L^{-1}M^{-1}A_k \right)^{-1}.$$

Notice the exact correspondence between the coefficients in the above equation and the matrix entries in the discrete linear Hamiltonian equations (5.2). In fact, this is the discrete Riccati equation that corresponds to the iteration defined by (5.2). See Ammar and Martin [2] for details on this correspondence.

To summarize the above observation, we have the following proposition.

PROPOSITION 5.4. *The discrete Hamilton–Jacobi equation* (3.20) *applied to the discrete linear Hamiltonian system* (5.2) *yields the discrete Riccati equation* (5.6).

In other words, the discrete Hamilton–Jacobi equation is a nonlinear generalization of the discrete Riccati equation.

**6. Relation to the Bellman equation.** In this section, we apply the above results to the optimal control setting. We will show that the (right) discrete Hamilton–Jacobi equation (3.17) gives the Bellman equation (discrete-time HJB equation) as a special case. This result gives a discrete analogue of the relationship between the H-J and HJB equations discussed in section 1.3.

**6.1. Discrete optimal control problem.** Let $q_{\mathrm{d}} := \{q_k\}_{k=0}^N$ be the state variables in a vector space $V \cong \mathbb{R}^n$ with $q_0$ and $q_N$ fixed and $u_{\mathrm{d}} := \{u_k\}_{k=0}^{N-1}$ be controls in the set $\mathcal{U} \subset \mathbb{R}^m$. With a given function $C_{\mathrm{d}} : V \times \mathcal{U} \to \mathbb{R}$, define the discrete cost functional

$$J_{\mathrm{d}} := \sum_{k=0}^{N-1} C_{\mathrm{d}}(q_k, u_k).$$

Then, we formulate the *standard discrete optimal control problem* as follows (see, e.g., [21, 9, 4, 17]).

PROBLEM 6.1 (standard discrete optimal control problem). *Minimize the discrete cost functional, i.e.,*

(6.1) $$\min_{u_{\mathrm{d}}} J_{\mathrm{d}} = \min_{u_{\mathrm{d}}} \sum_{k=0}^{N-1} C_{\mathrm{d}}(q_k, u_k),$$

*subject to the constraint,*

(6.2) $$q_{k+1} = f_{\mathrm{d}}(q_k, u_k).$$

**6.2. Necessary condition for optimality and the Bellman equation.** We would like to formulate the necessary condition for optimality. First, introduce the augmented discrete cost functional:

$$\hat{S}_{\mathrm{d}}(q_{\mathrm{d}}, p_{\mathrm{d}}, u_{\mathrm{d}}) := \sum_{k=0}^{N-1} \{C_{\mathrm{d}}(q_k, u_k) + p_{k+1} \cdot [q_{k+1} - f_{\mathrm{d}}(q_k, u_k)]\}$$

$$= \sum_{k=0}^{N-1} \left[ p_{k+1} \cdot q_{k+1} - \hat{H}_{\mathrm{d}}^+(q_k, p_{k+1}, u_k) \right],$$

where we introduced the costate $p_{\mathrm{d}} := \{p_k\}_{k=1}^N$ with $p_k \in V^*$, and also defined the *discrete control Hamiltonian*

(6.3) $$\hat{H}_{\mathrm{d}}^+(q_k, p_{k+1}, u_k) := p_{k+1} \cdot f_{\mathrm{d}}(q_k, u_k) - C_{\mathrm{d}}(q_k, u_k).$$

Then, the optimality condition (6.1) is restated as

$$\min_{q_{\mathrm{d}}, p_{\mathrm{d}}, u_{\mathrm{d}}} \hat{S}_{\mathrm{d}}(q_{\mathrm{d}}, p_{\mathrm{d}}, u_{\mathrm{d}}) = \min_{q_{\mathrm{d}}, p_{\mathrm{d}}, u_{\mathrm{d}}} \sum_{k=0}^{N-1} \left[ p_{k+1} \cdot q_{k+1} - \hat{H}_{\mathrm{d}}^+(q_k, p_{k+1}, u_k) \right].$$

In particular, extremality with respect to the control $u_{\mathrm{d}}$ implies

(6.4) $$D_3 \hat{H}_{\mathrm{d}}^+(q_k, p_{k+1}, u_k) = 0, \qquad k = 0, 1, \dots, N-1.$$

Now, we assume that $\hat{H}_{\mathrm{d}}^+$ is sufficiently regular so that this equation uniquely determines the optimal control $u_{\mathrm{d}}^* := \{u_k^*\}_{k=0}^{N-1}$; and therefore, $u_k^*$ is a function of $q_k$ and $p_{k+1}$, i.e., $u_k^* = u_k^*(q_k, p_{k+1})$. We then define

$$H_{\mathrm{d}}^+(q_k, p_{k+1}) := \max_{u_k} \hat{H}_{\mathrm{d}}^+(q_k, p_{k+1}, u_k)$$

$$= \max_{u_k}[p_{k+1} \cdot f_{\mathrm{d}}(q_k, u_k) - C_{\mathrm{d}}(q_k, u_k)]$$

(6.5) $$= p_{k+1} \cdot f_{\mathrm{d}}(q_k, u_k^*) - C_{\mathrm{d}}(q_k, u_k^*),$$

and also the *optimal discrete cost-to-go function*

$$J_{\mathrm{d}}^k(q_k) := \sum_{l=k}^{N-1} \{C_{\mathrm{d}}(q_l, u_l^*) + p_{l+1} \cdot [q_{l+1} - f_{\mathrm{d}}(q_l, u_l^*)]\}$$

$$= \sum_{l=k}^{N-1} \left[ p_{l+1} \cdot q_{l+1} - H_{\mathrm{d}}^+(q_l, p_{l+1}) \right]$$

(6.6) $$= S_{\mathrm{d}}^* - S_{\mathrm{d}}^k(q_k),$$

where $S_{\mathrm{d}}^*$ is the *optimal discrete cost functional*, i.e.,

$$S_{\mathrm{d}}^* := \hat{S}_{\mathrm{d}}(q_{\mathrm{d}}, p_{\mathrm{d}}, u_{\mathrm{d}}^*) = \sum_{k=0}^{N-1} \left[ p_{k+1} \cdot q_{k+1} - H_{\mathrm{d}}^+(q_k, p_{k+1}) \right]$$

and

$$S_{\mathrm{d}}^k(q_k) := \sum_{l=0}^{k-1} \left[ p_{l+1} \cdot q_{l+1} - H_{\mathrm{d}}^+(q_l, p_{l+1}) \right].$$

The above action sum has exactly the same form as (3.8) formulated in the framework of discrete Hamiltonian mechanics. Therefore, our theory now directly applies to this case: The corresponding right discrete Hamilton's equations (2.12) are, using the expression in (6.5),

$$
\begin{aligned}
q_{k+1} &= f_{\mathrm{d}}(q_k, u_k^*), \\
p_k &= p_{k+1} \cdot D_1 f_{\mathrm{d}}(q_k, u_k^*) - D_1 C_{\mathrm{d}}(q_k, u_k^*).
\end{aligned}
$$

Therefore, (3.16) gives the implicit definition of $f_k^+$ as follows:

$$
(6.7) \qquad f_k^+(q_k) = f_{\mathrm{d}}\big(q_k, u_k^*\big(q_k, DS_{\mathrm{d}}^{k+1}(f_k^+(q_k))\big)\big).
$$

Hence, the (right) discrete Hamilton–Jacobi equation (3.17) applied to this case gives

$$
S_{\mathrm{d}}^{k+1}(f_{\mathrm{d}}(q_k, u_k^*)) - S_{\mathrm{d}}^k(q_k) - DS_{\mathrm{d}}^{k+1}(f_{\mathrm{d}}(q_k, u_k^*)) \cdot f_{\mathrm{d}}(q_k, u_k^*) \\
+ H_{\mathrm{d}}^+\big(q_k, DS_{\mathrm{d}}^{k+1}(f_{\mathrm{d}}(q_k, u_k^*))\big) = 0,
$$

and again using the expression for the Hamiltonian in (6.5), this becomes

$$
\max_{u_k}\big[S_{\mathrm{d}}^{k+1}(f_{\mathrm{d}}(q_k, u_k)) - C_{\mathrm{d}}(q_k, u_k)\big] = S_{\mathrm{d}}^k(q_k).
$$

Since $S_{\mathrm{d}}^k(q_k) = S_{\mathrm{d}}^* - J_{\mathrm{d}}^k(q_k)$, we obtain

$$
(6.8) \qquad \min_{u_k}\big[J_{\mathrm{d}}^{k+1}(f_{\mathrm{d}}(q_k, u_k)) + C_{\mathrm{d}}(q_k, u_k)\big] = J_{\mathrm{d}}^k(q_k),
$$

which is the *Bellman equation* (see, e.g., Bellman [4, Chapter 2], [5] and Bertsekas [6, Chapter 1]).

*Remark* 6.2. Notice that the Bellman equation (6.8) is much simpler than the discrete Hamilton–Jacobi equations (3.17) and (3.20) because of the special form of the control Hamiltonian equation (6.5). Also, notice that, as shown in (6.7), the term $f_k^+(q_k)$ is written in terms of the given function $f_{\mathrm{d}}$. See Remark 3.6 for comparison.

**6.3. Relation between the discrete H-J and Bellman equations and its consequences.** Summarizing the observation made above, we have the following proposition.

PROPOSITION 6.3. *The right discrete Hamilton–Jacobi equation* (3.17) *applied to the Hamiltonian formulation of the standard discrete optimal control problem* 6.1 *gives the Bellman equation* (6.8).

This observation leads to the following well-known fact.

PROPOSITION 6.4. *Let* $J_{\mathrm{d}}^k(q_k)$ *be a solution to the Bellman equation* (6.8). *Then, the costate* $p_k$ *in the discrete maximum principle is given as follows:*

$$
p_k = -DJ_{\mathrm{d}}^k(c_k),
$$

*where* $c_{k+1} = f_{\mathrm{d}}(c_k, u_k^*)$ *with the optimal control* $u_k^*$.

*Proof.* The proposition follows from a reinterpretation of Theorem 4.1 through Proposition 6.3 with the relation $S_{\mathrm{d}}^k(q_k) = S_{\mathrm{d}}^* - J_{\mathrm{d}}^k(q_k)$.   $\square$

**7. Generalized Bellman equation with internal-stage controls.** In the previous section, we showed that the discrete Hamilton–Jacobi equation recovers the Bellman equation if we apply our theory to the Hamiltonian formulation of the standard discrete optimal control problem 6.1. In this section, we generalize the approach

to derive what may be considered as higher-order discrete-time approximations of the HJB equation (1.4). Namely, we derive a class of discrete control Hamiltonians that use higher-order approximations (a more general version of (6.3)) by employing the technique of Galerkin Hamiltonian variational integrators introduced by Leok and Zhang [28]; and then, we apply our theory to obtain a class of generalized Bellman equations that have controls at internal stages.

**7.1. Continuous-time optimal control problem.** Let us first briefly review the standard formulation of continuous-time optimal control problems. Let $q$ be the state variable in a vector space $V \cong \mathbb{R}^n$, $q_0$ and $q_T$ fixed in $V$, and $u$ be the control in the set $\mathcal{U} \subset \mathbb{R}^m$. With a given function $C : V \times \mathcal{U} \to \mathbb{R}$, define the cost functional

$$J := \int_0^T C(q(t), u(t)) \, dt.$$

Then, we formulate the *standard continuous-time optimal control problem* as follows.

PROBLEM 7.1 (standard continuous-time optimal control problem).  *Minimize the cost functional, i.e.,*

$$\min_{u(\cdot)} J = \min_{u(\cdot)} \int_0^T C(q(t), u(t)) \, dt,$$

*subject to the constraints,*

$$\dot{q} = f(q, u),$$

*and $q(0) = q_0$ and $q(T) = q_T$.*

A Hamiltonian structure comes into play with the introduction of the augmented cost functional:

$$\hat{S} := \int_0^T \left\{ C(q(t), u(t)) + p(t)[\dot{q}(t) - f(q(t), u(t))] \right\} dt$$

$$= \int_0^T \left[ p(t)\dot{q}(t) - \hat{H}(q(t), p(t), u(t)) \right] dt,$$

where we introduced the costate $p(t) \in V^*$, and also defined the *control Hamiltonian*,

(7.1) $$\hat{H}(q, p, u) := p \cdot f(q, u) - C(q, u).$$

**7.2. Galerkin Hamiltonian variational integrator.** Recall, from [28, Section 2.2], that the exact right discrete Hamiltonian is a type-two generating function for the original continuous-time Hamiltonian flow, defined by

(7.2) $$H_{\mathrm{d,ex}}^+(q_0, p_1) = \operatorname*{ext}_{\substack{(q,p)\in\mathcal{C}^1([0,h],T^*Q) \\ q(0)=q_0, p(h)=p_1}} \left\{ p_1 q_1 - \int_0^h [p(t)\dot{q}(t) - H(q(t), p(t))] \, dt \right\},$$

where $h$ is the time step; $\mathcal{C}^1([0,h], T^*Q)$ is the set of continuously differentiable curves on $T^*Q$ over the time interval $[0, h]$; an extremum is achieved for the exact solution of Hamilton's equations (1.1) that satisfy the specified boundary conditions. Therefore, it requires the exact solution $(q(t), p(t))$ to evaluate the the above integral, and so the exact discrete Hamiltonian cannot be practically computed in general.

The key idea of Galerkin Hamiltonian variational integrators [28] is to replace the set of curves $\mathcal{C}^1([0,h], T^*Q)$ by a certain finite-dimensional space so as to obtain a computable expression for a discrete Hamiltonian.

**7.3. Galerkin discrete control Hamiltonian.** Here, we would like to apply the above idea to the *control* Hamiltonian (7.1) to obtain a discrete control Hamiltonian.

Let $\mathcal{C}_d^s(V)$ be a finite-dimensional space of curves defined by

$$\mathcal{C}_d^s(V) := \left\{ t \mapsto \sum_{i=1}^{s} w^i \, \psi_i(t/h) \mid t \in [0, h], \ w^i \in V \text{ for each } i \in \{1, 2, \dots, s\} \right\},$$

with the basis functions $\{\psi_i : [0, 1] \to \mathbb{R}\}_{i=1}^{s}$.

1. Use the basis functions $\psi_i$ to approximate the velocity $\dot{q}$ over the interval $[0, h]$,

$$\dot{q}(\tau h) \approx \dot{q}_d(\tau h) = \sum_{i=1}^{s} w^i \psi_i(\tau),$$

where $\tau \in [0, 1]$ and $w^i \in V$ for each $i = 1, \dots, s$.

2. Integrate $\dot{q}_d(t)$ over $[0, \tau h]$, to obtain the approximation for the position $q$, i.e.,

$$q_d(\tau h) = q_d(0) + \int_0^{\tau h} \sum_{i=1}^{s} w^i \psi_i(t/h) \, dt = q_0 + h \sum_{i=1}^{s} w^i \int_0^{\tau} \psi_i(\rho) \, d\rho,$$

where we applied the boundary condition $q_d(0) = q_0$. Applying the boundary condition $q_d(h) = q_1$ at the other endpoint yields

$$q_1 = q_d(h) = q_0 + h \sum_{i=1}^{s} w^i \int_0^1 \psi_i(\rho) \, d\rho = q_0 + h \sum_{i=1}^{s} B_i w^i,$$

where $B_i := \int_0^1 \psi_i(\tau) \, d\tau$. Furthermore, we introduce the internal stages,

$$(7.3) \qquad Q^i(w) := q_d(c^i h) = q_0 + h \sum_{j=1}^{s} w^j \int_0^{c^i} \psi_j(\tau) \, d\tau = q_0 + h \sum_{j=1}^{s} A_j^i w^j,$$

where $A_j^i := \int_0^{c^i} \psi_j(\tau) \, d\tau$.

3. The exact discrete control Hamiltonian $\hat{H}_{d,ex}^+$ is defined as in (7.2):

$$\hat{H}_{d,ex}^+(q_0, p_1; u(\cdot)) := \operatorname*{ext}_{\substack{(q,p) \in \mathcal{C}^1([0,h], V \times V^*) \\ q(t_0) = q_0, \, p(t_1) = p_1}} \left\{ p_1 q_1 - \int_0^h \left[ p\dot{q} - \hat{H}(q, p, u) \right] dt \right\}.$$

Again this is practically not computable, and so we employ the following approximation: Use the numerical quadrature formula

$$\int_0^1 f(\rho) \, d\rho \approx \sum_{i=1}^{s} b_i \, f(c^i)$$

with constants $(b_i, c^i)$ and the finite-dimensional function space $\mathcal{C}_d^s(V)$ to construct the discrete control Hamiltonian $\hat{H}_d^+(q_0, p_1, U)$ as follows:

$$\hat{H}_d^+(q_0, p_1, U) := \operatorname*{ext}_{\substack{\dot{q}_d \in \mathcal{C}_d^s(V) \\ P^i \in V^*}} \left\{ p_1 q_d(h) - h \sum_{i=1}^{s} b_i \left[ p(c^i h)\dot{q}_d(c^i h) - \hat{H}(Q^i(w), P^i, U^i) \right] \right\}$$

$$= \operatorname*{ext}_{w, P} K(q_0, w, P, U, p_1),$$

where we set $P^i := p(c^i h)$ and $U^i := u(c^i h)$ and defined

$$K(q_0, w, P, U, p_1) := p_1 \cdot \left( q_0 + h \sum_{i=1}^{s} B_i w^i \right)$$

$$- h \sum_{i=1}^{s} b_i \left[ P^i \cdot \sum_{j=1}^{s} w^j \psi_j(c^i) - \hat{H} \left( Q^i(w), P^i, U^i \right) \right]$$

$$= p_1 \cdot \left( q_0 + h \sum_{i=1}^{s} B_i w^i \right)$$

$$- h \sum_{i=1}^{s} b_i \left\{ P^i \cdot \left[ \sum_{j=1}^{s} M_j^i w^j - f(Q^i(w), U^i) \right] + C(Q^i(w), U^i) \right\},$$

where we defined $M_j^i := \psi_j(c^i)$ and used the expression for the control Hamiltonian in (7.1); note that $P^i \in V^*$ and $w^i \in V$ for each $i = 1, \ldots, s$, and that $f$ takes values in $V$. In order to obtain an expression for $H_\mathrm{d}^+(q_0, p_1, U)$, we first compute the stationarity conditions for $K(q_0, w, P, U, p_1)$ under the fixed boundary condition $(q_0, p_1)$:

(7.4a)

$$0 = \frac{\partial K(q_0, w, P, U, p_1)}{\partial w^j} = hp_1 \cdot B_j - h \sum_{i=1}^{s} b_i \left[ M_j^i P^i - h A_j^i D_1 \hat{H} \left( q^i(w), P^i, U^i \right) \right],$$

(7.4b)

$$0 = \frac{\partial K(q_0, w, P, U, p_1)}{\partial P^i} = -h b_i \left[ \sum_{j=1}^{s} M_j^i w^j - f(Q^i(w), U^i) \right],$$

for $j = 1, \ldots, s$.

4. By solving the $2ns$ stationarity conditions (7.4), we can express the parameters $w$ and $P$ in terms of $q_0$, $p_1$, and $U$, i.e., $w = \tilde{w}(q_0, U)$ and $P = \tilde{P}(q_0, p_1, U)$: In particular, assuming $b_j \neq 0$ for each $j = 1, \ldots, s$, (7.4b) gives $w^j M_j^i = f(Q^i(w), U^i)$; this gives a set of $ns$ nonlinear equations[2] satisfied by $w = \tilde{w}(q_0, U)$.[3] Therefore, we have

$$K(q_0, \tilde{w}(q_0, U), P, U, p_1) = p_1 \cdot \left[ q_0 + h \sum_{j=1}^{s} B_j \, \tilde{w}^j(q_0, U) \right] - h \sum_{i=1}^{s} b_i \, C\big( Q^i(\tilde{w}(q_0, U)), U^i \big).$$

Notice that the internal-stage momenta $P^i$ disappear when we substitute $w = \tilde{w}(q_0, U)$. Therefore, we obtain the following Galerkin discrete control Hamiltonian:

$$\hat{H}_\mathrm{d}^+(q_0, p_1, U) := K \Big( q_0, \tilde{w}(q_0, U), \tilde{P}(q_0, p_1, U), U, p_1 \Big)$$

(7.5)

$$= p_1 \cdot \left[ q_0 + h \sum_{j=1}^{s} B_j \, \tilde{w}^j(q_0, U) \right] - h \sum_{i=1}^{s} b_i \, C\big( Q^i(\tilde{w}(q_0, U)), U^i \big).$$

---

[2]Recall that $w^j \in V$ and $f$ takes values in $V$.

[3]Note from (7.3) that $Q^i$ is written in terms of $q_0$ and $w$.
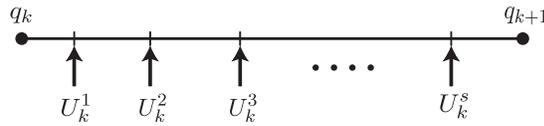
FIG. 2. *Internal-stage controls $\{U_k^i\}_{i=1}^s$ in the discrete-time intervals between $k$ and $k+1$.*

**7.4. The Bellman equation with internal-stage controls.** The Galerkin discrete control Hamiltonian (7.5) gives

$$(7.6) \quad \hat{H}_d^+(q_k, p_{k+1}, U_k^1, \ldots, U_k^s) := p_{k+1} \cdot f_d(q_k, U_k^1, \ldots, U_k^s) - C_d(q_k, U_k^1, \ldots, U_k^s),$$

with

$$(7.7) \qquad f_d(q_k, U_k^1, \ldots, U_k^s) := q_k + h \sum_{i=1}^s \tilde{w}^i(q_k, U_k^1, \ldots, U_k^s)\, B_i,$$

and

$$(7.8) \qquad C_d(q_k, U_k^1, \ldots, U_k^s) := h \sum_{i=1}^s b_i\, C\big(Q^i(\tilde{w}(q_k, U_k^1, \ldots, U_k^s)), U_k^i\big).$$

This is a generalized version of (6.3) with internal-stage controls $\{U_k^i\}_{i=1}^s$ as opposed to a single control $u_k$ per time step (see Figure 2). Now assume that

$$(7.9) \qquad \frac{\partial}{\partial U_k^i} \hat{H}_d^+(q_k, p_{k+1}, U_k^1, \ldots, U_k^s) = 0, \qquad i = 1, \ldots, s,$$

is solvable for $\{U_k^i\}_{i=1}^s$ to give the optimal internal-stage controls $\{U_k^{*,i}\}_{i=1}^s$. Then, we may apply the same argument as in section 6: In particular, the right discrete Hamilton–Jacobi equation (3.17) applied to this case gives the following *Bellman equation with internal-stage controls*:

$$(7.10) \qquad \min_{U_k^1, \ldots, U_k^s} \big[ J_d^{k+1}(f_d(q_k, U_k^1, \ldots, U_k^s)) + C_d(q_k, U_k^1, \ldots, U_k^s) \big] = J_d^k(q_k).$$

The following example shows that the standard Bellman equation (6.8) follows as a special case.

*Example* 7.2 (the standard Bellman equation). Let $s = 1$, and select

$$\psi_1(\tau) = 1, \qquad b_1 = 1, \qquad c^1 = 0.$$

Then, we have $B_1 = 1$, $A_1^1 = 0$, and $M_1^1 = 1$. Hence, (7.3) gives $Q^1 = q_k$ (we set the endpoints $(q_0, q_1)$ to be $(q_k, q_{k+1})$ here), and (7.4b) gives

$$w^1 = f(q_k, U_k^1).$$

However, the control $U_k^1$ is defined as follows (we shift the time intervals from $[0, h]$ to $[t_k, t_k + h]$ here):

$$U_k^1 := u(t_k + c^1 h) = u(t_k) =: u_k.$$

So, we have $w^1 = f(q_k, u_k)$, and thus, (7.7) and (7.8) give

$$f_{\mathrm{d}}(q_k, u_k) = q_k + h\, f(q_k, u_k)$$

and

$$C_{\mathrm{d}}(q_k, u_k) = h\, C(q_k, u_k),$$

respectively. Notice that this approximation gives the forward-Euler discretization of the standard continuous-time optimal control problem 7.1 to yield the standard discrete optimal control problem 6.1. In fact, the Bellman equation with internal-stage controls, (7.10), reduces to the standard Bellman equation (6.8):

$$(7.11) \qquad \min_{u_k}\left[J_{\mathrm{d}}^{k+1}(f_{\mathrm{d}}(q_k, u_k)) + C_{\mathrm{d}}(q_k, u_k)\right] = J_{\mathrm{d}}^k(q_k).$$

*Remark* 7.3. Higher-order approximations with any number of internal stages $s$ are possible as long as (7.4b) is solvable for $w$. See [28] for various different choices of discretizations.

**7.5. Application to the Heisenberg system.** Let us now apply the above results to a simple optimal control problem to illustrate the result.

*Example* 7.4 (the Heisenberg system; see, e.g., Brockett [8] and Bloch [7]). Consider the following optimal control problem: For a fixed time $T > 0$,

$$\min_{u(\cdot),\, v(\cdot)} \int_0^T \frac{1}{2}(u^2 + v^2)\, dt,$$

subject to the constraint,

$$\dot{x} = u, \qquad \dot{y} = v, \qquad \dot{z} = uy - vx.$$

This is the standard continuous-time optimal control problem 7.1 with $V = \mathbb{R}^3$, $\mathcal{U} = \mathbb{R}^2$, $q = (x, y, z)$, and

$$f(x, y, z, u, v) = \begin{pmatrix} u \\ v \\ uy - vx \end{pmatrix}, \qquad C(x, y, z, u, v) = \frac{1}{2}(u^2 + v^2).$$

If we apply the choice of the discretization in Example 7.2, we have the standard Bellman equation

$$\min_{u_k, v_k}\left[S_{\mathrm{d}}^{k+1}(f_{\mathrm{d}}(q_k, u_k, v_k)) - C_{\mathrm{d}}(q_k, u_k, v_k)\right] - S_{\mathrm{d}}^k(q_k) = 0,$$

with $q_k := (x_k, y_k, z_k)$, where

$$f_{\mathrm{d}}(q_k, u_k, v_k) = \begin{pmatrix} x_k + h\, u_k \\ y_k + h\, v_k \\ z_k + h\,(u_k y_k - v_k x_k) \end{pmatrix}, \qquad C_{\mathrm{d}}(q_k, u_k, v_k) = \frac{h}{2}(u_k^2 + v_k^2).$$

Now, if we choose $s = 2$, and select

$$(\psi_1(\tau), \psi_2(\tau)) = (1, \cos(\pi\tau)), \qquad b = (b_1, b_2) = \left(\frac{1}{2}, \frac{1}{2}\right), \qquad c = (c^1, c^2) = (0, 1).$$

Then, we have

$$B = (B_1, B_2) = (1, 0), \qquad A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad M = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Leok and Zhang [28, Example 4.4] show that this choice of discretization corresponds to the Störmer–Verlet method (see, e.g., Marsden and West [30]). The Bellman equation with internal-stage controls, (7.10), then becomes

$$\min_{u_k^1, v_k^1, u_k^2, v_k^2} \left[ J_{\mathrm{d}}^{k+1} \left( f_{\mathrm{d}} \left( q_k, u_k^1, v_k^1, u_k^2, v_k^2 \right) \right) + C_{\mathrm{d}} \left( q_k, u_k^1, v_k^1, u_k^2, v_k^2 \right) \right] = J_{\mathrm{d}}^k(q_k),$$

where

$$f_{\mathrm{d}} \left( q_k, u_k^1, v_k^1, u_k^2, v_k^2 \right) = \begin{pmatrix} x_k + h \left( u_k^1 + u_k^2 \right)/2 \\ y_k + h \left( v_k^1 + v_k^2 \right)/2 \\ z_k + h \left( \dfrac{u_k^1 + u_k^2}{2} \, y_k - \dfrac{v_k^1 + v_k^2}{2} \, x_k + \dfrac{u_k^2 v_k^1 - u_k^1 v_k^2}{4} \right) \end{pmatrix}$$

and

$$C_{\mathrm{d}} \left( q_k, u_k^1, v_k^1, u_k^2, v_k^2 \right) = \frac{h}{2} \left[ \frac{(u_k^1)^2 + (u_k^2)^2}{2} + \frac{(v_k^1)^2 + (v_k^2)^2}{2} \right].$$

**8. Conclusion and future work.** We developed a discrete-time analogue of Hamilton–Jacobi theory starting from the discrete variational Hamiltonian mechanics formulated by Lall and West [23]. We reinterpreted and extended the discrete Hamilton–Jacobi equation given by Elnatanov and Schiff [13] in the language of discrete mechanics. Furthermore, we showed that the discrete Hamilton–Jacobi equation reduces to the discrete Riccati equation with a quadratic Hamiltonian, and also that it specializes to the Bellman equation of dynamic programming if applied to standard discrete optimal control problems. These results are discrete analogues of the corresponding known results in the continuous-time theory. Application to discrete optimal control also revealed that the discrete Hamilton–Jacobi theorem 4.1 specializes to a well-known result in discrete optimal control theory. We also used a Galerkin-type approximation to derive Galerkin discrete control Hamiltonians. This technique gave an explicit formula for discrete control Hamiltonians in terms of the constructs in the original continuous-time optimal control problem. By viewing the Bellman equation as a special case of the discrete Hamilton–Jacobi equation, we could introduce the discretization technique for discrete Hamiltonian mechanics into the discrete optimal control setting; this lead us to a class of Bellman equations with controls at internal stages.

We are interested in the following topics for future work:
- *Application to integrable discrete systems*: Theorem 4.1 gives a discrete analogue of the theory behind the technique of solution by separation of variables, i.e., the theorem relates a solution of the discrete Hamilton–Jacobi equations with that of the discrete Hamilton's equations. An interesting question then is whether or not separation of variables applies to integrable discrete systems, e.g., discrete rigid bodies of Moser and Veselov [31] and various others discussed by Suris [34, 35].
- *Development of numerical methods based on the discrete Hamilton–Jacobi equation*: Hamilton–Jacobi equation has been used to develop structured integrators for Hamiltonian systems. Ge and Marsden [14] developed a numerical

method that preserves momentum maps and Poisson brackets of Lie–Poisson systems by solving the Lie–Poisson Hamilton–Jacobi equation approximately. See also Channell and Scovel [11] (and references therein) for a survey of structured integrators based on the Hamilton–Jacobi equation. The present theory, being inherently discrete in time, potentially provides a variant of such numerical methods.

- *Extension to discrete nonholonomic and Dirac mechanics*: The present work is concerned only with unconstrained systems. Extensions to nonholonomic and Dirac mechanics, more specifically discrete-time versions of the nonholonomic Hamilton–Jacobi theory [19, 12, 32, 10, 33] and Dirac Hamilton–Jacobi theory [27], are another direction for future research.

- *Relation to the power method and iterations on the Grassmannian manifold*: Ammar and Martin [2] established links between the power method, iterations on the Grassmannian manifold, and the Riccati equation. The discussion on iterations of Lagrangian subspaces and its relation to the Riccati equation in sections 5.1 and A.2 is a special case of such links. On the other hand, Proposition 5.4 suggests that the discrete Hamilton–Jacobi equation is a generalization of the Riccati equation. We are interested in exploring possible further links implied by the generalization.

- *Galerkin discrete optimal control problems*: The Galerkin discrete control Hamiltonians may be considered to be a means of formulating discrete optimal control problems with higher-order of approximation to a continuous-time optimal control problem. This idea generalizes the Runge–Kutta discretizations of optimal control problems (see, e.g., Hager [18] and references therein). In fact, Leok and Zhang [28] showed that their method recovers the symplectic-partitioned Runge–Kutta method. Therefore, this approach is expected to provide higher-order structure-preserving numerical methods for optimal control problems.

## Appendix A. Discrete linear Hamiltonian systems.

**A.1. Discrete linear Hamiltonian systems.** Suppose that the configuration space $Q$ is an $n$-dimensional vector space, and that the discrete Hamiltonian $H_{\mathrm{d}}^+$ or $H_{\mathrm{d}}^-$ is quadratic as in (5.1). Also assume that the corresponding discrete Hamiltonian map $\tilde{F}_{L_{\mathrm{d}}} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$ is invertible. Then, the discrete Hamilton's equations (2.12) or (2.15) reduce to the discrete linear Hamiltonian system

$$(\mathrm{A.1}) \qquad z_{k+1} = A_{L_{\mathrm{d}}} z_k,$$

where $z_k \in \mathbb{R}^{2n}$ is a coordinate expression for $(q_k, p_k) \in Q \oplus Q^*$ and $A_{L_{\mathrm{d}}} : Q \oplus Q^* \to Q \oplus Q^*$ is the matrix representation of the map $\tilde{F}_{L_{\mathrm{d}}}$ in the same basis. Since $\tilde{F}_{L_{\mathrm{d}}}$ is symplectic, it follows that $A_{L_{\mathrm{d}}}$ is an $2n \times 2n$ symplectic matrix, i.e.,

$$(\mathrm{A.2}) \qquad A_{L_{\mathrm{d}}}^T \mathbb{J} A_{L_{\mathrm{d}}} = \mathbb{J},$$

where the matrix $\mathbb{J}$ is defined by

$$\mathbb{J} := \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

with $I$ the $n \times n$ identity matrix.

**A.2. Lagrangian subspaces and Lagrangian affine spaces.** Let us recall the definition of a Lagrangian subspace.

DEFINITION A.1. *Let $V$ be a symplectic vector space with a symplectic form $\Omega$. A subspace $\mathcal{L}$ of $V$ is said to be* Lagrangian *if $\Omega(v, w) = 0$ for any $v, w \in \mathcal{L}$ and $\dim \mathcal{L} = \dim V/2$.*

We introduce the following definition for later convenience.

DEFINITION A.2. *A subset $\tilde{\mathcal{L}}_b$ of a symplectic vector space $V$ is called a* Lagrangian affine space *if $\tilde{\mathcal{L}}_b = b + \mathcal{L}$ for some element $b \in V$ and a Lagrangian subspace $\mathcal{L} \subset V$.*

The following fact is well known (see, e.g., Jurdjevic [22, Theorem 6 on p. 417]).

PROPOSITION A.3. *Let $\mathcal{L}$ be a Lagrangian subspace of $V$ and $\Phi : V \to V$ be a symplectic transformation. Then, for any $k \in \mathbb{N}$, the image of $\mathcal{L}$ under the $k$-fold composition of $\Phi$, i.e.,*

$$\Phi^k(\mathcal{L}) = \underbrace{\Phi \circ \cdots \circ \Phi}_{k}(\mathcal{L}),$$

*is also a Lagrangian subspace of $V$.*

A similar result holds for Lagrangian affine spaces.

PROPOSITION A.4. *Let $\tilde{\mathcal{L}}_b = b + \mathcal{L}$ be a Lagrangian affine space of $V$ and $\Phi : V \to V$ be a symplectic transformation. Then $\Phi^k(\tilde{\mathcal{L}}_b)$ is also a Lagrangian affine space of $V$ for any $k \in \mathbb{N}$. More explicitly, we have*

$$\Phi^k(\tilde{\mathcal{L}}_b) = \Phi^k b + \Phi^k(\mathcal{L}).$$

*Proof.* The proposition follows from a straightforward calculation.   □

**A.3. Generating functions.** Now, consider the case where $V = Q \oplus Q^*$ to apply the results from section A.2 to the setting in section A.1. This is a symplectic vector space with the symplectic form $\Omega : (Q \oplus Q^*) \times (Q \oplus Q^*) \to \mathbb{R}$ defined by

$$\Omega : (v, w) \mapsto v^T \mathbb{J} w.$$

The key result here regarding Lagrangian subspaces on $Q \oplus Q^*$ is the following.

PROPOSITION A.5. *A Lagrangian subspace of $Q \oplus Q^*$ that is transversal to $\{0\} \oplus Q^*$ is the graph of an exact one-form on $Q$, i.e., $\mathcal{L} = \operatorname{graph} dS$ for some function $S : Q \to \mathbb{R}$, which has the form*

(A.3)
$$S(q) = \frac{1}{2} \langle Aq, q \rangle + C$$

*with some symmetric linear map $A : Q \to Q^*$ and an arbitrary real scalar constant $C$. Moreover, the correspondence between the Lagrangian subspaces and such functions (modulo the constant term) is one-to-one.*

*Proof.* First, recall that a Lagrangian submanifold of $T^*Q$ that projects diffeomorphically onto $Q$ is the graph of a closed one-form on $Q$ (see Abraham and Marsden [1, Proposition 5.3.15 and the subsequent paragraph on p. 410]). In our case, $Q$ is a vector space, and so the cotangent bundle $T^*Q$ is identified with the direct sum $Q \oplus Q^*$. Now, a Lagrangian subspace of $Q \oplus Q^*$ that is transversal to $\{0\} \oplus Q^*$ projects diffeomorphically onto $Q$, therefore it is the graph of a closed one-form. Then, by the Poincaré lemma, it follows that any such Lagrangian subspace $\mathcal{L}$ is identified with the graph of an exact one-form $dS$ with some function $S$ on $Q$, i.e., $\mathcal{L} = \operatorname{graph} dS$.

However, as shown in, e.g., Jurdjevic [22, Theorem 3 on p. 233], the space of Lagrangian subspaces that are transversal to $\{0\} \oplus Q^*$ is in one-to-one correspondence with the space of all symmetric maps $A : Q \to Q^*$, with the correspondence given by $\mathcal{L} = \mathrm{graph}\, A$. Hence, $\mathrm{graph}\, dS = \mathrm{graph}\, A$, or more specifically,

$$dS(q) = A_{ij} q^j \, dq^i.$$

This implies that $S$ has the form

$$S(q) = \frac{1}{2} A_{ij} q^i q^j + C,$$

with an arbitrary real scalar constant $C$. □

COROLLARY A.6. *Let $\tilde{\mathcal{L}}_{z_0} = z_0 + \mathcal{L}$ be a Lagrangian affine space, where $z_0 = (q_0, p_0)$ is an element in $Q \oplus Q^*$ and $\mathcal{L}$ is a Lagrangian subspace of $Q \oplus Q^*$ that is transversal to $\{0\} \oplus Q^*$. Then, $\tilde{\mathcal{L}}_{z_0}$ is the graph of an exact one-form $d\tilde{S}$ with a function $\tilde{S} : Q \to \mathbb{R}$ of the form*

(A.4) $$\tilde{S}(q) = \frac{1}{2} \langle Aq, q \rangle + \langle p_0 - Aq_0, q \rangle + C,$$

*with some symmetric linear map $A : Q \to Q^*$ and an arbitrary real scalar constant $C$.*

*Proof.* From Proposition A.5, there exists a function $S : Q \to \mathbb{R}$ of the form equation (A.3) with some symmetric linear map $A : Q \to Q^*$ such that $\mathcal{L} = \mathrm{graph}\, dS = \mathrm{graph}\, A$. Let us define $\tilde{S} : Q \to R$ by

$$\tilde{S}(q) := S(q - q_0) + \langle p_0, q \rangle,$$

from which (A.4) follows by direct calculation. Then,

$$d\tilde{S}(q) = A(q - q_0) + p_0,$$

and thus

$$\begin{aligned}
\mathrm{graph}\, d\tilde{S} &= \{(q, d\tilde{S}(q)) \mid q \in Q\} \\
&= \{(q, A(q - q_0) + p_0) \mid q \in Q\} \\
&= (q_0, p_0) + \{(q - q_0, A(q - q_0)) \mid q \in Q\} \\
&= z_0 + \mathrm{graph}\, A \\
&= z_0 + \mathcal{L} \\
&= \tilde{\mathcal{L}}_{z_0}. \quad □
\end{aligned}$$

REFERENCES

[1] R. ABRAHAM AND J. E. MARSDEN, *Foundations of Mechanics*, 2nd ed., Addison-Wesley, Reading, MA, 1978.
[2] G. AMMAR AND C. MARTIN, *The geometry of matrix eigenvalue methods*, Acta Appl. Math., 5 (1986), pp. 239–278.
[3] V. I. ARNOLD, *Mathematical Methods of Classical Mechanics*, Springer, New York, 1989.

[4]   R. Bellman, *Introduction to the Mathematical Theory of Control Processes*, Vol. 2, Academic Press, New York, 1971.

[5]   R. Bellman, *Dynamic Programming*, Princeton University Press, Princeton, NJ, 1972.

[6]   D. P. Bertsekas, *Dynamic Programming and Optimal Control*, Vol. 1, Athena Scientific, Nashua, NH, 2005.

[7]   A. M. Bloch, *Nonholonomic Mechanics and Control*, Springer, New York, 2003.

[8]   R. W. Brockett, *Control theory and singular Riemannian geometry*, in New Directions in Applied Mathematics, P. J. Hilton and G. S. Young, eds., Springer, New York, 1981, pp. 11–27.

[9]   J. A. Cadzow, *Discrete calculus of variations*, Internat. J. Control, 11 (1970), pp. 393–407.

[10]  J. F. Cariñena, X. Gracia, G. Marmo, E. Martínez, M. C. Munõz Lecanda, and N. Román-Roy, *Geometric Hamilton–Jacobi theory for nonholonomic dynamical systems*, Int. J. Geom. Methods Mod. Phys., 7 (2010), pp. 431–454.

[11]  P. J. Channell and C. Scovel, *Symplectic integration of Hamiltonian systems*, Nonlinearity, 3 (1990), pp. 231–259.

[12]  M. de León, J. C. Marrero, and D. Martín de Diego, *Linear almost Poisson structures and Hamilton–Jacobi equation. Applications to nonholonomic mechanics*, J. Geom. Mech., 2 (2010), pp. 159–198.

[13]  N. A. Elnatanov and J. Schiff, *The Hamilton–Jacobi difference equation*, Funct. Differ. Equ., 3 (1996), pp. 279–286.

[14]  Z. Ge and J. E. Marsden, *Lie–Poisson Hamilton–Jacobi theory and Lie–Poisson integrators*, Phys. Lett. A, 133 (1988), pp. 134–139.

[15]  I. M. Gelfand and S. V. Fomin, *Calculus of Variations*, Dover, New York, 2000.

[16]  H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*, 3rd ed., Addison-Wesley, Reading, MA, 2001.

[17]  V. Guibout and A. M. Bloch, *A discrete maximum principle for solving optimal control problems*, in Proceedings of the 43rd IEEE Conference on Decision and Control, 2004, pp. 1806–1811.

[18]  W. W. Hager, *Runge–Kutta methods in optimal control and the transformed adjoint system*, Numer. Math., 87 (2000), pp. 247–282.

[19]  D. Iglesias-Ponte, M. de León, and D. Martín de Diego, *Towards a Hamilton–Jacobi theory for nonholonomic mechanical systems*, J. Phys. A, 41 (2008), article 015205.

[20]  S. M. Jalnapurkar, M. Leok, J. E. Marsden, and M. West, *Discrete Routh reduction*, J. Phys. A, 39 (2006), pp. 5521–5544.

[21]  B. W. Jordan and E. Polak, *Theory of a class of discrete optimal control systems*, J. Electron. Control, 17 (1964), pp. 694–711.

[22]  V. Jurdjevic, *Geometric Control Theory*, Cambridge University Press, Cambridge, U.K., 1997.

[23]  S. Lall and M. West, *Discrete variational Hamiltonian mechanics*, J. Phys. A, 39 (2006), pp. 5509–5519.

[24]  C. Lanczos, *The Variational Principles of Mechanics*, 4th ed., Dover, New York, 1986.

[25]  M. Leok, *Foundations of Computational Geometric Mechanics*, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 2004.

[26]  M. Leok, J. E. Marsden, and A. Weinstein, *A discrete theory of connections on principal bundles*, Preprint, 2004; available online from http://arxiv.org/abs/math/0508338.

[27]  M. Leok, T. Ohsawa, and D. Sosa, *Hamilton–Jacobi theory for degenerate Lagrangian systems with constraints*, in preparation.

[28]  M. Leok and J. Zhang, *Discrete Hamiltonian variational integrators*, IMA J. Numer. Anal., 2011.

[29]  J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*, Springer, New York, 1999.

[30]  J. E. Marsden and M. West, *Discrete mechanics and variational integrators*, Acta Numer., (2001), pp. 357–514.

[31]  J. Moser and A. P. Veselov, *Discrete versions of some classical integrable systems and factorization of matrix polynomials*, Comm. Math. Phys., 139 (1991), pp. 217–243.

[32]  T. Ohsawa and A. M. Bloch, *Nonholonomic Hamilton–Jacobi equation and integrability*, J. Geom. Mech., 1 (2009), pp. 461–481.

[33]  T. Ohsawa, O. E. Fernandez, A. M. Bloch, and D. V. Zenkov, *Nonholonomic Hamilton–Jacobi theory via Chaplygin Hamiltonization*, J. Geom. Phys., 61 (2011), pp. 1263–1291.

[34]  Y. B. Suris, *The problem of integrable discretization: Hamiltonian Approach*, Birkhäuser, Basel, 2003.

[35]  Y. B. Suris, *Discrete Lagrangian models*, in Discrete Integrable Systems, Lecture Notes in Physics 644, Springer, New York, 2004, pp. 111–184.