Nonholonomic and Discrete Hamilton–Jacobi Theory

by

Tomoki Ohsawa

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Applied and Interdisciplinary Mathematics)
in The University of Michigan
2010

Doctoral Committee:

Professor Anthony M. Bloch, Co-Chair
Professor N. Harris McClamroch, Co-Chair
Professor Alejandro Uribe-Ahumada
Associate Professor Melvin Leok, University of California, San Diego

We have done considerable mountain climbing. Now we are in the rarefied atmosphere of theories of excessive beauty and we are nearing a high plateau on which geometry, optics, mechanics, and wave mechanics meet on common ground. Only concentrated thinking, and a considerable amount of re-creation, will reveal the full beauty of our subject in which the last word has not yet been spoken.

— Cornelius Lanczos on Hamilton–Jacobi Theory [42]

To my parents

Acknowledgments

First of all, I would like to thank my advisor Tony Bloch, who has always been patient, supportive, and encouraging. Tony gave countless ideas, suggestions, and insights on my studies, which, through my often slow reactions, thinking, and interpretations, crystallized into this thesis.

The thesis would not be possible without Melvin Leok, who supported and supervised my visits at Purdue, Caltech, and UCSD. Melvin was always willing to have hours of discussions with me, which not only deepened my understanding but also provided me with ideas on research, and helped me in many aspects of academic life.

I am grateful to Harris McClamroch for being the co-advisor of the thesis, and Alejandro Uribe, who kindly agreed to serve as a committee member. To my academic brothers Dmitry Zenkov and Oscar Fernandez, who listened to me carefully and encouraged me to study nonholonomic mechanics. To Martín de Diego, Juan Carlos Marrero, Diana Sosa, and Jingjing Zhang, who gave helpful comments on my results. To Nathalie Fiveland, Emily Garvan, Tara McQueen, Warren Noone, Bert Ortiz, and Jessica Taylor, who were always there to help us grad students, and played the central role in making the department a very friendly place for us to be.

I would also like to thank the organizers of the Second and Third International Summer Schools on Geometry, Mechanics, and Control for their very generous financial support. The summer schools in the beautiful Canary Islands and L'Ametlla de Mar were absolutely fun and full of inspiration, giving me invaluable opportunities to meet and discuss with dozens of people who share interests in geometric mechanics and geometric control theory.

I had so many, well sometimes too many, great conversations with my officemates Henry Boateng, Catherine Kublik, and Richard Vasques. To Rizwan Khan and Greg McNulty: I will never forget the fun times, with and without math, at the Math Atrium!

Yoshifumi Suzuki, the *sempai* of mine at Michigan, helped me in many ways since my very first day in Ann Arbor, and showed me what it means to have a life in the

small town in the Midwest.

Wooi-Chen Ng and his friends (too many to mention all of them!) brought so much fun to my life in West Lafayette during the summer of 2008. Without them, I would have drowned in the sea of commutative diagrams!

Bin Cheng, the wonderful roommate who was always there to chat to, helped me survive my stressful last year at Michigan. A special thanks to his cat Huniu, who, perhaps unwillingly, served as a live example of the Falling Cat Theorem. After all, she seems to know geometric mechanics better than me!

Last, but not the least, I would like to thank my parents for letting me live my own life, for always being understanding toward their particularly stubborn son, and for being my parents.

Table of Contents

List of Tables

List of Figures

Figure

Chapter 1

Introduction

1.1 Hamilton–Jacobi Theory

Hamilton–Jacobi theory for continuous-time unconstrained systems is well understood from both the classical and geometric points of view. In classical mechanics [see, e.g., 3; 27; 42; 47], 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, and suppose that $(\hat{q}(s), \hat{p}(s)) \in T^*Q$ is a solution of Hamilton's equations

$$\dot{q} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial q}. \tag{1.1}$$

Calculate the action integral along the solution starting from $s = 0$ and ending at $s = t$ with $t > 0$:

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

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

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

This is the *(time-dependent) Hamilton–Jacobi 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 flow) that describe the dynamics defined by Hamilton's equations.

Furthermore, with a specified energy E, define $W : Q \to \mathbb{R}$ by

$$W(q) = S(q, t) - E\,t,$$

where on the right-hand side, t is seen as a function of the endpoint q. Then Eq. (1.3) turns into the *(time-independent) Hamilton–Jacobi equation.*

$$H\left(q, \frac{\partial W}{\partial q}\right) = E. \tag{1.4}$$

Hamilton–Jacobi theory plays a significant role in Hamiltonian dynamics. In particular, the fact that solving the Hamilton–Jacobi equation gives a generating function for the family of canonical transformation of the dynamics is the theoretical basis for the powerful technique of exact integration of Hamilton's equations (1.1) [see, e.g., 3; 27; 42] that are often employed with the technique of separation of variables. In fact, Arnold [3, §47, p. 261] states that this technique (which he refers to Jacobi's theorem) "is the most powerful method known for exact integration, and many problems which were solved by Jacobi cannot be solved by other methods."

The idea of Hamilton–Jacobi theory is also useful in optimal control theory [see, e.g., 34]. Namely, an argument similar to the above derivation of the Hamilton–Jacobi equation applied to optimal control problems yields the Hamilton–Jacobi–Bellman equation, which is a partial differential equation satisfied by the optimal cost function. It can also be shown that the costate of the optimal solution is related to the solution of the Hamilton–Jacobi–Bellman equation.

1.2 Extension to Nonholonomic Mechanics

One of our objectives is to extend Hamilton–Jacobi theory to nonholonomic systems, that is, mechanical systems with non-integrable velocity constraints. Nonholonomic mechanics deals with such systems by extending the ideas of Lagrangian and Hamiltonian mechanics [see, e.g., 7]. However, it is often not straightforward to do so, since a mechanical system loses some properties that are common to (conventional) Lagrangian and Hamiltonian systems when one adds nonholonomic constraints.

1.2.1 Nonholonomic Hamilton–Jacobi Theory

Since Hamilton–Jacobi theory is developed based on the Hamiltonian picture of dynamics, a natural starting point in extending Hamilton–Jacobi theory to nonholonomic systems is a Hamiltonian formulation of nonholonomic mechanics. Bates and Sniatycki [4] and van der Schaft and Maschke [58] generalized the definition of Hamiltonian system to the almost-symplectic and almost-Poisson formulations, respectively [see also 7; 38; 39]. As is shown in these papers, adding nonholonomic constraints to a Hamiltonian system renders the flow of the system non-symplectic. In fact, van der Schaft and Maschke [58] showed that the condition for the almost-Poisson Hamiltonian system to be (strictly) Poisson is equivalent to the system being holonomic. This implies that the conventional Hamilton–Jacobi theory does not directly apply to nonholonomic mechanics, since the (strict) symplecticity is critical in the theory. In fact, the Hamilton–Jacobi equation is a PDE for generating functions that yield symplectic maps for the flows of the dynamics.

There are some previous attempts to extend Hamilton–Jacobi theory to nonholonomic mechanics, such as Pavon [52]. However, as pointed out by Iglesias-Ponte et al. [30], these results are based on a variational approach, which does not apply to nonholonomic setting. See de León et al. [19] for details. Iglesias-Ponte et al. [30] proved a nonholonomic Hamilton–Jacobi theorem that shares the geometric view with the unconstrained theory by Abraham and Marsden [1]. The recent work by de León et al. [19] developed a new geometric framework for systems defined with linear almost Poisson structures. Their result generalizes Hamilton–Jacobi theory to the linear almost Poisson settings, and also specializes and provides geometric insights into nonholonomic mechanics.

Our work refines the result of Iglesias-Ponte et al. [30] so that nonholonomic Hamilton–Jacobi theory can be applied to exact integration of the equations of motion for nonholonomic systems.

1.2.2 Chaplygin Hamiltonization and Nonholonomic H–J Theory

There is an alternative less direct approach to nonholonomic Hamilton–Jacobi theory using the so-called Chaplygin Hamiltonization. The Chaplygin Hamiltonization [see, e.g., 16; 24; 25] is a method of transforming nonholonomic systems (which are not strictly Hamiltonian) into Hamiltonian systems. The conventional Hamilton–Jacobi theory applied to the transformed system gives the (conventional) Hamilton–Jacobi equation (1.3) or (1.4) related to the original nonholonomic dynamics. This approach is shown to give the same solutions as the direct approach mentioned above for some solvable nonholonomic systems. However, it is not clear how the direct approach is related to the Hamiltonization-based approach, since one approach is concerned with the original dynamics whereas the other with the Hamiltonized dynamics.

Our work relates the two approaches by first formulating the Chaplygin Hamiltonization in an intrinsic fashion. The intrinsic formulation clarifies the geometry involved in the Chaplygin Hamiltonization, which is often discussed locally in coordinates. We show that a link between the two different approaches to nonholonomic Hamilton–Jacobi theory comes out rather naturally from the geometric picture.

1.3 Discrete-Time Formulation

The second part of the thesis is also concerned with Hamilton–Jacobi theory, but the work is independent of the nonholonomic one; our focus turns to the development of a discrete-time version of Hamilton–Jacobi theory that fits into the framework of so-called discrete mechanics.

1.3.1 Discrete Mechanics

Discrete mechanics, a discrete-time version of Lagrangian and Hamiltonian mechanics, provides not only a systematic view of structure-preserving integrators but also a discrete-time counterpart to the theory of Lagrangian and Hamiltonian mechanics [see, e.g., 48; 55; 56]. 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. In other words, 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.

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 [48]. 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., 50; 55; 56] and also those of reduction and connections [31; 43; 46]. Whereas the main topic in discrete mechanics is the development of structure-preserving algorithms for Lagrangian and Hamiltonian systems [see, e.g., 48], 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 (nonsingular) discrete optimal control problems. In fact, as stated in Marsden and West [48], discrete mechanics is inspired by discrete formulations of optimal control problems (see, for example, Jordan and Polak [32] and Cadzow [11]).

1.3.2 Discrete Hamilton–Jacobi Theory

We develop the discrete-time version of Hamilton–Jacobi theory in a way analogous to that of the continuous-time counterpart. Much of the ideas are essentially a translation of ideas and concepts in continuous time to the discrete-time setting. Specifically, our starting point is a discrete counterpart of the derivation of the Hamilton–Jacobi equation from the action integral sketched in Section 1.1. We also relate the solutions of the resulting discrete Hamilton–Jacobi equation with those of the discrete Hamilton's equations. This is again a discrete analogue of the result mentioned in Section 1.1.

The theory specializes to linear discrete Hamiltonian systems and (regular) discrete optimal control problems, and the discrete Hamilton–Jacobi equation gives the discrete Riccati equation and the discrete Hamilton–Jacobi–Bellman equation, respectively. Furthermore, some results in discrete Hamilton–Jacobi theory are shown to reduce to some well-known results in discrete optimal control theory.

Chapter 2

Basic Concepts in Geometric Mechanics and Discrete Mechanics

2.1 Geometric Mechanics

This section reviews basic notions and results of geometric mechanics following Arnold [3], Marsden and Ratiu [47], and Bloch [7].

2.1.1 Hamiltonian Mechanics on Symplectic Manifolds

Let P be a symplectic manifold, that is, a manifold with a symplectic form Ω, i.e., a closed non-degenerate two-form on P. Given a Hamiltonian $H : P \to \mathbb{R}$, the corresponding *Hamiltonian vector field* $X_H \in \mathfrak{X}(P)$ is defined by *Hamilton's equations*

$$i_{X_H}\Omega = dH. \tag{2.1}$$

The most important case in terms of applications is the one with P being the cotangent bundle T^*Q of a differentiable manifold Q called the *configuration space (or manifold)* and each point in Q represents a configuration of the mechanical systems of interest. Let $\pi_Q : T^*Q \to Q$ be the cotangent bundle projection, and define the *standard symplectic one-form* Θ on T^*Q as follows: For any $p_q \in T^*Q$ and $v_{p_q} \in T_{p_q}T^*Q$,

$$\left\langle \Theta(p_q), v_{p_q} \right\rangle = \left\langle p_q, T\pi_Q(v_{p_q}) \right\rangle. \tag{2.2}$$

Then the two-form Ω on T^*Q defined by

$$\Omega := -d\Theta \tag{2.3}$$

is a symplectic form on T^*Q; it is called the *standard symplectic form*. So any cotangent bundle is a symplectic manifold. Let $\dim Q = n$ and $(q^1, \ldots, q^n)$ be local coordinates for Q. This induces the basis $\{dq^i\}_{i=1}^n$ for T_q^*Q. Then a point in T^*Q has the local coordinate expression $(q^1, \ldots, q^n, p_1, \ldots, p_n)$. Then we have $\Theta = p_i \, dq^i$ and $\Omega = dq^i \wedge dp_i$, and Eq. (2.1) gives the conventional form of Hamilton's equations:

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q^i}. \tag{2.4}$$

One of the most fundamental properties of Hamiltonian systems is conservation of the Hamiltonian:

Proposition 2.1.1. *The Hamiltonian function H is conserved along the flow of X_H:*

$$X_H[H] = dH(X_H) = 0. \tag{2.5}$$

2.1.2 Hamiltonian Mechanics on Poisson Manifolds

One can also take the Poisson point of view, instead of the symplectic one.

Definition 2.1.2. A *Poisson bracket* on a manifold P is a bilinear map $\{\,\cdot\,,\,\cdot\,\} : C^\infty(P) \times C^\infty(P) \to \mathbb{R}$ that satisfies

(i) anti-commutativity: $\{f, g\} = -\{g, f\}$,

(ii) Jacobi's identity: $\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0$,

(iii) Leibniz's rule: $\{fg, h\} = f\{g, h\} + g\{f, h\}$.

A manifold P endowed with a Poisson bracket is called a *Poisson manifold*.

Hamilton's equations on a Poisson manifold P is defined as follows: For any $F : P \to \mathbb{R}$,

$$\dot{F} = \{F, H\}. \tag{2.6}$$

Any symplectic manifold (P, Ω) is a Poisson manifold with the Poisson bracket defined by

$$\{F, G\} := \Omega(X_F, X_G). \tag{2.7}$$

for any $F, G : P \to \mathbb{R}$, where X_F and X_G are the vector fields on P defined by $i_{X_F}\Omega = dF$ and $i_{X_G}\Omega = dG$.

2.1.3 Momentum Maps and Noether's Theorem

Let G be a Lie group and $\mathfrak{g}$ its Lie algebra. Consider a symplectic group action of G on a symplectic manifold (P, Ω), i.e., $\Phi_g^P : P \to P;\ z \mapsto gz$ such that $(\Phi_g^P)^*\Omega = \Omega$. For any element $\xi \in \mathfrak{g}$, we define the infinitesimal generator $\xi_P \in \mathfrak{X}(P)$ as follows:

$$\xi_P(z) := \left.\frac{d}{dt}\Phi^P_{\exp t\xi}(z)\right|_{t=0} \tag{2.8}$$

for any $z \in P$, where $\exp : \mathfrak{g} \to G$ is defined by $\exp\xi := \gamma_\xi(1)$; the curve $\gamma_\xi : \mathbb{R} \to G$ is the solution with initial condition $\gamma_\xi(0) = e$ of the differential equation

$$\frac{d}{dt}\gamma_\xi(t) = T_e L_{\gamma_\xi(t)}(\xi), \tag{2.9}$$

where $L_h : G \to G;\ g \mapsto hg$ is the left translation map, i.e., γ_ξ is an integral curve of the vector field on G defined by the left-invariant extension of $\xi \in \mathfrak{g}$.

Definition 2.1.3. Suppose there exists a linear map $J_{(\cdot)} : \mathfrak{g} \to C^\infty(P)$ such that

$$i_{\xi_P}\Omega = dJ_\xi, \tag{2.10}$$

for any $\xi \in \mathfrak{g}$, i.e., the infinitesimal generator ξ_P is the Hamiltonian vector field associated to the function $J_\xi : P \to \mathbb{R}$. Then the *momentum map* $\mathbf{J} : P \to \mathfrak{g}^*$ is defined by

$$\langle \mathbf{J}(z), \xi \rangle = J_\xi(z), \tag{2.11}$$

for any $z \in P$.

If P is a cotangent bundle, i.e., $P = T^*Q$, and given a Lie group G acting on Q, i.e., we have $\Phi_g : Q \to Q;\ q \mapsto gq$. Then there is a natural corresponding symplectic action on its cotangent bundle T^*Q: For any $g \in G$ and $q \in Q$, let $\Phi_g^{T^*Q} : T_q^*Q \to T_{gq}^*Q$ be the cotangent lift [see, e.g., 47, Section 6.3] of Φ_g, i.e., $\Phi_g^{T^*Q} := T_q^*\Phi_{g^{-1}}$ where

$$\left\langle T_q^*\Phi_{g^{-1}}(p_q), v_{gq} \right\rangle = \langle p_q, T_q\Phi_{g^{-1}}(v_{gq}) \rangle \tag{2.12}$$

for any $p_q \in T_q^*Q$ and $v_{gq} \in T_{gq}Q$. Since the cotangent lift of any diffeomorphism is symplectic [see, e.g., 47, Proposition 6.3.2 on p. 170], the action $T^*\Phi_{g^{-1}} : T^*Q \to T^*Q$ is automatically symplectic. Moreover, we have an explicit expression for the map J_ξ:

Proposition 2.1.4. *If P is a cotangent bundle, i.e., $P = T^*Q$, then the map*

$J_{(\cdot)} : \mathfrak{g} \to C^\infty(T^*Q)$ is given by

$$J_\xi(p_q) := \langle p_q, \xi_Q(q) \rangle \,, \tag{2.13}$$

and thus the momentum map $\mathbf{J} : T^*Q \to \mathfrak{g}^*$ is defined by

$$\langle \mathbf{J}(p_q), \xi \rangle = \langle p_q, \xi_Q(q) \rangle \tag{2.14}$$

for any $p_q \in T^*Q$. Furthermore, $\mathbf{J}$ is equivariant, i.e.,

$$\mathbf{J} \circ T^*\Phi_g = \mathrm{Ad}_g^* \circ \mathbf{J} \tag{2.15}$$

for any $g \in G$.

Proof. See, e.g., Marsden and Ratiu [47, Theorem 12.1.4 on p. 386]. □

Definition 2.1.5. Given a Hamiltonian $H : P \to \mathbb{R}$, a Lie group G is called a *symmetry group* of the Hamiltonian system Eq. (2.1) if $H \circ \Phi_g^P = H$ for any $g \in G$.

Theorem 2.1.6 (Noether's Theorem)**.** *Consider a Hamiltonian system on a symplectic manifold P with Hamiltonian $H : P \to \mathbb{R}$. Suppose G is a symmetry group of the Hamiltonian system. Then the corresponding momentum map $\mathbf{J}$ is conserved along the Hamiltonian flow defined by X_H, i.e.,*

$$\mathbf{J} \circ \varphi_t = \mathbf{J}$$

for $t \in \mathbb{R}$, where $\varphi_t : P \to P$ is the flow of X_H.

2.2 Integrability of Hamiltonian Systems

One of the most important aspects of Hamiltonian mechanics is the question of their integrability or exact solvability of Hamilton's equations. The following theorem establishes the so-called *complete integrability* of Hamiltonian systems, and "covers all the problems of dynamics which have been integrated to the present day." [3]; so the assumptions of the theorem is often recognized as the definition of integrability of Hamiltonian systems:

Theorem 2.2.1 (Liouville–Arnold [3])**.** *Let (P, Ω) be a $2n$-dimensional symplectic manifold and $H : P \to \mathbb{R}$ a Hamiltonian. Suppose that there exist $F_i : P \to \mathbb{R}$*

for $i = 1, \ldots, n$ *with* $F_1 := H$ *such that* $\{F_i, F_j\} = 0$ *for any* $1 \le i, j \le n$. *Let* $\mathbf{f} = (f_1, \ldots, f_n) \in \mathbb{R}^n$ *and define*

$$M_{\mathbf{f}} := \{ z \in P \mid F_i(z) = f_i,\ i = 1, \ldots, n \} .$$

Then

(1) $M_{\mathbf{f}}$ *is a smooth manifold invariant under the flow of* X_H.

(2) If $M_{\mathbf{f}}$ *is compact and connected, then it is diffeomorphic to the* n*-dimensional torus*

$$\mathbb{T}^n = \underbrace{\mathbb{S}^1 \times \cdots \times \mathbb{S}^1}_{n} = \{(\varphi_1, \ldots, \varphi_n)\}.$$

(3) The flow of X_H *is a conditionally periodic motion on* $M_{\mathbf{f}} \cong \mathbb{T}^n$, *i.e., we have*

$$\frac{d\varphi_i}{dt} = \omega_i(\mathbf{f})$$

with some $\omega_i : \mathbb{R}^n \to \mathbb{R}$ *for* $i = 1, \ldots, n$,

(4) Hamilton's equations (2.1) *can be integrated by quadratures.*

Proof. See Arnold [3, §49]. □

Definition 2.2.2. The variables $(\varphi_1, \ldots, \varphi_n)$ defined above are called the *angle variables.* Those variables $(I_1, \ldots, I_n)$ so that the coordinate system (φ, I) for P is symplectic, i.e., the standard symplectic form Ω is locally written as $\Omega = d\varphi_i \wedge dI_i$, are called the *action variables.* The coordinate system (φ, I) is called the *action-angle variables.*

In the action-angle variables, Hamilton's equations (2.1) reduces to the form

$$\frac{d\varphi_i}{dt} = \omega_i(\mathbf{I}), \qquad \frac{dI_i}{dt} = 0, \tag{2.16}$$

and therefore

$$\varphi_i(t) = \omega_i(\mathbf{I})\, t + \varphi_i(0), \tag{2.17}$$

where $\mathbf{I} := (I_1, \ldots, I_n) \in \mathbb{R}^n$ is a constant vector.

2.3 Hamilton–Jacobi Theory

This section gives a brief account of Hamilton–Jacobi theory. We give the geometric description of Hamilton–Jacobi theory of Abraham and Marsden [1] as well as a brief survey of the method of separation of variables for the Hamilton–Jacobi equation to solve Hamilton's equations. The link between the Liouville–Arnold Theorem (Theorem 2.2.1) and Hamilton–Jacobi theory lies in the action-angle variables defined above; in practice the action-angle variables can be found through separation of variables for the Hamilton–Jacobi equation [see, e.g., 33, §6.2].

2.3.1 The Hamilton–Jacobi Theorem

Theorem 2.3.1 (Hamilton–Jacobi)**.** *Consider a Hamiltonian system Eq.* (2.1) *defined on the cotangent bundle T^*Q of a connected differentiable manifold Q with the standard symplectic form Ω. Let $W : Q \to \mathbb{R}$ be a smooth function defined on Q. Then the following are equivalent:*

(i) For every curve $c(t)$ in Q satisfying

$$\dot{c}(t) = T\pi_Q \cdot X_H(dW \circ c(t)), \tag{2.18}$$

the curve $t \mapsto dW \circ c(t)$ is an integral curve of X_H.

(ii) The function W satisfies the Hamilton–Jacobi equation*:*

$$H \circ dW = E, \tag{2.19}$$

where E is a constant.

Proof. See Abraham and Marsden [1, Theorem 5.2.4 on p. 381]. □

2.3.2 Separation of Variables

Let us briefly show how separation of variables works in solving the time-independent Hamilton–Jacobi equation (1.4). One first assumes that the function W can be split into pieces, each of which depends only on some subset of the variables q, e.g.,

$$W(q) = W_1(q_1) + W_2(q_2),$$

for $W_1, W_2 : Q \to \mathbb{R}$, and $q = (q_1, q_2)$. Then this sometimes helps us split the Hamilton–Jacobi equation (3.1) as follows:

$$H_1\left(q_1, \frac{\partial W_1}{\partial q_1}\right) = H_2\left(q_2, \frac{\partial W_2}{\partial q_2}\right),$$

with some functions $H_1, H_2 : T^*Q \to \mathbb{R}$. The left-hand side depends only on q_1 whereas the right-hand side depends only on q_2; this implies that both sides must be constant:

$$H_1\left(q_1, \frac{\partial W_1}{\partial q_1}\right) = H_2\left(q_2, \frac{\partial W_2}{\partial q_2}\right) = C.$$

Then we can solve them to obtain $\partial W_1/\partial q_1$ and $\partial W_2/\partial q_2$ separately, and hence dW. Now Theorem 2.3.1 implies that substituting this dW into Eq. 2.18 gives the set of equations that defines the dynamics on Q. So the problem of solving Hamilton's equations 2.4, which is a set of $2n$ ODEs, reduces to that of solving the set of n ODEs shown in Eq. 2.18, and it often turns out that one can solve Eq. 2.18 by quadrature.

Let us show a simple example of how this method works:

Example 2.3.2 (The plane central-force problem; Example 6.1 of José and Saletan [33])**.** Consider the Hamiltonian system whose configuration manifold is a plane, i.e., $Q = \mathbb{R}^2 = \{(r, \theta)\}$ with the Hamiltonian

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2}\right) + V(r).$$

The Hamilton–Jacobi equation (2.19) is

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial W}{\partial \theta}\right)^2\right] + V(r) = E.$$

Assuming that the function W is written as

$$W(r, \theta) = W_r(r) + W_\theta(\theta),$$

we obtain

$$2mr^2[E - V(r)] - r^2\left(\frac{dW_r}{dr}\right)^2 = \left(\frac{dW_\theta}{d\theta}\right)^2.$$

Since the left-hand side depends only on r whereas the right-hand side depends only on θ, we obtain

$$2mr^2[E - V(r)] - r^2\left(\frac{dW_r}{dr}\right)^2 = \left(\frac{dW_\theta}{d\theta}\right)^2 = C^2,$$

with some constant C, and hence

$$\frac{dW_r}{dr} = \sqrt{2m[E - V(r)] - \frac{C^2}{r^2}}, \qquad \frac{dW_\theta}{d\theta} = C,$$

assuming $\partial W_r/\partial r$ is positive. Then Eq. (2.18) gives

$$\dot{r} = \frac{1}{m}\sqrt{2m[E - V(r)] - \frac{C^2}{r^2}}, \qquad \dot{\theta} = \frac{C}{mr^2},$$

which are solved by quadrature.

2.4 Nonholonomic Mechanics

Nonholonomic mechanics is an extension of Lagrangian and Hamiltonian mechanics that incorporates so-called *nonholonomic constraints*, or in other words, non-integrable constraints[1].

2.4.1 Nonholonomic Constraints

Many mechanical systems have some form of constraints. There are essentially two kinds of constraints: holonomic and nonholonomic constraints. In loose terms, they are classified as follows: A holonomic constraint restricts the dynamics only in terms of position, or in other words, it tells *where the dynamics should be.* On the other hand, a nonholonomic constraint does it in terms of velocity only, or it tells *in which direction the dynamics should go.* Typical examples of nonholonomic constraints are those imposed by rolling and sliding of the mechanical systems. Such systems often arise in engineering problems, e.g., systems with wheels like cars and bicycles and those with sliding parts like sleighs (see Fig. 2.1).

From the geometric point of view, a holonomic constraint restricts the dynamics on a submanifold S of the configuration manifold Q; on the other hand, a nonholonomic constraint restricts the dynamics $q(t)$ on the subbundle $\mathcal{D} \subset TQ$ defined by

[1]We will later discuss what we mean by "non-integrable".

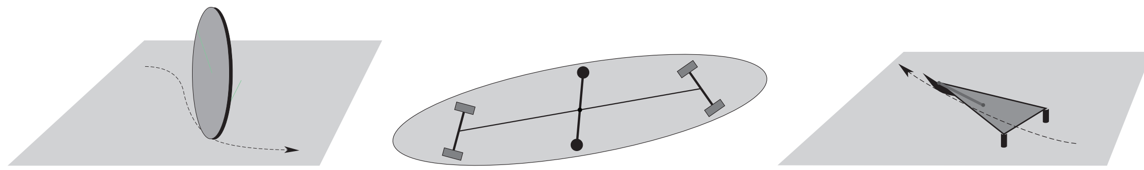

Figure 2.1: Examples of nonholonomic systems: Rolling disk, Snakeboard, and Sleigh.

the constraints, i.e., $\dot{q}(t) \in \mathcal{D}_{q(t)}$, and $\mathcal{D}$ is non-integrable in the sense that there is no local submanifold whose tangent space is given by $\mathcal{D}$ (see Fig. 2.2). This cru-

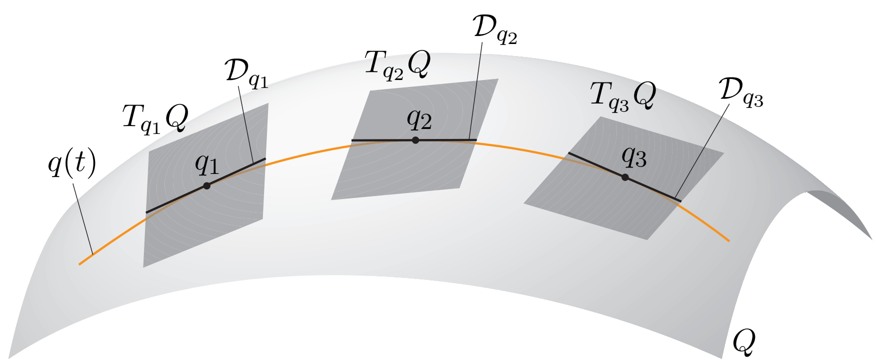

Figure 2.2: Distribution $\mathcal{D}$ defined by nonholonomic constraints.

cial difference makes nonholonomic constraints much more difficult to deal with than holonomic ones. For holonomic constraints, one can simply choose the submanifold S as the new configuration manifold and do (unconstrained) mechanics on it. However, for nonholonomic constraints, no such straightforward workaround is available and thus one has to extend Lagrangian and Hamiltonian mechanics so that it can deal with such constraints.

2.4.2 Hamiltonian Formulation of Nonholonomic Mechanics

Hamiltonian approaches to nonholonomic mechanical systems are developed by, for example, Bates and Sniatycki [4] and van der Schaft and Maschke [58]. See also Koon and Marsden [38, 39] and Bloch [7].

Consider a mechanical system on a differentiable manifold Q with Lagrangian $L : TQ \to \mathbb{R}$. Suppose that the system has nonholonomic constraints given by the distribution

$$\mathcal{D} := \left\{ v \in TQ \mid \omega^s(v) = A_i^s v^i = 0,\ s = 1, \dots, m \right\}. \tag{2.20}$$

Then the Lagrange–d'Alembert principle gives the equations of motion [see, e.g., 7, Chapter 5]:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = \lambda_s A_i^s, \tag{2.21}$$

where λ_s are Lagrange multipliers and $\omega^s = A_i^s\, dq^i$ are linearly independent non-exact one-forms on Q. The Legendre transformation of this set of equations gives the Hamiltonian formulation of nonholonomic systems. Specifically, define the Legendre transform $\mathbb{F}L : TQ \to T^*Q$ by

$$\langle \mathbb{F}L(v_q), w_q \rangle = \left. \frac{d}{d\varepsilon} L(v_q + \varepsilon\, w_q) \right|_{\varepsilon=0},$$

for $v_q, w_q \in T_qQ$. We assume that the Lagrangian is hyperregular, i.e., the Legendre transform $\mathbb{F}L$ is a diffeomorphism. Also define the Hamiltonian $H : T^*Q \to \mathbb{R}$ by

$$H(q,p) := \langle p, \dot{q} \rangle - L(q, \dot{q}),$$

where $\dot{q} = (\mathbb{F}L)^{-1}(p)$ on the right-hand side. Then we can rewrite Eq. (2.21) as follows:

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q^i} + \lambda_s A_i^s, \tag{2.22}$$

with the constraint equations

$$\omega^s(\dot{q}) = \omega^s\!\left(\frac{\partial H}{\partial p_i} \frac{\partial}{\partial q^i} \right) = 0 \quad \text{for} \quad s = 1, \dots, m. \tag{2.23}$$

Equations (2.22) and (2.23) define *Hamilton's equations for nonholonomic systems.* We can also write this system in the intrinsic form in the following way: Suppose that $X_H^{\mathrm{nh}} = \dot{q}^i \partial_{q^i} + \dot{p}_i \partial_{p_i}$ is the vector field on T^*Q that defines the flow of the system, Ω is the standard symplectic form on T^*Q, and $\pi_Q : T^*Q \to Q$ is the cotangent bundle projection. Then we can write Hamilton's equations for nonholonomic systems (2.22) and (2.23) in the following intrinsic form:

$$i_{X_H^{\mathrm{nh}}} \Omega = dH - \lambda_s \pi_Q^* \omega^s, \tag{2.24}$$

along with

$$T\pi_Q(X_H^{\mathrm{nh}}) \in \mathcal{D} \quad \text{or} \quad \omega^s(T\pi_Q(X_H^{\mathrm{nh}})) = 0 \text{ for } s = 1, \dots, m. \tag{2.25}$$

Introducing the *constrained momentum space* $\mathcal{M} := \mathbb{F}L(\mathcal{D}) \subset T^*Q$, the above constraints may be replaced by the following:

$$p \in \mathcal{M}. \tag{2.26}$$

2.4.3 Completely Nonholonomic Constraints

Let us introduce a special class of nonholonomic constraints that applies to all the examples in this thesis.

Definition 2.4.1 (Vershik and Gershkovich [59]; see also Montgomery [49])**.** A distribution $\mathcal{D} \subset TQ$ is said to be *completely nonholonomic* (or *bracket-generating*) if $\mathcal{D}$ along with all of its iterated Lie brackets $[\mathcal{D}, \mathcal{D}], [\mathcal{D}, [\mathcal{D}, \mathcal{D}]], \dots$ spans the tangent bundle TQ.

Let us also introduce the following notion for convenience:

Definition 2.4.2. Let Q be the configuration manifold of a mechanical system. Then nonholonomic constraints on the system are said to be *completely nonholonomic* if the distribution $\mathcal{D} \subset TQ$ defined by the nonholonomic constraints is completely nonholonomic (or bracket-generating).

One of the most important results concerning completely nonholonomic distributions is the following[2]:

Theorem 2.4.3 (Chow's Theorem)**.** *Let Q be a connected differentiable manifold. If a distribution $\mathcal{D} \subset TQ$ is completely nonholonomic, then any two points on Q can be joined by a horizontal path.*

We will need the following result that easily follows from Chow's Theorem:

Proposition 2.4.4. *Let Q be a connected differentiable manifold and $\mathcal{D} \subset TQ$ be a completely nonholonomic distribution. Then there is no non-zero exact one-form in the annihilator $\mathcal{D}^\circ \subset T^*Q$.*

Proof. Chow's Theorem says that, for any two points q_0 and q_1 in Q, there exists a curve $c : [0, T] \to Q$ with some $T > 0$ such that $c(0) = q_0$ and $c(T) = q_1$, and also $\dot{c}(t) \in \mathcal{D}_{c(t)}$ for any $t \in (0, T)$. Now let df be an exact one-form in the annihilator $\mathcal{D}^\circ$. Then by Stokes' theorem, we have

$$f(q_1) - f(q_0) = \int_0^T df(\dot{c}(t))\, dt = 0,$$

[2]See, e.g., Montgomery [49] for a proof.

where $df(\dot{c}(t)) = 0$ because $df \in \mathcal{D}^\circ$ and $\dot{c}(t) \in \mathcal{D}_{c(t)}$. Since q_0 and q_1 are arbitrary and Q is connected, this implies that f is constant on Q. □

2.4.4 Regularity of Nonholonomic Systems

We also introduce the notion of *regularity* of nonholonomic systems. Again all the nonholonomic systems treated in this thesis are regular.

Consider a nonholonomic system with a hyperregular Lagrangian $L : TQ \to \mathbb{R}$ and a constant-dimensional distribution $\mathcal{D} \subset TQ$ defined by nonholonomic constraints. For any $v_q \in TQ$ define a bilinear form $B_L(v_q) : T_qQ \times T_qQ \to \mathbb{R}$ by

$$B_L(v_q)(u_q, w_q) := \frac{\partial^2}{\partial\varepsilon_1 \partial\varepsilon_2} L(v_q + \varepsilon_1 u_q + \varepsilon_2 w_q)\bigg|_{\varepsilon_1=\varepsilon_2=0} = D_2 D_2 L(q, v) \cdot (u_q, w_q).$$

Then hyperregularity of the Lagrangian implies that the associated map $B_L^\flat(v_q) : T_qQ \to T_q^*Q$ defined by

$$\left\langle B_L^\flat(v_q)(u_q), w_q \right\rangle := B_L(v_q)(u_q, w_q)$$

is an isomorphism. Thus we can define a bilinear form $W_L : T_q^*Q \times T_q^*Q \to \mathbb{R}$ by

$$W_L(v_q)(\alpha_q, \beta_q) := \left\langle \alpha_q, (B_L^\flat)^{-1}(\beta_q) \right\rangle.$$

Definition 2.4.5 (de León and Martín de Diego [18]; see also de León et al. [20])**.** In the above setup, suppose that the annihilator $\mathcal{D}^\circ$ is spanned by the one-forms $\{\omega^s\}_{s=1}^m$. Then the nonholonomic system is said to be *regular* if the matrices $(\mathcal{C}_L^{rs}(v))$ defined by

$$\mathcal{C}_L^{rs}(v) := -W_L(v)(\omega^r, \omega^s) \tag{2.27}$$

are nonsingular for any $v \in \mathcal{D}$.

For a mechanical system whose Lagrangian is kinetic minus potential energy, regularity follows automatically:

Proposition 2.4.6 (Cariñena and Rañada [13]; see also de León and Martín de Diego [18])**.** *If the Lagrangian $L : TQ \to \mathbb{R}$ has the form*

$$L(q, v) = \frac{1}{2} g_q(v, v) - V(q), \tag{2.28}$$

with g being a Riemannian metric on Q, then the nonholonomic system is regular.

Proof. In this case $D_2D_2L(q,v)(u_q,w_q) = g_q(u_q,w_q)$, and so W_L is defined by the inverse g^{ij} of the matrix g_{ij}. Since g_{ij} is positive-definite, so is the inverse g^{ij}; hence it follows that W_L is positive-definite. A positive-definite matrix restricted to a subspace is again positive-definite, and so $\mathcal{C}_L^{rs}$ is positive-definite and hence nondegenerate. □

In the Hamiltonian setting with the form of Lagrangian in Eq. (2.28), we have the following result:

Proposition 2.4.7 (Bates and Sniatycki [4])**.** *Suppose that the Lagrangian is of the form in Eq.* (2.28)*. Let $\mathcal{F}$ be the distribution on T^*Q defined by*

$$\mathcal{F} := \{v \in TT^*Q \mid T\pi_Q(v) \in \mathcal{D}\}, \tag{2.29}$$

and then define a distribution $\mathcal{H}$ on $\mathcal{M} := \mathbb{F}L(\mathcal{D})$ by

$$\mathcal{H} := \mathcal{F} \cap T\mathcal{M}. \tag{2.30}$$

Then the standard symplectic form Ω restricted to $\mathcal{H}$ is nondegenerate.

Proof. See Bates and Sniatycki [4, Theorem on p. 105]. □

2.5 Discrete Mechanics

This section briefly reviews some key results of discrete mechanics following Marsden and West [48] and Lall and West [41].

2.5.1 Discrete Lagrangian Mechanics

A discrete Lagrangian flow $\{q_k\}$ for $k = 0, 1, \ldots, N$ on an n-dimensional differentiable manifold Q is defined by the following discrete variational principle: Let $S_{\rm d}^N$ be the following action sum of the discrete Lagrangian $L_{\rm d} : Q \times Q \to \mathbb{R}$:

$$S_{\rm d}^N(\{q_k\}_{k=0}^N) := \sum_{k=0}^{N-1} L_{\rm d}(q_k, q_{k+1}). \tag{2.31}$$

More precisely, given a Lagrangian $L : TQ \to \mathbb{R}$ for a continuous-time system, the corresponding discrete Lagrangian $L_{\rm d} : Q \times Q \to \mathbb{R}$ is an approximation of the *exact* *discrete Lagrangian* $L_{\mathrm{d}}^{\mathrm{ex}} : Q \times Q \to \mathbb{R}$ defined by

$$L_{\mathrm{d}}^{\mathrm{ex}}(q_k, q_{k+1}) := \int_{t_k}^{t_{k+1}} L(q(t), \dot{q}(t))\, dt,$$

where $q : [t_k, t_{k+1}] \to Q$ is the solution of the (continuous-time) Euler–Lagrange equation with the boundary conditions $q(t_k) = q_k$ and $q(t_{k+1}) = q_{k+1}$.

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

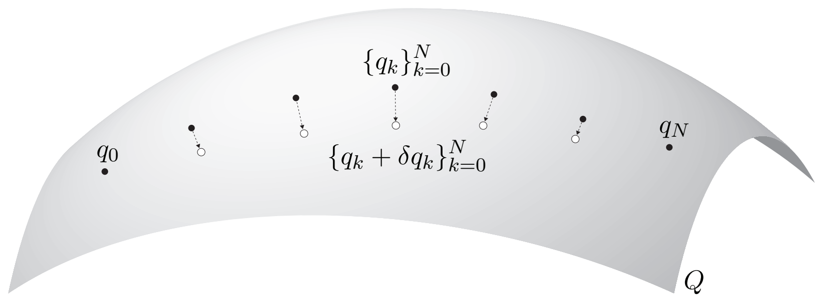

Figure 2.3: Discrete variations on configuration manifold Q.

equations:

$$D_2 L_{\mathrm{d}}(q_{k-1}, q_k) + D_1 L_{\mathrm{d}}(q_k, q_{k+1}) = 0. \tag{2.32}$$

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

$$F_{L_{\mathrm{d}}} : (q_{k-1}, q_k) \mapsto (q_k, q_{k+1}). \tag{2.33}$$

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

$$\Theta_{L_{\mathrm{d}}}^{+} : (q_k, q_{k+1}) \mapsto D_2 L_{\mathrm{d}}(q_k, q_{k+1})\, dq_{k+1}, \tag{2.34a}$$

$$\Theta_{L_{\mathrm{d}}}^{-} : (q_k, q_{k+1}) \mapsto -D_1 L_{\mathrm{d}}(q_k, q_{k+1})\, dq_k. \tag{2.34b}$$

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

$$\Omega_{L_{\mathrm{d}}}(q_k, q_{k+1}) := d\Theta_{L_{\mathrm{d}}}^{+} = d\Theta_{L_{\mathrm{d}}}^{-} = D_1 D_2 L_{\mathrm{d}}(q_k, q_{k+1})\, dq_k \wedge dq_{k+1}. \tag{2.35}$$

Specifically, we have

$$(F_{L_{\mathrm{d}}})^* \Omega_{L_{\mathrm{d}}} = \Omega_{L_{\mathrm{d}}}. \tag{2.36}$$

2.5.2 Discrete Hamiltonian Mechanics

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

$$\mathbb{F}L_{\mathrm{d}}^{+} : (q_k, q_{k+1}) \mapsto (q_{k+1}, D_2 L_{\mathrm{d}}(q_k, q_{k+1})), \tag{2.37a}$$

$$\mathbb{F}L_{\mathrm{d}}^{-} : (q_k, q_{k+1}) \mapsto (q_k, -D_1 L_{\mathrm{d}}(q_k, q_{k+1})). \tag{2.37b}$$

Then we find that the discrete Lagrangian symplectic forms Eq. (2.34) and (2.35) are pull-backs by these maps of the standard symplectic form on T^*Q:

$$\Theta_{L_{\mathrm{d}}}^{\pm} = (\mathbb{F}L_{\mathrm{d}}^{\pm})^*\Theta, \qquad \Omega_{L_{\mathrm{d}}}^{\pm} = (\mathbb{F}L_{\mathrm{d}}^{\pm})^*\Omega. \tag{2.38}$$

Let us define the momenta

$$p_{k,k+1}^{-} := -D_1 L_{\mathrm{d}}(q_k, q_{k+1}), \qquad p_{k,k+1}^{+} := D_2 L_{\mathrm{d}}(q_k, q_{k+1}). \tag{2.39}$$

Then the discrete Euler–Lagrange equations (2.32) become simply $p_{k-1,k}^{+} = p_{k,k+1}^{-}$. So defining

$$p_k := p_{k-1,k}^{+} = p_{k,k+1}^{-}, \tag{2.40}$$

one can rewrite the discrete Euler–Lagrange equations (2.32) as follows:

$$\begin{aligned} p_k &= -D_1 L_{\mathrm{d}}(q_k, q_{k+1}), \\ p_{k+1} &= D_2 L_{\mathrm{d}}(q_k, q_{k+1}). \end{aligned} \tag{2.41}$$

Notice that this can be interpreted as a symplectic map generated by the Type I generating function L_{d} [27].

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

$$\tilde{F}_{L_{\mathrm{d}}} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1}). \tag{2.42}$$

One may relate this map with the discrete Legendre transforms in Eq. (2.37) as follows:

$$\tilde{F}_{L_{\mathrm{d}}} = \mathbb{F}L_{\mathrm{d}}^{+} \circ (\mathbb{F}L_{\mathrm{d}}^{-})^{-1}. \tag{2.43}$$

The diagram below summarizes the relations between $F_{L_{\mathrm{d}}}$, $\tilde{F}_{L_{\mathrm{d}}}$, and $\mathbb{F}L_{\mathrm{d}}^{\pm}$ (see Marsden and West [48] for details).

$$
\begin{tikzcd}
Q \times Q \arrow[rr, "F_{L_\mathrm{d}}"] \arrow[dl, "\mathbb{F}L_\mathrm{d}^-"'] \arrow[dr, "\mathbb{F}L_\mathrm{d}^+"] & & Q \times Q \arrow[dl, "\mathbb{F}L_\mathrm{d}^-"'] \arrow[dr, "\mathbb{F}L_\mathrm{d}^+"] & \\
T^*Q \arrow[r, "\tilde{F}_{L_\mathrm{d}}"'] & T^*Q \arrow[rr, "\tilde{F}_{L_\mathrm{d}}"'] & & T^*Q
\end{tikzcd}
\qquad
\begin{tikzcd}
 & (q_0, q_1) \arrow[rr, mapsto] \arrow[dl, mapsto] \arrow[dr, mapsto] & & (q_1, q_2) \arrow[dl, mapsto] \arrow[dr, mapsto] & \\
(q_0, p_0) \arrow[rr, mapsto] & & (q_1, p_1) \arrow[rr, mapsto] & & (q_2, p_2)
\end{tikzcd}
\tag{2.44}
$$

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

$$
(\tilde{F}_{L_\mathrm{d}})^* \Omega = \Omega. \tag{2.45}
$$

This is the Hamiltonian description of the dynamics defined by the discrete Euler–Lagrange equation (2.32) introduced by Marsden and West [48]. However, notice 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 [41] 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 the first kind, we can apply Legendre transforms to the discrete Lagrangian to find the corresponding generating function of type two or three [27]. In fact, they turn out to be a natural Hamiltonian counterpart to the discrete Lagrangian mechanics described above. Specifically, with the discrete Legendre transform

$$
p_{k+1} = \mathbb{F}L_\mathrm{d}^+(q_k, q_{k+1}) = D_2 L_\mathrm{d}(q_k, q_{k+1}), \tag{2.46}
$$

we can define the following *right discrete Hamiltonian*:

$$
H_\mathrm{d}^+(q_k, p_{k+1}) = p_{k+1} \cdot q_{k+1} - L_\mathrm{d}(q_k, q_{k+1}), \tag{2.47}
$$

which is a generating function of the second kind [27]. 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*

$$
q_{k+1} = D_2 H_\mathrm{d}^+(q_k, p_{k+1}), \tag{2.48a}
$$

$$
p_k = D_1 H_\mathrm{d}^+(q_k, p_{k+1}). \tag{2.48b}
$$

Similarly, with the discrete Legendre transform

$$p_k = \mathbb{F}L_{\mathrm{d}}^{-}(q_k, q_{k+1}) = -D_1 L_{\mathrm{d}}(q_k, q_{k+1}), \tag{2.49}$$

we can define the following *left discrete Hamiltonian*:

$$H_{\mathrm{d}}^{-}(p_k, q_{k+1}) = -p_k \cdot q_k - L_{\mathrm{d}}(q_k, q_{k+1}), \tag{2.50}$$

which is a generating function of the third kind [27]. Then we have the *left discrete Hamilton's equations*

$$q_k = -D_1 H_{\mathrm{d}}^{-}(p_k, q_{k+1}), \tag{2.51a}$$
$$p_{k+1} = -D_2 H_{\mathrm{d}}^{-}(p_k, q_{k+1}). \tag{2.51b}$$

Leok and Zhang [44] 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.

Chapter 3

Nonholonomic Hamilton–Jacobi Theory

3.1 Introduction

This chapter develops an extension of Hamilton–Jacobi theory to nonholonomic mechanics based on the Hamiltonian formulation of nonholonomic mechanics discussed in Section 2.4.2. Much of the ideas in the proof of the nonholonomic Hamilton–Jacobi theorem come from identifying both the similarities and differences between the nonholonomic and unconstrained Hamilton's equations.

3.1.1 Nonholonomic Hamilton–Jacobi Theory

The previous work by Iglesias-Ponte et al. [30] and de León et al. [19] is of theoretical importance in its own right. However, it is still unknown if the theorems are applicable to the problem of exactly integrating the equations of motion of nonholonomic systems in a similar way to the conventional theory. To see this let us briefly discuss the difference between the unconstrained Hamilton–Jacobi equation and the nonholonomic ones mentioned above. First recall the conventional unconstrained theory: Let Q be a configuration space, T^*Q be its cotangent bundle, and $H : T^*Q \to \mathbb{R}$ be the Hamiltonian; then the Hamilton–Jacobi equation can be written as a *single* equation:

$$H\left(q, \frac{\partial W}{\partial q}\right) = E, \tag{3.1a}$$

or

$$H \circ dW(q) = E, \tag{3.1b}$$

for an unknown *function* $W : Q \to \mathbb{R}$. On the other hand, the nonholonomic Hamilton–Jacobi equations in [30] have the following form:

$$d(H \circ \gamma)(q) \in \mathcal{D}^{\circ}, \tag{3.2}$$

where $\gamma : Q \to T^*Q$ is an unknown *one-form*, and $\mathcal{D}^{\circ}$ is the annihilator of the distribution $\mathcal{D} \subset TQ$ defined by the nonholonomic constraints. While it is clear that Eq. (3.2) reduces to Eq. (3.1) for the special case that there are no constraints[1], Eq. (3.2) in general gives a set of partial differential equations for γ as opposed to a single equation like Eq. (3.1).

Having this difference in mind, let us now consider the following question: Is separation of variables applicable to the nonholonomic Hamilton–Jacobi equation? It is not clear how the approach shown in Section 2.3.2 applies to the nonholonomic Hamilton–Jacobi Equation (3.2). Furthermore, there are additional conditions on the solution γ which do not exist in the conventional theory.

3.1.2 Integrability of Nonholonomic Systems

Integrability of Hamiltonian systems is an interesting question that has a close link with Hamilton–Jacobi theory. For integrability of unconstrained Hamiltonian systems, the Arnold–Liouville theorem (Theorem 2.2.1) stands as the definitive work.

For nonholonomic mechanics, however, the Arnold–Liouville theorem does not directly apply, since the nonholonomic flow is not Hamiltonian and so the key ideas in the Arnold–Liouville theorem lose their effectiveness. Kozlov [40] gave certain conditions for integrability of nonholonomic systems with invariant measures. However, it is important to remark that there are examples that do not have invariant measures but are still integrable, such as the Chaplygin sleigh [see, e.g., 6; 7]. Also it is unknown how this result may be related to nonholonomic Hamilton–Jacobi theory, which does not have an apparent relationship with invariant measures.

3.1.3 Main Results

The goal of the present work is to fill the gap between the unconstrained and nonholonomic Hamilton–Jacobi theory by showing applicability of separation of variables to nonholonomic systems, and also to discuss integrability of them. For that purpose,

[1] $\mathcal{D} = TQ$ and hence $\mathcal{D}^{\circ} = 0$ and identifying the one-form γ with dW

we would like to first reformulate the nonholonomic Hamilton–Jacobi theorem from an intrinsic point of view[2]. We show that the nonholonomic Hamilton–Jacobi equation (3.2) reduces to a single equation $H \circ \gamma = E$. This result resolves the differences between unconstrained and nonholonomic Hamilton–Jacobi equations mentioned in Section 3.1.1, and makes it possible to apply separation of variables to nonholonomic systems. Furthermore, the intrinsic proof helps us identify the difference from the unconstrained theory by Abraham and Marsden [1] and find the conditions on the solution γ arising from nonholonomic constraints that are more practical than (although equivalent to, as pointed out by Sosa [53]) those of Iglesias-Ponte et al. [30]. It turns out that these conditions are not only useful in finding the solutions of the Hamilton–Jacobi equation by separation of variables, but also provide a way to integrate the equations of motion of a system to which separation of variables does not apply.

3.1.4 Outline

In Section 3.2 we formulate and prove the nonholonomic Hamilton–Jacobi theorem. The theorem and proof are an extension of the one by Abraham and Marsden [1] to the nonholonomic setting. In doing so we identify the differences from the unconstrained theory; this in turn gives the additional conditions arising from the nonholonomic constraints.

We apply the nonholonomic Hamilton–Jacobi theorem to several examples in Section 3.3. We first apply the technique of separation of variables to solve the nonholonomic Hamilton–Jacobi equation to obtain exact solutions of the motions of the vertical rolling disk and knife edge on an inclined plane. We then take the snakeboard and Chaplygin sleigh as examples to which separation of variables does not apply, and show another way of employing the nonholonomic Hamilton–Jacobi theorem to exactly integrate the equations of motion.

3.2 Nonholonomic Hamilton–Jacobi Theorem

We would like to refine the result of Iglesias-Ponte et al. [30] with a particular attention to applications to exact integration of the equations of motion. Specifically, we would like to take an intrinsic approach (see [30] for the coordinate-based approach)

[2] A coordinate-based proof is given in [30]

to clarify the difference from the (unconstrained) Hamilton–Jacobi theorem of Abraham and Marsden [1] (Theorem 5.2.4). A significant difference from the result by Iglesias-Ponte et al. [30] is that the nonholonomic Hamilton–Jacobi equation is given as a single algebraic equation $H \circ \gamma = E$ just as in the unconstrained Hamilton–Jacobi theory, as opposed to a set of differential equations $d(H \circ \gamma) \in \mathcal{D}^{\circ}$.

Theorem 3.2.1 (Nonholonomic Hamilton–Jacobi)**.** *Consider a nonholonomic system defined on a connected differentiable manifold Q with a Lagrangian of the form Eq.* (2.28) *and a completely nonholonomic constraint distribution $\mathcal{D} \subset TQ$. Let $\gamma : Q \to T^*Q$ be a one-form that satisfies*

$$\gamma(q) \in \mathcal{M}_q \text{ for any } q \in Q, \tag{3.3}$$

and

$$d\gamma|_{\mathcal{D}\times\mathcal{D}} = 0, \;\text{ i.e., } d\gamma(v, w) = 0 \text{ for any } v, w \in \mathcal{D}. \tag{3.4}$$

Then the following are equivalent:

(i) For every curve $c(t)$ in Q satisfying

$$\dot{c}(t) = T\pi_Q \cdot X_H(\gamma \circ c(t)), \tag{3.5}$$

the curve $t \mapsto \gamma \circ c(t)$ is an integral curve of $X_H^{\rm nh}$, where X_H is the Hamiltonian vector field of the unconstrained system with the same Hamiltonian, i.e., $i_{X_H}\Omega = dH$.

(ii) The one-form γ satisfies the nonholonomic Hamilton–Jacobi equation*:*

$$H \circ \gamma = E, \tag{3.6}$$

where E is a constant.

The following lemma, which is a slight modification of Lemma 5.2.5 of Abraham and Marsden [1], is the key to the proof of the above theorem:

Lemma 3.2.2. *For any one-form γ on Q that satisfies Eq.* (3.4) *and any $v, w \in \mathcal{F}$, the following equality holds:*

$$\Omega(T(\gamma \circ \pi_Q) \cdot v, w) = \Omega(v, w - T(\gamma \circ \pi_Q) \cdot w). \tag{3.7}$$

Proof. Notice first that $v - T(\gamma \circ \pi_Q) \cdot v$ is vertical for any $v \in TT^*Q$:

$$\begin{aligned} T\pi_Q \cdot (v - T(\gamma \circ \pi_Q) \cdot v) &= T\pi_Q(v) - T(\pi_Q \circ \gamma \circ \pi_Q) \cdot v \\ &= T\pi_Q(v) - T\pi_Q(v) = 0, \end{aligned}$$

where we used the relation $\pi_Q \circ \gamma \circ \pi_Q = \pi_Q$. Hence

$$\Omega(v - T(\gamma \circ \pi_Q) \cdot v, w - T(\gamma \circ \pi_Q) \cdot w) = 0,$$

and thus

$$\Omega(T(\gamma \circ \pi_Q) \cdot v, w) = \Omega(v, w - T(\gamma \circ \pi_Q) \cdot w) + \Omega(T(\gamma \circ \pi_Q) \cdot v, T(\gamma \circ \pi_Q) \cdot w).$$

However, the second term on the right-hand side vanishes:

$$\Omega(T(\gamma \circ \pi_Q) \cdot v, T(\gamma \circ \pi_Q) \cdot w) = \gamma^*\Omega(T\pi_Q(v), T\pi_Q(w)) = -d\gamma(T\pi_Q(v), T\pi_Q(w)) = 0,$$

where we used the fact that for any one-form β on Q, $\beta^*\Omega = -d\beta$ with β on the left-hand side being regarded as a map $\beta : Q \to T^*Q$ [see 1, Proposition 3.2.11 on p. 179], and the assumption that $d\gamma|_{\mathcal{D}\times\mathcal{D}} = 0$; note that $v, w \in \mathcal{F}$ implies $T\pi_Q(v), T\pi_Q(w) \in \mathcal{D}$. $\square$

Let us state another lemma:

Lemma 3.2.3. *The unconstrained Hamiltonian vector field X_H evaluated on the constrained momentum space $\mathcal{M}$ is in the distribution $\mathcal{F}$, i.e.,*

$$X_H(\alpha_q) \in \mathcal{F}_{\alpha_q} \text{ for any } \alpha_q \in \mathcal{M}_q.$$

Proof. We want to show that $T\pi_Q(X_H(\alpha_q))$ is in $\mathcal{D}_q$. First notice that

$$T\pi_Q(X_H(\alpha_q)) = \frac{\partial H}{\partial p_i}(\alpha_q)\,\frac{\partial}{\partial q^i} = \mathbb{F}H(\alpha_q),$$

where we defined $\mathbb{F}H : T^*Q \to TQ$ by

$$\langle \beta_q, \mathbb{F}H(\alpha_q) \rangle = \frac{d}{d\varepsilon} H(\alpha_q + \varepsilon\,\beta_q)\bigg|_{\varepsilon=0}.$$

However, because the Lagrangian L is hyperregular, we have $\mathbb{F}H = (\mathbb{F}L)^{-1}$ and thus

$$T\pi_Q(X_H(\alpha_q)) = (\mathbb{F}L)^{-1}(\alpha_q).$$

Now, by the definition of $\mathcal{M}$, $\alpha_q \in \mathcal{M}$ implies $\alpha_q \in \mathbb{F}L(\mathcal{D}_q)$, which gives $(\mathbb{F}L)^{-1}(\alpha_q) \in \mathcal{D}_q$ by the hyperregularity of L. Hence the claim follows. □

Proof of Theorem 3.2.1. Let us first show that (ii) implies (i). Assume (ii) and let $p(t) := \gamma \circ c(t)$, where $c(t)$ satisfies Eq. (3.5). Then

$$\begin{aligned}\dot{p}(t) &= T\gamma(\dot{c}(t)) \\ &= T\gamma \circ T\pi_Q \cdot X_H(\gamma \circ c(t)) \\ &= T(\gamma \circ \pi_Q) \cdot X_H(\gamma \circ c(t)).\end{aligned} \tag{3.8}$$

Therefore, using Lemmas 3.2.2 and 3.2.3, we obtain, for any $w \in \mathcal{F}$,

$$\begin{aligned}\Omega(T(\gamma \circ \pi_Q) \cdot X_H(p(t)), w) &= \Omega(X_H(p(t)), w - T(\gamma \circ \pi_Q) \cdot w) \\ &= \Omega(X_H(p(t)), w) - \Omega(X_H(p(t)), T(\gamma \circ \pi_Q) \cdot w).\end{aligned}$$

For the first term on the right-hand side, notice that for any $w \in \mathcal{F}$,

$$\Omega(X_H^{\mathrm{nh}}, w) = dH \cdot w - \lambda_s \pi_Q^* \omega^s(w) = dH \cdot w = \Omega(X_H, w).$$

Also for the second term,

$$\Omega(X_H(p(t)), T(\gamma \circ \pi_Q) \cdot w) = dH(p(t)) \cdot T(\gamma \circ \pi_Q) \cdot w = d(H \circ \gamma)(c(t)) \cdot T\pi_Q(w).$$

So we now have

$$\Omega(T(\gamma \circ \pi_Q) \cdot X_H(p(t)), w) = \Omega(X_H^{\mathrm{nh}}(p(t)), w) - d(H \circ \gamma)(c(t)) \cdot T\pi_Q(w). \tag{3.9}$$

However, the nonholonomic Hamilton–Jacobi equation (3.6) implies that the second term on the right-hand side vanishes. Thus we have

$$\Omega(T(\gamma \circ \pi_Q) \cdot X_H(p(t)), w) = \Omega(X_H^{\mathrm{nh}}(p(t)), w) \tag{3.10}$$

for any $w \in \mathcal{F}_{p(t)}$. Now $T(\gamma \circ \pi_Q) \cdot X_H \in T\mathcal{M}$ since γ takes values in $\mathcal{M}$; also

$T(\gamma \circ \pi_Q) \cdot X_H(p(t)) \in \mathcal{F}_{p(t)}$ because

$$T\pi_Q \circ T(\gamma \circ \pi_Q) \cdot X_H(p(t)) = T(\pi_Q \circ \gamma \circ \pi_Q) \cdot X_H(p(t)) = T\pi_Q \cdot X_H(p(t)) \in \mathcal{D},$$

using Lemma 3.2.3 again. Therefore $T(\gamma \circ \pi_Q) \cdot X_H(p(t)) \in \mathcal{H}_{p(t)}$. On the other hand, $X_H^{\rm nh}(p(t)) \in \mathcal{H}_{p(t)}$ as well: $X_H^{\rm nh}(p(t)) \in T_{p(t)}\mathcal{M}$ because $\mathcal{M}$ is an invariant manifold of the nonholonomic flow defined by $X_H^{\rm nh}$ and also $X_H^{\rm nh}(p(t)) \in \mathcal{F}_{p(t)}$ due to Eq. (2.25). Now, in Eq. (3.10), w is an arbitrary element in $\mathcal{F}_{p(t)}$ and thus Eq. (3.10) holds for any $w \in \mathcal{H}_{p(t)}$ because $\mathcal{H} \subset \mathcal{F}$. However, according to Proposition 2.4.7, Ω restricted to $\mathcal{H}$ is nondegenerate. So we obtain

$$T(\gamma \circ \pi_Q) \cdot X_H(p(t)) = X_H^{\rm nh}(p(t)),$$

and hence Eq. (3.8) gives

$$\dot{p}(t) = X_H^{\rm nh}(p(t)).$$

This means that $p(t)$ gives an integral curve of $X_H^{\rm nh}$. Thus (ii) implies (i).

Conversely, assume (i); let $c(t)$ be a curve in Q that satisfies Eq. (3.5) and set $p(t) := \gamma \circ c(t)$. Then $p(t)$ is an integral curve of $X_H^{\rm nh}$ and so

$$\dot{p}(t) = X_H^{\rm nh}(p(t)).$$

However, from the definition of $p(t)$ and Eq. (3.5),

$$\dot{p}(t) = T\gamma(\dot{c}(t)) = T\gamma \circ T\pi_Q \cdot X_H(p(t)) = T(\gamma \circ \pi_Q) \cdot X_H(p(t)).$$

Therefore we get

$$X_H^{\rm nh}(p(t)) = T(\gamma \circ \pi_Q) \cdot X_H(p(t)).$$

In view of Eq. (3.9), we get, for any $w \in TT^*Q$ such that $T\pi_Q(w) \in \mathcal{D}$,

$$d(H \circ \gamma)(c(t)) \cdot T\pi_Q(w) = 0,$$

but this implies $d(H \circ \gamma)(c(t)) \cdot v = 0$ for any $v \in \mathcal{D}_{c(t)}$, or $d(H \circ \gamma)(c(t)) \in \mathcal{D}^{\circ}_{c(t)}$. However, this further implies $d(H \circ \gamma)(q) = 0$ for any $q \in Q$: For an arbitrary point $q \in Q$, consider a curve $c(t)$ that satisfies Eq. (3.5) such that $c(0) = q$. Then this gives $d(H \circ \gamma)(q) \in \mathcal{D}^{\circ}_q$. Therefore $d(H \circ \gamma) \in \mathcal{D}^{\circ}$ on Q, but then Proposition 2.4.4 implies that $d(H \circ \gamma) = 0$ because $\mathcal{D}$ is assumed to be completely nonholonomic. Therefore we have $H \circ \gamma = E$ with some constant E, which is the nonholonomic Hamilton–Jacobi equation (3.6). □

Remark 3.2.4. The condition on $d\gamma$, Eq. (3.4), stated in the above theorem is equivalent to the one in [30] as pointed out by Sosa [53] [see also 49, Lemma 4.6 on p. 51]. However Eq. (3.4) gives a simpler geometric interpretation and also is easily implemented in applications. To be specific, the condition in [30] states that there exist one-forms $\{\beta^i\}_{i=1}^m$ such that

$$d\gamma = \sum_{s=1}^{m} \beta^s \wedge \omega^s, \tag{3.11}$$

which does not easily translate into direct expressions for the conditions on γ. On the other hand, Eq. (3.4) is equivalent to

$$d\gamma(v_a, v_b) = 0 \text{ for any } a \neq b, \tag{3.12}$$

where $\{v_a\}_{a=1}^{n-m}$ spans the distribution $\mathcal{D}$. Clearly the above equations give direct expressions for the conditions on γ. We will see later in Section 3.3 that the above equations play an important role in exact integration.

Remark 3.2.5. Table 3.1 compares Theorem 3.2.1 with the unconstrained Hamilton–Jacobi theorem of Abraham and Marsden [1] (Theorem 2.3.1). Note that Eq. (3.4) is trivially satisfied for the unconstrained case: Recall that γ is replaced by an exact one-form dW in this case. Since $\mathcal{D} = TQ$ by assumption, we have $d\gamma|_{\mathcal{D}\times\mathcal{D}} = d\gamma = d(dW) = 0$ and thus this does not impose any condition on dW. Notice also that

Table 3.1: Comparison between unconstrained and nonholonomic Hamilton–Jacobi theorems.

	Nonholonomic	**Unconstrained**
Generating Function	None	$W : Q \to \mathbb{R}$
One-form	$\gamma : Q \to \mathcal{M} \subset T^*Q$	$dW : Q \to T^*Q$
Condition	$d\gamma\|_{\mathcal{D}\times\mathcal{D}} = 0$	$ddW = 0$ (trivial)
Hamilton–Jacobi Eq.	$H \circ \gamma(q) = E$	$H \circ dW(q) = E$

if $\mathcal{D} = TQ$, then the condition $d\gamma|_{\mathcal{D}\times\mathcal{D}} = 0$ implies that γ is closed, and so locally exact by the Poincaré lemma; hence the (local) existence of the generating function W such that $\gamma = dW$ follows.

Remark 3.2.6. See Cariñena et al. [14] for a Lagrangian version of Theorem 3.2.1, and de León et al. [19] for an extension to a more general framework, i.e., systems defined with linear almost Poisson structures.

3.3 Application to Exactly Integrating Equations of Motion

3.3.1 Applying the Nonholonomic Hamilton–Jacobi Theorem to Exact Integration

Theorem 3.2.1 suggests a way to use the solution of the Hamilton–Jacobi equation to integrate the equations of motion. Namely,

Step 1. Find a solution $\gamma(q)$ of the Hamilton–Jacobi equation

$$H \circ \gamma(q) = E, \tag{3.13}$$

that satisfies the conditions $\gamma(q) \in \mathcal{M}_q$ and $d\gamma|_{\mathcal{D}\times\mathcal{D}} = 0$;

Step 2. Substitute the solution $\gamma(q)$ into Eq. (3.5) to obtain the set of first-order ODEs defined in the configuration Q:

$$\dot{c}(t) = T\pi_Q \cdot X_H(\gamma \circ c(t)), \tag{3.14a}$$

or, in coordinates,

$$\dot{c}(t) = \frac{\partial H}{\partial p}(\gamma \circ c(t)); \tag{3.14b}$$

Step 3. Solve the ODEs (3.14) to find the curve $c(t)$ in the configuration space Q. Then $\gamma \circ c(t)$ gives the dynamics in the phase space T^*Q.

Figure 3.1 depicts the idea of this procedure.

In the following sections, we apply this procedure to several examples of nonholonomic systems. In any of the examples to follow, it is easy to check that the constraints are completely nonholonomic (see Definition 2.4.1), and also that the Lagrangian takes the form in Eq. (2.28) and hence the system is regular in the sense of Definition 2.4.5.

3.3.2 Examples with Separation of Variables

Let us first illustrate through a very simple example how the above procedure works with the method of separation of variables.

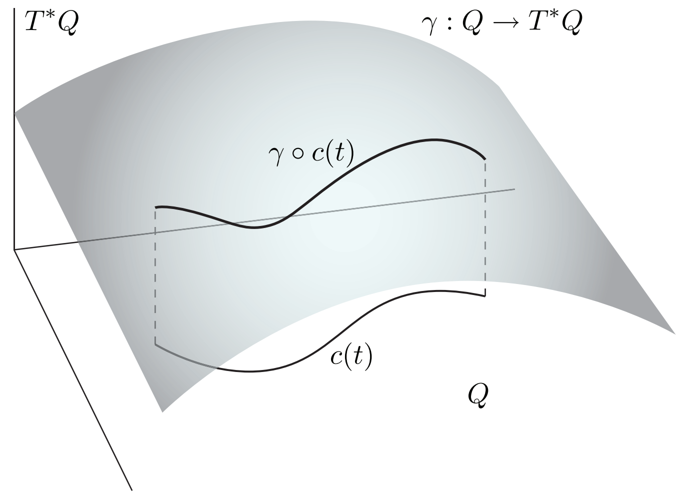

Figure 3.1: Schematic of an implication of the nonholonomic Hamilton–Jacobi theorem.

Example 3.3.1 (The vertical rolling disk; see, e.g., Bloch [7])**.** Consider the motion of the vertical rolling disk of radius R shown in Fig. 3.2. The configuration space is

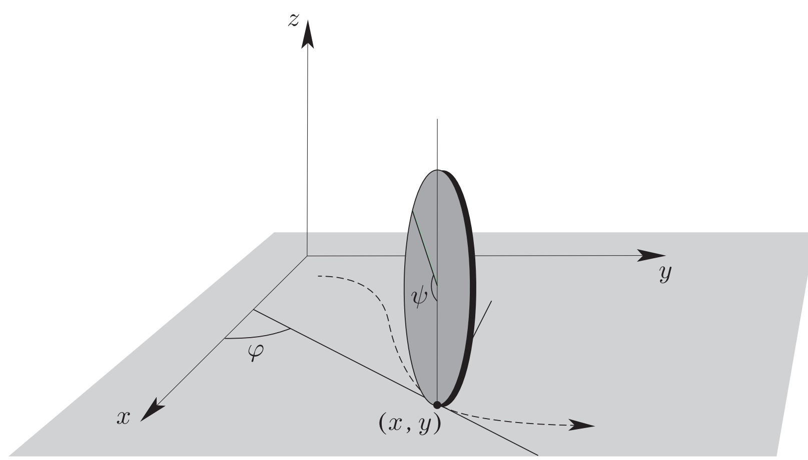

Figure 3.2: Vertical rolling disk.

$Q = SE(2) \times \mathbb{S}^1 = \{(x, y, \varphi, \psi)\}$. Suppose that m is the mass of the disk, I is the moment of inertia of the disk about the axis perpendicular to the plane of the disk, and J is the moment of inertia about an axis in the plane of the disk (both axes passing through the disk's center). The velocity constraints are

$$\dot{x} = R \cos\varphi\, \dot{\psi}, \qquad \dot{y} = R \sin\varphi\, \dot{\psi}, \tag{3.15}$$

or in terms of constraint one-forms,

$$\omega^1 = dx - R\cos\varphi\, d\psi, \qquad \omega^2 = dy - R\sin\varphi\, d\psi. \tag{3.16}$$

The Hamiltonian $H : T^*Q \to \mathbb{R}$ is given by

$$H = \frac{1}{2}\left(\frac{p_x^2 + p_y^2}{m} + \frac{p_\varphi^2}{J} + \frac{p_\psi^2}{I}\right). \tag{3.17}$$

The nonholonomic Hamilton–Jacobi equation (3.6) is

$$H \circ \gamma = E, \tag{3.18}$$

where E is a constant (the total energy). Let us construct an ansatz for Eq. (3.18). The momentum constraint $p \in \mathcal{M} = \mathbb{F}L(\mathcal{D})$ gives $p_x = mR\cos\varphi\, p_\psi/I$ and $p_y = mR\sin\varphi\, p_\psi/I$, and so we can write $\gamma : Q \to \mathcal{M}$ as

$$\gamma = \frac{mR}{I}\cos\varphi\, \gamma_\psi(x, y, \varphi, \psi)\, dx + \frac{mR}{I}\sin\varphi\, \gamma_\psi(x, y, \varphi, \psi)\, dy \\ + \gamma_\varphi(x, y, \varphi, \psi)\, d\varphi + \gamma_\psi(x, y, \varphi, \psi)\, d\psi. \tag{3.19}$$

Now we assume the following ansatz:

$$\gamma_\varphi(x, y, \varphi, \psi) = \gamma_\varphi(\varphi). \tag{3.20}$$

Then the condition $d\gamma|_{\mathcal{D}\times\mathcal{D}} = 0$ in Eq. (3.4) gives

$$\frac{\partial \gamma_\psi}{\partial \varphi} = 0, \tag{3.21}$$

and so

$$\gamma_\psi(x, y, \varphi, \psi) = \gamma_\psi(x, y, \psi). \tag{3.22}$$

So Eq. (3.18) becomes

$$\frac{1}{2}\left(\frac{\gamma_\varphi(\varphi)^2}{J} + \frac{I + mR^2}{I^2}\gamma_\psi(x, y, \psi)^2\right) = E. \tag{3.23}$$

The first term in the parentheses depends only on φ, whereas the second depends on

x, y, and ψ. This implies that both of them must be constant:

$$\gamma_{\varphi}(\varphi) = \gamma_{\varphi}^{0}, \qquad \gamma_{\psi}(x, y, \psi) = \gamma_{\psi}^{0}, \tag{3.24}$$

where γ_{φ}^{0} and γ_{ψ}^{0} are the constants determined by the initial condition such that

$$\frac{1}{2}\left(\frac{1}{J}(\gamma_{\varphi}^{0})^{2} + \frac{I + mR^{2}}{I^{2}}(\gamma_{\psi}^{0})^{2}\right) = E.$$

Then Eq. (3.5) becomes

$$\dot{x} = \frac{\gamma_{\psi}^{0} R}{I}\cos\varphi, \qquad \dot{y} = \frac{\gamma_{\psi}^{0} R}{I}\sin\varphi, \qquad \dot{\varphi} = \frac{\gamma_{\varphi}^{0}}{J}, \qquad \dot{\psi} = \frac{\gamma_{\psi}^{0}}{I}, \tag{3.25}$$

which are integrated easily to give the solution

$$\begin{aligned} x(t) &= c_1 + \frac{JR\,\gamma_{\psi}^{0}}{I\,\gamma_{\varphi}^{0}}\,\sin\!\left(\frac{\gamma_{\varphi}^{0}}{J}\,t + \varphi_0\right), \\ y(t) &= c_2 - \frac{JR\,\gamma_{\psi}^{0}}{I\,\gamma_{\varphi}^{0}}\,\cos\!\left(\frac{\gamma_{\varphi}^{0}}{J}\,t + \varphi_0\right), \\ \varphi(t) &= \varphi_0 + \frac{\gamma_{\varphi}^{0}}{J}\,t, \qquad \psi(t) = \psi_0 + \frac{\gamma_{\psi}^{0}}{I}\,t, \end{aligned} \tag{3.26}$$

where c_1, c_2, φ_0, and ψ_0 are all constants.

Separation of variables for unconstrained Hamilton–Jacobi equations often deals with problems with potential forces, e.g., a harmonic oscillator and the Kepler problem. Let us show that separation of variables works also for the following simple nonholonomic system with a potential force:

Example 3.3.2 (The knife edge; see, e.g., Bloch [7]). Consider a plane slanted at an angle α from the horizontal and let (x, y) represent the position of the point of contact of the knife edge with respect to a fixed Cartesian coordinate system on the plane (see Fig. 3.3). The configuration space is $Q = SE(2) = \{(x, y, \varphi)\}$. Suppose that the mass of the knife edge is m, and the moment of inertia about the axis perpendicular to the inclined plane through its contact point is J. The velocity constraint is

$$\sin\varphi\,\dot{x} - \cos\varphi\,\dot{y} = 0, \tag{3.27}$$

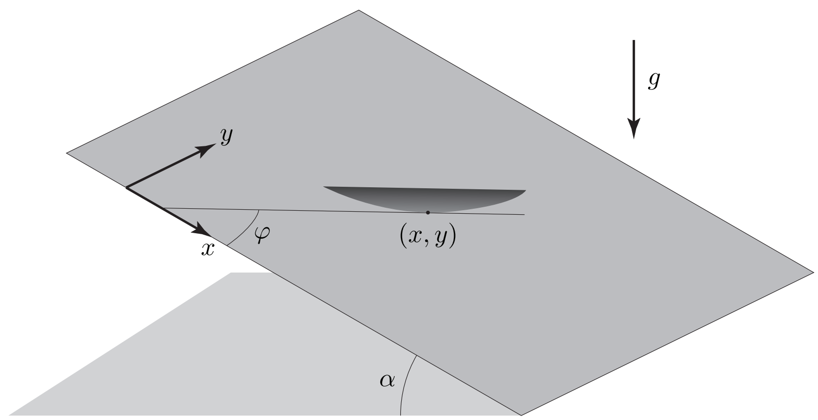

Figure 3.3: Knife edge on inclined plane.

and so the constraint one-form is

$$\omega^1 = \sin\varphi\, dx - \cos\varphi\, dy. \tag{3.28}$$

The Hamiltonian $H : T^*Q \to \mathbb{R}$ is given by

$$H = \frac{1}{2}\left(\frac{p_x^2 + p_y^2}{m} + \frac{p_\varphi^2}{J}\right) - mgx\sin\alpha. \tag{3.29}$$

The nonholonomic Hamilton–Jacobi equation (3.6) is

$$H \circ \gamma = E, \tag{3.30}$$

where E is a constant (the total energy). Let us construct an ansatz for Eq. (3.30). The momentum constraint $p \in \mathcal{M} = \mathbb{F}L(\mathcal{D})$ gives

$$p_y = \tan\varphi\, p_x,$$

and so we can write $\gamma : Q \to \mathcal{M}$ as

$$\gamma = \gamma_x(x, y, \varphi)\, dx + \tan\varphi\, \gamma_x(x, y, \varphi)\, dy + \gamma_\varphi(x, y, \varphi)\, d\varphi. \tag{3.31}$$

Now we assume the following ansatz:

$$\gamma_\varphi(x, y, \varphi) = \gamma_\varphi(\varphi). \tag{3.32}$$

Then the condition $d\gamma|_{\mathcal{D}\times\mathcal{D}} = 0$ in Eq. (3.4) gives

$$\frac{\partial \gamma_x}{\partial \varphi} = -\tan\varphi\,\gamma_x. \tag{3.33}$$

Integration of this equation yields

$$\gamma_x(x,y,\varphi) = f(x,y)\cos\varphi, \tag{3.34}$$

with some function $f(x,y)$. Then Eq. (3.30) becomes

$$\frac{1}{2}\left[\frac{f(x,y)^2}{m} - (2mg\sin\alpha)\,x + \frac{\gamma_\varphi(\varphi)^2}{J}\right] = E. \tag{3.35}$$

The first two terms in the brackets depend only on x and y, whereas the third depends only on φ. This implies that

$$\gamma_\varphi(\varphi) = \gamma_\varphi^0 \tag{3.36}$$

with some constant γ_φ^0, and $f(x,y)$ satisfies

$$\frac{1}{2}\left[\frac{f(x,y)^2}{m} - (2mg\sin\alpha)\,x + \frac{(\gamma_\varphi^0)^2}{J}\right] = E. \tag{3.37}$$

Let us suppose that sleigh is sliding downward in Fig. 3.3. Then we should have $\gamma_x \geq 0$ for $0 < \varphi < \pi/2$. From Eq. (3.34) we see that $f(x,y) \geq 0$, and hence choose the branch

$$f(x,y) = \sqrt{m\left(2E - \frac{(\gamma_\varphi^0)^2}{J}\right) + (2m^2 g\sin\alpha)\,x}. \tag{3.38}$$

Then Eq. (3.5) becomes

$$\begin{gathered} \dot{x} = \frac{\cos\varphi}{\sqrt{m/2}}\sqrt{\left(E - \frac{(\gamma_\varphi^0)^2}{2J}\right) + (mg\sin\alpha)\,x}, \\ \dot{y} = \frac{\sin\varphi}{\sqrt{m/2}}\sqrt{\left(E - \frac{(\gamma_\varphi^0)^2}{2J}\right) + (mg\sin\alpha)\,x}, \qquad \dot{\varphi} = \frac{\gamma_\varphi^0}{J}, \end{gathered} \tag{3.39}$$

Let us choose the initial condition

$$(x(0), y(0), \varphi(0), \dot{x}(0), \dot{y}(0), \dot{\varphi}(0)) = (0,0,0,0,0,\omega),$$

where $\omega := \gamma_\varphi^0/J$. Then we obtain

$$x(t) = \frac{g \sin\alpha}{2\omega^2} \sin^2(\omega t), \qquad y(t) = \frac{g \sin\alpha}{2\omega^2}\left(\omega t - \frac{1}{2}\sin(2\omega t)\right), \qquad \varphi(t) = \omega t. \tag{3.40}$$

These are the solution obtained in Bloch [7, Section 1.6].

3.3.3 Examples without Separation of Variables

In the unconstrained theory, separation of variables seems to be the only practical way of solving the Hamilton–Jacobi equation. However notice that separation of variables implies the existence of conserved quantities (or at least one) independent of the Hamiltonian, which often turn out to be the momentum maps arising from the symmetry of the system. This means that the integrability argument based on separation of variables is possible only if there are sufficient number of conserved quantities independent of the Hamiltonian [see, e.g., 42, §VIII.3]. This is consistent with the Arnold–Liouville theorem, and as a matter of fact, separation of variables can be used to identify the action-angle variables [see, e.g., 33, §6.2].

The above two examples show that we have a similar situation on the nonholonomic side as well. In each of these two examples we found conserved quantities (which are not the Hamiltonian) from the Hamilton–Jacobi equation by separation of variables as in the unconstrained theory. So again the existence of sufficient number of conserved quantities is necessary for application of separation variables. However, this condition can be more restrictive for nonholonomic systems since, for nonholonomic systems, momentum maps are replaced by momentum equations, which in general do not give conservation laws [8].

An interesting question to ask is then: What can we do when separation of variables does not seem to be working? In the unconstrained theory, there are cases where one can come up with a new set of coordinates in which one can apply separation of variables. An example is the use of elliptic coordinates in the problem of attraction by two fixed centers [3, §47.C]. The question of existence of such coordinates for nonholonomic examples is interesting to consider. However, we would like to take a different approach based on what we already have. Namely we illustrate how the nonholonomic Hamilton–Jacobi theorem can be used for those examples to which we cannot apply separation of variables. The key idea is to utilize the condition $d\gamma|_{\mathcal{D}\times\mathcal{D}} = 0$, which does not exist in the unconstrained theory as shown in Remark 3.2.5.

Example 3.3.3 (The Snakeboard; see, e.g., Bloch et al. [8]). Consider the motion of the snakeboard shown in Fig. 3.4. Let m be the total mass of the board, J

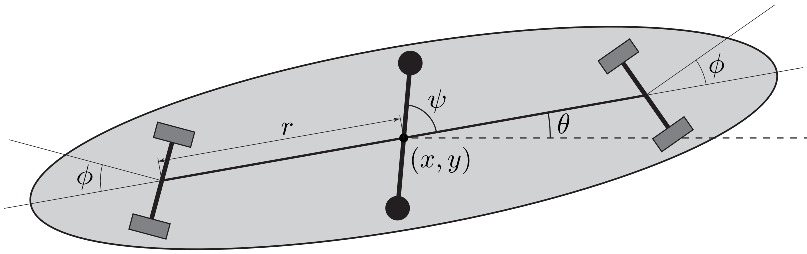

Figure 3.4: The Snakeboard.

the inertia of the board, J_0 the inertia of the rotor, J_1 the inertia of each of the wheels, and assume the relation $J + J_0 + 2J_1 = mr^2$. The configuration space is $Q = SE(2) \times \mathbb{S}^1 \times \mathbb{S}^1 = \{(x, y, \theta, \psi, \phi)\}$ and the Hamiltonian $H : T^*Q \to \mathbb{R}$ is given by

$$H = \frac{1}{2m}(p_x^2 + p_y^2) + \frac{1}{2J_0}p_\psi^2 + \frac{1}{2(mr^2 - J_0)}(p_\theta - p_\psi)^2 + \frac{1}{4J_1}p_\phi^2. \tag{3.41}$$

The velocity constraints are

$$\dot{x} + r\cot\phi\,\cos\theta\,\dot{\theta} = 0, \qquad \dot{y} + r\cot\phi\,\sin\theta\,\dot{\theta} = 0, \tag{3.42}$$

and thus the constraint distribution is written as

$$\mathcal{D} = \left\{ v = (\dot{x}, \dot{y}, \dot{\theta}, \dot{\psi}, \dot{\phi}) \in TQ \mid \omega^s(v) = 0,\ s = 1, 2 \right\}, \tag{3.43}$$

where

$$\omega^1 = dx + r\cot\phi\,\cos\theta\,d\theta, \qquad \omega^2 = dy + r\cot\phi\,\sin\theta\,d\theta. \tag{3.44}$$

The nonholonomic Hamilton–Jacobi equation (3.6) is

$$H \circ \gamma = E. \tag{3.45}$$

Let us construct an ansatz for Eq. (3.45). The momentum constraint $p \in \mathcal{M} = \mathbb{F}L(\mathcal{D})$ gives

$$p_x = -\frac{mr}{mr^2 - J_0}\cot\phi\,\cos\theta\,(p_\theta - p_\psi), \quad p_y = -\frac{mr}{mr^2 - J_0}\cot\phi\,\sin\theta\,(p_\theta - p_\psi),$$

and so we can write $\gamma : Q \to \mathcal{M}$ as

$$\gamma = -\frac{mr}{mr^2 - J_0}\cot\phi\,(\gamma_\theta - \gamma_\psi)(\cos\theta\,dx + \sin\theta\,dy) + \gamma_\theta\,d\theta + \gamma_\psi\,d\psi + \gamma_\phi\,d\phi. \tag{3.46}$$

Now we assume the following ansatz:

$$\gamma_\psi(x, y, \theta, \psi, \phi) = \gamma_\psi(\psi), \qquad \gamma_\phi(x, y, \theta, \psi, \phi) = \gamma_\phi(\phi). \tag{3.47}$$

Then the nonholonomic Hamilton–Jacobi equation (3.45) becomes

$$\frac{mr^2}{2(mr^2 - J_0)^2} \cot^2\phi\,(\gamma_\theta - \gamma_\psi)^2 + \frac{1}{2J_0}\gamma_\psi^2 + \frac{1}{2(mr^2 - J_0)}\,(\gamma_\theta - \gamma_\psi)^2 + \frac{1}{4J_1}\gamma_\phi^2 = E. \tag{3.48}$$

Solving this for γ_θ, we have

$$\gamma_\theta(x, y, \theta, \psi, \phi) = \gamma_\psi(\psi) + \frac{(mr^2 - J_0)\sin\phi}{\sqrt{(mr^2 - J_0\sin^2\phi)/2}}\sqrt{E - \frac{\gamma_\psi(\psi)^2}{2J_0} - \frac{\gamma_\phi(\phi)^2}{4J_1}} \tag{3.49}$$

and substituting the result and Eq. (3.47) into the condition $d\gamma|_{\mathcal{D}\times\mathcal{D}} = 0$ in Eq. (3.4) gives

$$\frac{d}{d\phi}\left[\gamma_\phi(\phi)^2\right] = 0,$$

$$\sin\phi\left[J_0\sqrt{E - \frac{\gamma_\psi(\psi)}{2J_0} - \frac{\gamma_\phi(\phi)}{4J_1}}\,\sin\phi - \sqrt{(mr^2 - J_0\sin^2\phi)/2}\,\gamma_\psi(\psi)\right]\gamma_\psi'(\psi) = 0.$$

Therefore it follows that

$$\gamma_\phi(\phi) = \gamma_\phi^0, \qquad \gamma_\psi(\psi) = \gamma_\psi^0$$

with some constants γ_ϕ^0 and γ_ψ^0. Hence Eq. (3.49) becomes

$$\gamma_\theta(x, y, \theta, \psi, \phi) = \gamma_\psi(\psi) + \frac{(mr^2 - J_0)\,C\sin\phi}{g(\phi)},$$

where we defined

$$C := \sqrt{E - \frac{(\gamma_\psi^0)^2}{2J_0} - \frac{(\gamma_\phi^0)^2}{4J_1}}, \qquad g(\phi) := \sqrt{(mr^2 - J_0\sin^2\phi)/2}.$$

Then Eq. (3.5) gives

$$\begin{aligned}
&\dot{x} = -\frac{C\,r\cos\theta\,\cos\phi}{g(\phi)}, \qquad \dot{y} = -\frac{C\,r\sin\theta\,\cos\phi}{g(\phi)}, \\
&\dot{\theta} = \frac{C\sin\phi}{g(\phi)}, \qquad \dot{\psi} = \frac{\gamma_\psi^0}{J_0} - \frac{C\sin\phi}{g(\phi)}, \qquad \dot{\phi} = \frac{\gamma_\phi^0}{2J_1}.
\end{aligned} \tag{3.50}$$

This result is consistent with that of Koon and Marsden [38] obtained by reduction of Hamilton's equations for nonholonomic systems. It is also clear from the above expressions that the solution is obtained by a quadrature.

In the above example we found conserved quantities through $d\gamma|_{\mathcal{D}\times\mathcal{D}} = 0$ instead of separation of variables. In the following example, we cannot identify conserved quantities even through $d\gamma|_{\mathcal{D}\times\mathcal{D}} = 0$; nevertheless we can still integrate the equations of motion.

Example 3.3.4 (The Chaplygin sleigh; see, e.g., Bloch [7])**.** Consider the motion of the Chaplygin sleigh shown in Fig. 3.5. Let m be the mass, I the moment of inertia

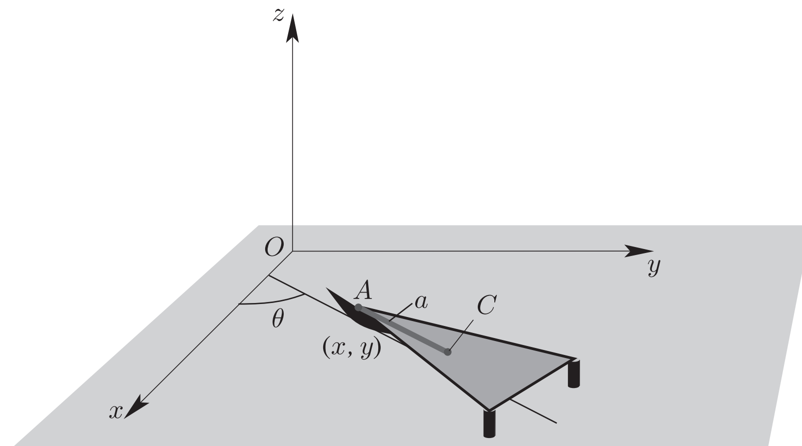

Figure 3.5: The Chaplygin sleigh.

about the center of mass C, a be the distance from the center of mass C to the contact point A of the edge. The configuration space is $Q = SE(2) = \{(x, y, \theta)\}$, where the coordinates (x, y) give the position of the contact point of the edge (not the center of mass). The velocity constraint is

$$\sin\theta\,\dot{x} - \cos\theta\,\dot{y} = 0, \tag{3.51}$$

and so the constraint one-form is

$$\omega^1 = \sin\theta\, dx - \cos\theta\, dy. \tag{3.52}$$

The Hamiltonian $H : T^*Q \to \mathbb{R}$ is given by

$$\begin{aligned} H = \frac{Ma^2\sin^2\theta + J}{2JM}\, p_x^2 + \frac{Ma^2\cos^2\theta + J}{2JM}\, p_y^2 \\ + \frac{1}{2J}\, p_\theta^2 - \frac{a^2\sin\theta\cos\theta}{J}\, p_x\, p_y + \frac{a}{J}(\sin\theta\, p_x - \cos\theta\, p_y)\, p_\theta. \end{aligned} \tag{3.53}$$

The nonholonomic Hamilton–Jacobi equation (3.6) is

$$H \circ \gamma = E, \tag{3.54}$$

where E is a constant (the total energy). Let us construct an ansatz for Eq. (3.54). The momentum constraint $p \in \mathcal{M} = \mathbb{F}L(\mathcal{D})$ gives

$$p_y = \tan\theta\, p_x + \frac{aM\sec\theta}{J + a^2M}\, p_\theta,$$

and so we can write $\gamma : Q \to \mathcal{M}$ as

$$\gamma = \gamma_x(x,y,\theta)\, dx + \left[\tan\theta\, \gamma_x(x,y,\theta) + \frac{aM\sec\theta}{J + a^2M}\, \gamma_\theta(x,y,\theta)\right] dy + \gamma_\theta(x,y,\theta)\, d\theta. \tag{3.55}$$

Now we assume the following ansatz:

$$\gamma_\theta(x,y,\theta) = \gamma_\theta(\theta). \tag{3.56}$$

Then the condition $d\gamma|_{\mathcal{D}\times\mathcal{D}} = 0$ in Eq. (3.4) gives

$$(J + a^2M)\sec\theta\left(\frac{\partial\gamma_x}{\partial\theta} + \tan\theta\, \gamma_x\right) + aM\tan\theta\left(\frac{d\gamma_\theta}{d\theta} + \tan\theta\, \gamma_\theta\right) = 0. \tag{3.57}$$

On the other hand, the Hamilton–Jacobi equation (3.54) becomes

$$\begin{aligned} \frac{1}{4}\sec\theta\left[\frac{2\sec\theta}{M}\, \gamma_x(x,y,\theta)^2 + \frac{4a\tan\theta}{J + a^2M}\, \gamma_x(x,y,\theta)\, \gamma_\theta(\theta) \right. \\ \left. + \frac{(J + 2a^2M + J\cos 2\theta)\sec\theta}{(J + a^2M)^2}\, \gamma_\theta(\theta)^2\right] = E. \end{aligned} \tag{3.58}$$

It is impossible to separate the variables as we did in the examples in Examples 3.3.1

and 3.3.3, since we cannot isolate the terms that depend only on θ. Instead we solve the above equation for γ_x and substitute the result into Eq. (3.57). Then we obtain

$$\frac{d\gamma_\theta}{d\theta} = -a\sqrt{M\left(2E - \frac{\gamma_\theta^2}{J + a^2 M}\right)}.$$

Solving this ODE gives

$$\gamma_\theta(\theta) = (J + a^2 M)\,\omega\cos\left(\sqrt{\frac{a^2 M}{J + a^2 M}}\,\theta\right), \tag{3.59}$$

where we assumed that $x'(0) = y'(0) = 0$, $\theta(0) = 0$, and $\theta'(0) = \omega$ and also that $|\theta(t)| < \pi/2$; note that the angular velocity ω is related to the total energy by the equation $E = (J + a^2 M)\,\omega/2$. Then the equation for $\theta(t)$ in Eq. (3.5) becomes

$$\dot{\theta} = \omega\cos\left(\sqrt{\frac{a^2 M}{J + a^2 M}}\,\theta\right), \tag{3.60}$$

which, with $\theta(0) = 0$, gives

$$\theta(t) = \frac{2}{b}\arctan\left[\tanh\left(\frac{b}{2}\,\omega t\right)\right], \tag{3.61}$$

where we set $b := \sqrt{a^2 M/(J + a^2 M)}$. Substituting this back into Eq. (3.59), we obtain

$$\gamma_\theta(t) = (J + a^2 M)\,\omega\,\mathrm{sech}\left(\sqrt{\frac{a^2 M}{J + a^2 M}}\,\omega t\right), \tag{3.62}$$

which is the solution obtained by Bloch [6] [see also 7, Section 8.6].

Chapter 4

Chaplygin Hamiltonization and Nonholonomic Hamilton–Jacobi Theory

4.1 Introduction

This chapter approaches nonholonomic Hamilton–Jacobi theory from a different perspective from the previous chapter, and also establishes a link between those two approaches. Specifically, we first employ the technique called the Chaplygin Hamiltonization to transform a certain class of nonholonomic systems into Hamiltonian systems, and then apply the conventional Hamilton–Jacobi theory to the resulting Hamiltonian systems to obtain what we would like to call the *Chaplygin Hamilton–Jacobi equation*. The main result in this chapter is an explicit formula that relates the solutions of the Chaplygin Hamilton–Jacobi equation with those of the nonholonomic Hamilton–Jacobi equation in the previous chapter.

4.1.1 Direct vs. Indirect Approaches

The indirect approach via Chaplygin Hamiltonization has both advantages and disadvantages. The main advantage is that we have a conventional Hamilton–Jacobi equation and thus the separation of variables argument applies in a rather straightforward manner compared to the direct approach in the previous chapter. A disadvantage is that the Chaplygin Hamiltonization works only for limited nonholonomic systems; and even if it does, the relationship between the Hamilton–Jacobi equation and the original nonholonomic system is not transparent, since one has to inverse-transform the information in the Hamiltonized systems. Nevertheless Hamiltonization is known to be a powerful technique of integration of nonholonomic systems [9; 16; 23–25], and hence it is interesting to make a connection between the approach in the previous chapter and the one with Hamiltonization.

Let us briefly summarize the differences between two approaches. Recall from the previous chapter that nonholonomic Hamilton–Jacobi theory gives the following set of equations for a one-form $\gamma : Q \to \mathcal{M} \subset T^*Q$ defined on the original configuration space Q:

$$H \circ \gamma = E, \qquad d\gamma|_{\mathcal{D}\times\mathcal{D}} = 0,$$

where the Hamiltonian H is a function on T^*Q. On the other hand, the Chaplygin Hamiltonization first reduces the system by identifying it as a so-called Chaplygin system with a symmetry group G, and then Hamiltonize the system on the cotangent bundle $T^*(Q/G)$ of the reduced configuration space Q/G. The resulting Chaplygin Hamilton–Jacobi equation is an equation for a function $\bar{W} : Q/G \to \mathbb{R}$:

$$\bar{H}_{\mathrm{C}} \circ d\bar{W} = E,$$

with another Hamiltonian $\bar{H}_{\mathrm{C}}$ defined on $T^*(Q/G)$. Therefore the difference lies not only in the forms of the equations (former one involves a one-form that is not even closed, whereas the latter an exact one-form), but also in the spaces on which the equations are defined. Furthermore, the Chaplygin Hamilton–Jacobi equation corresponds to the Hamiltonized dynamics and is related to the original nonholonomic one in a rather indirect way. Therefore, on the surface, there does not seem to be an apparent relationship between the two approaches.

4.1.2 Main Results

The main goal of this chapter is to establish a link between the two distinct approaches to nonholonomic Hamilton–Jacobi theory. To that end, we first formulate the Chaplygin Hamiltonization in an intrinsic manner to elucidate the geometry involved in the Hamiltonization. This gives a slight generalization of the Chaplygin Hamiltonization by Fedorov and Jovanović [24] and also an intrinsic account of the necessary and sufficient condition for symplectizing a certain class of nonholonomic systems. These results are also related to existence of invariant measure in nonholonomic systems. We then identify a sufficient condition for the Chaplygin Hamiltonization. The sufficient condition turns out to be identical to one of those for another kind of Hamiltonization (which renders the systems "conformal symplectic" [29]) obtained by Stanchenko [54] and Cantrijn et al. [12]. We then give an explicit formula that translates the solutions of the Chaplygin Hamilton–Jacobi equation into those of the nonholonomic Hamilton–Jacobi equation. Interestingly, it turns out that the sufficient condition plays an important role here as well. We show, through a couple of examples, that the solutions of the Chaplygin Hamilton–Jacobi equation are, through the formula, identical to those obtained by the direct approach in the previous chapter.

4.1.3 Outline

We first review, in Section 4.2, the so-called Chaplygin systems and their reduction following Koiller [36], Ehlers et al. [21], and Hochgerner and García-Naranjo [29]. Section 4.3 treats the Chaplygin Hamiltonization of such systems intrinsically, making links with existence of invariant measures, and also identifies the necessary and sufficient condition for symplectization and a sufficient condition for the Chaplygin Hamiltonization. In Section 4.4, we formulate the Chaplygin Hamilton–Jacobi equation and give an explicit formula that relates the solutions of it to those of the nonholonomic Hamilton–Jacobi equation (3.6). Section 4.5 takes two examples, the vertical rolling disk and knife edge, to illustrate that the formula, combined with separation of variables for the Chaplygin Hamilton–Jacobi equation, gives the solutions obtained in Examples 3.3.1 and 3.3.2 in the previous chapter.

4.2 Chaplygin Systems

Consider a nonholonomic system on an n-dimensional configuration manifold Q with a constraint distribution $\mathcal{D} \subset TQ$ with $\dim \mathcal{D}_q = n - m$, and also with the Lagrangian $L : TQ \to \mathbb{R}$ of the form

$$L(v_q) = \frac{1}{2} g_q(v_q, v_q) - V(q) \tag{4.1}$$

with the kinetic energy metric g defined on Q. Define the Legendre transform $\mathbb{F}L : TQ \to T^*Q$ by

$$\langle \mathbb{F}L(v_q), w_q \rangle = \left. \frac{d}{d\varepsilon} L(v_q + \varepsilon\, w_q) \right|_{\varepsilon=0} = g_q(v_q, w_q) = \left\langle g_q^\flat(v_q), w_q \right\rangle$$

for any $v_q, w_q \in T_qQ$. where the last equality defines $g^\flat : TQ \to T^*Q$; hence we have $\mathbb{F}L = g^\flat$. Also define the Hamiltonian $H : T^*Q \to \mathbb{R}$ by

$$H(p_q) := \langle p_q, v_q \rangle - L(v_q),$$

where $v_q = (\mathbb{F}L)^{-1}(p_q)$ on the right-hand side.

Suppose that the system is a so-called *Chaplygin system*: Consider a free and proper group action of a Lie group G on Q, i.e., we have $\Phi : G \times Q \to Q$ or $\Phi_h : Q \to Q$ for any $h \in G$; we assume that the system has a symmetry under the group action, and also that each tangent space is the direct sum of the the constraint distribution and the tangent space to the orbit of the group action, i.e., we have, for any $q \in Q$,

$$T_qQ = \mathcal{D}_q \oplus T_q\mathcal{O}_q, \tag{4.2}$$

where $\mathcal{O}_q$ is the orbit through q of the G-action on Q, i.e.,

$$\mathcal{O}_q = \{\Phi_g(q) \in Q \mid g \in G\}.$$

Therefore the dimension of the Lie group G must be m. This setup gives rise to the principal bundle $\pi : Q \to Q/G =: \bar{Q}$ and the connection $\mathcal{A} : TQ \to \mathfrak{g}$, with $\mathfrak{g}$ being the Lie algebra of G, such that $\ker \mathcal{A} = \mathcal{D}$. So the above decomposition is now written as

$$T_qQ = \ker \mathcal{A}_q \oplus \ker T_q\pi, \tag{4.3}$$

and any vector $v_q \in T_qQ$ can be decomposed into the horizontal and vertical parts:

$$v_q = \mathrm{hor}(v_q) + \mathrm{ver}(v_q), \tag{4.4a}$$

with

$$\mathrm{hor}(v_q) = v_q - (\mathcal{A}_q(v_q))_Q(q), \qquad \mathrm{ver}(v_q) = (\mathcal{A}_q(v_q))_Q(q), \tag{4.4b}$$

where $\xi_Q \in \mathfrak{X}(Q)$ is the infinitesimal generator of $\xi \in \mathfrak{g}$. Furthermore, for any $q \in Q$ and $\bar{q} := \pi(q) \in \bar{Q}$, the map $T_q\pi|_{\mathcal{D}_q} : \mathcal{D}_q \to T_{\bar{q}}\bar{Q}$ is a linear isomorphism, and hence we have the horizontal lift

$$\mathrm{hl}_q^{\mathcal{D}} : T_{\bar{q}}\bar{Q} \to \mathcal{D}_q; \quad \bar{v}_{\bar{q}} \mapsto (T_q\pi|_{\mathcal{D}_q})^{-1}(\bar{v}_{\bar{q}}). \tag{4.5}$$

We employ the following shorthand notation for horizontal lifts:

$$v_q^{\mathrm{h}} := \mathrm{hl}_q^{\mathcal{D}}(\bar{v}_{\bar{q}}). \tag{4.6}$$

This gives rise to the reduced Lagrangian $\bar{L} := L \circ \mathrm{hl}^{\mathcal{D}}$, or more explicitly,

$$\bar{L} : T\bar{Q} \to \mathbb{R}; \quad \bar{v}_{\bar{q}} \mapsto \frac{1}{2}\bar{g}_{\bar{q}}(\bar{v}_{\bar{q}}, \bar{v}_{\bar{q}}) - \bar{V}(\bar{q}), \tag{4.7}$$

where $\bar{g}$ is the metric on the reduced space $\bar{Q}$ induced by g as follows:

$$\bar{g}_{\bar{q}}(\bar{v}_{\bar{q}}, \bar{w}_{\bar{q}}) := g_q\big(\mathrm{hl}_q^{\mathcal{D}}(\bar{v}_{\bar{q}}), \mathrm{hl}_q^{\mathcal{D}}(\bar{v}_{\bar{q}})\big) = g_q(v_q^{\mathrm{h}}, w_q^{\mathrm{h}}), \tag{4.8}$$

and the reduced potential $\bar{V} : \bar{Q} \to \mathbb{R}$ is defined such that $V = \bar{V} \circ \pi$.

This geometric structure is carried over to the Hamiltonian side (see Ehlers et al. [21]). Specifically, we define the horizontal lift $\mathrm{hl}_q^{\mathcal{M}} : T_{\bar{q}}^*\bar{Q} \to \mathcal{M}_q$ by

$$\mathrm{hl}_q^{\mathcal{M}} := \mathbb{F}L_q \circ \mathrm{hl}_q^{\mathcal{D}} \circ (\mathbb{F}\bar{L})_{\bar{q}}^{-1} = g_q^{\flat} \circ \mathrm{hl}_q^{\mathcal{D}} \circ (\bar{g}^{\flat})_{\bar{q}}^{-1}, \tag{4.9}$$

or the diagram below commutes.

$$\begin{tikzcd}
\mathcal{D}_q \arrow[r, "\mathbb{F}L_q"] & \mathcal{M}_q \\
T_{\bar{q}}\bar{Q} \arrow[u, "\mathrm{hl}_q^{\mathcal{D}}"] & T_{\bar{q}}^*\bar{Q} \arrow[l, "(\mathbb{F}\bar{L})_{\bar{q}}^{-1}"] \arrow[u, dashed, "\mathrm{hl}_q^{\mathcal{M}}"']
\end{tikzcd} \tag{4.10}$$

Again we employ the following shorthand notation:

$$\alpha_q^{\mathrm{h}} := \mathrm{hl}_q^{\mathcal{M}}(\bar{\alpha}_{\bar{q}}) \tag{4.11}$$

for any $\bar{\alpha}_{\bar{q}} \in T_{\bar{q}}^*\bar{Q}$.

We also define the reduced Hamiltonian $\bar{H} : T^*\bar{Q} \to \mathbb{R}$ by

$$\bar{H} := H \circ \mathrm{hl}^{\mathcal{M}}; \tag{4.12}$$

it is easy to check that this definition coincides with the following one by using the reduced Lagrangian $\bar{L}$:

$$\bar{H}(p_{\bar{q}}) := \langle p_{\bar{q}}, v_{\bar{q}} \rangle - \bar{L}(v_{\bar{q}}),$$

with $v_{\bar{q}} = (\mathbb{F}\bar{L})_{\bar{q}}^{-1}(p_{\bar{q}})$.

Performing the nonholonomic reduction of Koiller [36] (see also Bates and Sniatycki [4], Ehlers et al. [21], and Hochgerner and García-Naranjo [29]), we obtain the following reduced Hamilton's equations for Chaplygin systems:

$$i_{\bar{X}}\bar{\Omega} = d\bar{H} + i_{\bar{X}}\Xi, \tag{4.13}$$

where $\bar{X}$ is a vector field on $T^*\bar{Q}$ and $\bar{\Omega}$ is the standard symplectic form on $T^*\bar{Q}$; the two-form Ξ on $T^*\bar{Q}$ is defined as follows: For any $\bar{\alpha}_{\bar{q}} \in T^*_{\bar{q}}\bar{Q}$ and $Y_{\bar{\alpha}_{\bar{q}}}, Z_{\bar{\alpha}_{\bar{q}}} \in T_{\bar{\alpha}_{\bar{q}}}T^*\bar{Q}$, let $\bar{Y}_{\bar{q}} := T\pi_{\bar{Q}}(Y_{\bar{\alpha}_{\bar{q}}})$ and $\bar{Z}_{\bar{q}} := T\pi_{\bar{Q}}(Z_{\bar{\alpha}_{\bar{q}}})$ where $\pi_{\bar{Q}} : T^*\bar{Q} \to \bar{Q}$ is the cotangent bundle projection, and then set

$$\begin{aligned}
\Xi_{\bar{\alpha}_{\bar{q}}}(Y_{\bar{\alpha}_{\bar{q}}}, Z_{\bar{\alpha}_{\bar{q}}}) &:= \left\langle \mathbf{J} \circ \mathrm{hl}_q^{\mathcal{M}}(\bar{\alpha}_{\bar{q}}), \mathcal{B}_q\big(\mathrm{hl}_q^{\mathcal{D}}(\bar{Y}_{\bar{q}}), \mathrm{hl}_q^{\mathcal{D}}(\bar{Z}_{\bar{q}})\big) \right\rangle \\
&= \left\langle \mathbf{J}(\alpha_q^{\mathrm{h}}), \mathcal{B}_q(Y_q^{\mathrm{h}}, Z_q^{\mathrm{h}}) \right\rangle,
\end{aligned} \tag{4.14}$$

where $\mathbf{J} : T^*Q \to \mathfrak{g}^*$ is the momentum map corresponding to the G-action, and $\mathcal{B}$ is the curvature two-form of the connection $\mathcal{A}$. This is well-defined, since the Ad-equivariant properties of the momentum map $\mathbf{J}$ and the curvature $\mathcal{B}$ cancel each other [37]: Writing $hq := \Phi_h(q)$ for any $h \in G$, we have, using Lemma 4.A.1 and the G-equivariance of the momentum map $\mathbf{J}$ and the curvature $\mathcal{B}$,

$$\begin{aligned}
\left\langle \mathbf{J}(\alpha_{hq}^{\mathrm{h}}), \mathcal{B}_{hq}(Y_{hq}^{\mathrm{h}}, Z_{hq}^{\mathrm{h}}) \right\rangle &= \left\langle \mathbf{J}\big(T_q^*\Phi_{h^{-1}}(\alpha_q^{\mathrm{h}})\big), \Phi_h^*\mathcal{B}_q(Y_q^{\mathrm{h}}, Z_q^{\mathrm{h}}) \right\rangle \\
&= \left\langle \mathrm{Ad}_{h^{-1}}^* \mathbf{J}(\alpha_q^{\mathrm{h}}), \mathrm{Ad}_h \mathcal{B}_q(Y_q^{\mathrm{h}}, Z_q^{\mathrm{h}}) \right\rangle \\
&= \left\langle \mathbf{J}(\alpha_q^{\mathrm{h}}), \mathcal{B}_q(Y_q^{\mathrm{h}}, Z_q^{\mathrm{h}}) \right\rangle.
\end{aligned}$$

4.3 Chaplygin Hamiltonization of Nonholonomic Systems

This section discusses the so-called Hamiltonization of the reduced dynamics defined by Eq. (4.13). The results here are mostly a summary of some of the key results of Stanchenko [54], Cantrijn et al. [12], Fedorov and Jovanović [24], and Fernandez et al. [25]. However, our exposition is slightly different from them, and also touches on those aspects that are not found in the above papers.

4.3.1 Hamiltonization and Existence of Invariant Measure

We first discuss the relationship between Hamiltonization and existence of invariant measure for nonholonomic systems. The next subsection will show how to Hamiltonize the reduced system Eq. (4.13) explicitly.

Let $f : T^*\bar{Q} \to \mathbb{R}$ be a smooth function with $f > 0$ and also constant on each fiber, i.e., $f(\alpha_{\bar{q}}) = f(\beta_{\bar{q}})$ for any $\alpha_{\bar{q}}, \beta_{\bar{q}} \in T^*_{\bar{q}}\bar{Q}$. Therefore we can write, with a slight abuse of notation, $f(\alpha_{\bar{q}}) = f(\bar{q})$; so f may be seen as a function on $\bar{Q}$. Now consider the vector field

$$\bar{X}/f = \frac{1}{f}\bar{X} \in \mathfrak{X}(T^*\bar{Q}),$$

and let $\Phi_t^{\bar{X}/f} : T^*\bar{Q} \to T^*\bar{Q}$ be the flow defined by this vector field, i.e., for any $\alpha_{\bar{q}} \in T^*\bar{Q}$,

$$\left.\frac{d}{dt}\Phi_t^{\bar{X}/f}(\alpha_{\bar{q}})\right|_{t=0} = (\bar{X}/f)(\alpha_{\bar{q}}) = \frac{1}{f(\alpha_{\bar{q}})}\,\bar{X}(\alpha_{\bar{q}}).$$

Now consider the map $\Psi_f : T^*\bar{Q} \to T^*\bar{Q}$ defined by

$$\Psi_f : \alpha \mapsto f\alpha,$$

which is clearly a diffeomorphism with the inverse $\Psi_f^{-1} = \Psi_{1/f} : T^*\bar{Q} \to T^*\bar{Q};\ \alpha \mapsto \alpha/f$, and define $\Phi_t^{\bar{Y}} : T^*\bar{Q} \to T^*\bar{Q}$ by

$$\Phi_t^{\bar{Y}} := \Psi_f \circ \Phi_t^{\bar{X}/f} \circ \Psi_f^{-1} = \Psi_f \circ \Phi_t^{\bar{X}/f} \circ \Psi_{1/f},$$

or the diagram below commutes.

$$\begin{tikzcd}
T^*\bar{Q} \arrow[r, "\Phi_t^{\bar{X}/f}"] & T^*\bar{Q} \arrow[d, "\Psi_f"] & \alpha/f \arrow[r, mapsto] & \Phi_t^{\bar{X}/f}(\alpha/f) \arrow[d, mapsto] \\
T^*\bar{Q} \arrow[u, "\Psi_f^{-1}=\Psi_{1/f}"] \arrow[r, dashed, "\Phi_t^{\bar{Y}}"'] & T^*\bar{Q} & \alpha \arrow[u, mapsto] \arrow[r, dashed, mapsto] & \Phi_t^{\bar{Y}}(\alpha)
\end{tikzcd} \tag{4.15}$$

Then we have the vector filed $\bar{Y} \in \mathfrak{X}(T^*\bar{Q})$ corresponding to the flow $\Phi_t^{\bar{Y}}$, which is the pull-back of $\bar{X}/f$ by $\Psi_f^{-1} = \Psi_{1/f}$:

$$\begin{aligned}
\bar{Y}(\alpha_{\bar{q}}) &:= \left.\frac{d}{dt}\Phi_t^{\bar{Y}}(\alpha_{\bar{q}})\right|_{t=0} \\
&= \left.\frac{d}{dt}\Psi_f \circ \Phi_t^{\bar{X}/f} \circ \Psi_f^{-1}(\alpha_{\bar{q}})\right|_{t=0} \\
&= T\Psi_f \cdot (\bar{X}/f)(\Psi_f^{-1}(\alpha_{\bar{q}})) \\
&= (\Psi_f^{-1})^*(\bar{X}/f)(\alpha_{\bar{q}}) \\
&= \Psi_{1/f}^*(\bar{X}/f)(\alpha_{\bar{q}}),
\end{aligned} \tag{4.16}$$

for any $\alpha_{\bar{q}} \in T^*\bar{Q}$; notice also that the third line in the above equation shows that $\bar{X}/f$ and $\bar{Y}$ are Ψ_f-related:

$$T\Psi_f \circ (\bar{X}/f) = \bar{Y} \circ \Psi_f.$$

The following theorem relates the symplecticity of the vector field $\bar{Y}$ and the existence of an invariant measure for the reduced system Eq. (4.13):

Theorem 4.3.1. *If $\bar{Y} \in \mathfrak{X}(T^*\bar{Q})$ is symplectic, i.e., $\pounds_{\bar{Y}}\bar{\Omega} = 0$, then the reduced system Eq. (4.13) has the invariant measure $f^{\bar{n}-1}\bar{\Lambda}$, where $\bar{n} := \dim \bar{Q} = n - m$ and $\bar{\Lambda}$ being the Liouville volume form*

$$\bar{\Lambda} := \frac{(-1)^{\bar{n}(\bar{n}-1)/2}}{\bar{n}!}\underbrace{\bar{\Omega} \wedge \cdots \wedge \bar{\Omega}}_{\bar{n}} = dq^1 \wedge \cdots \wedge dq^{\bar{n}} \wedge dp_1 \wedge \cdots \wedge dp_{\bar{n}}.$$

In other words, we have

$$\pounds_{\bar{X}}(f^{\bar{n}-1}\bar{\Lambda}) = 0.$$

This theorem is a slight generalization of the following:

Corollary 4.3.2 (Fedorov and Jovanović [24])**.** *If $\bar{Y} \in \mathfrak{X}(T^*\bar{Q})$ is Hamiltonian, i.e., $i_{\bar{Y}}\bar{\Omega} = d\bar{H}_{\mathrm{C}}$ for some $\bar{H}_{\mathrm{C}} : T^*\bar{Q} \to \mathbb{R}$, then the reduced nonholonomic dynamics Eq. (4.13) has the invariant measure $f^{\bar{n}-1}\bar{\Lambda}$.*

Proof. Follows easily from Cartan's formula:

$$\pounds_{\bar{Y}}\bar{\Omega} = d(i_{\bar{Y}}\bar{\Omega}) + i_{\bar{Y}}d\bar{\Omega} = dd\bar{H}_{\mathrm{C}} = 0. \qquad \square$$

Definition 4.3.3. We would like to call such $\bar{H}_{\mathrm{C}} : T^*\bar{Q} \to \mathbb{R}$ a *Chaplygin Hamiltonian.*

We state a couple of lemmas before proving Theorem 4.3.1.

Lemma 4.3.4. *Let $f : T^*\bar{Q} \to \mathbb{R}$ be a smooth function that is constant on the fibers, i.e., $f(\alpha_{\bar{q}}) = f(\beta_{\bar{q}})$ for any $\alpha_{\bar{q}}, \beta_{\bar{q}} \in T^*_{\bar{q}}\bar{Q}$. Then*

$$\Psi_f^*\bar{\Omega} = f\bar{\Omega} - df \wedge \bar{\Theta},$$

and

$$\underbrace{(\Psi_f^*\bar{\Omega}) \wedge \cdots \wedge (\Psi_f^*\bar{\Omega})}_{\bar{n}} = f^{\bar{n}}\, \underbrace{\bar{\Omega} \wedge \cdots \wedge \bar{\Omega}}_{\bar{n}}.$$

Proof. Let $\bar{\Theta}$ be the symplectic one-form on $T^*\bar{Q}$, i.e., $\bar{\Omega} = -d\bar{\Theta}$. Let us first calculate $\Psi_f^*\bar{\Theta}$: We have, for any $\alpha \in T^*\bar{Q}$ and $v \in T_\alpha T^*\bar{Q}$,

$$\begin{aligned}
(\Psi_f^*\bar{\Theta})_\alpha(v) &= \bar{\Theta}_{\Psi_f(\alpha)}(T\Psi_f(v)) \\
&= \langle \Psi_f(\alpha), T\pi_Q \cdot T\Psi_f(v) \rangle \\
&= \langle f\alpha, T(\pi_Q \circ \Psi_f)(v) \rangle \\
&= f\, \langle \alpha, T\pi_Q(v) \rangle \\
&= f\, \bar{\Theta}_\alpha(v),
\end{aligned}$$

where we used the fact that Ψ_f is fiber-preserving, i.e., $\pi_Q \circ \Psi_f = \pi_Q$. Hence we have $\Psi_f^*\bar{\Theta} = f\bar{\Theta}$, and thus

$$\begin{aligned}
\Psi_f^*\bar{\Omega} &= \Psi_f^*(-d\bar{\Theta}) \\
&= -d(\Psi_f^*\bar{\Theta}) \\
&= -d(f\bar{\Theta}) \\
&= -df \wedge \bar{\Theta} - f d\bar{\Theta} \\
&= f\bar{\Omega} - df \wedge \bar{\Theta}.
\end{aligned}$$

Therefore, using the fact that $\alpha \wedge \beta = \beta \wedge \alpha$ for any two-forms α and β, we have

$$\begin{aligned}
(\Psi_f^*\bar{\Omega}) \wedge \cdots \wedge (\Psi_f^*\bar{\Omega}) &= f^{\bar{n}}\bar{\Omega} \wedge \cdots \wedge \bar{\Omega} \\
&\quad + \sum_{k=1}^{\bar{n}} \binom{\bar{n}}{k} (-1)^k f^{\bar{n}-k}\, \underbrace{\bar{\Omega} \wedge \cdots \wedge \bar{\Omega}}_{\bar{n}-k} \wedge \underbrace{(df \wedge \bar{\Theta}) \wedge \cdots \wedge (df \wedge \bar{\Theta})}_{k}.
\end{aligned}$$

Let us show that the second term vanishes. Since f is constant on fibers, we have

$$df = \frac{\partial f}{\partial q^a}\, dq^a.$$

Therefore

$$df \wedge \bar{\Theta} = p_b \frac{\partial f}{\partial q^a}\, dq^a \wedge dq^b$$

and thus $df \wedge \bar{\Theta}$ does not contain any term with dp_a's. On the other hand, $\underbrace{\bar{\Omega} \wedge \cdots \wedge \bar{\Omega}}_{\bar{n}-k}$ contains only $\bar{n} - k$ of dp_a's. Therefore the $2\bar{n}$-form

$$\underbrace{\bar{\Omega} \wedge \cdots \wedge \bar{\Omega}}_{\bar{n}-k} \wedge \underbrace{(df \wedge \bar{\Theta}) \wedge \cdots \wedge (df \wedge \bar{\Theta})}_{k}$$

contains only $\bar{n} - k$ of dp_a's, and thus $\bar{n} + k$ of dq^a's, which implies that this $2\bar{n}$-form must vanish. □

Lemma 4.3.5. *Let M be a differentiable manifold, μ be a volume form on M, X a vector field on M, and $f \in C^\infty(M)$ a positive function. Then the following identity holds:*

$$\operatorname{div}_{f\mu}(X) = \operatorname{div}_\mu(fX). \tag{4.17}$$

Proof. The following identities hold [see, e.g., 1, Proposition 2.5.23 on p. 130]:

$$\operatorname{div}_{f\mu}(X) = \operatorname{div}_\mu(X) + \frac{1}{f} X[f], \qquad \operatorname{div}_\mu(fX) = f \operatorname{div}_\mu(X) + X[f].$$

Multiplying the first identity by f and taking the difference of both sides, we have

$$f \operatorname{div}_{f\mu}(X) = f \operatorname{div}_\mu(fX).$$

The claim follows since f is positive. □

Proof of Theorem 4.3.1. As shown in Eq. (4.3.1), the vector fields $\bar{X}/f$ and $\bar{Y}$ are Ψ_f-related. Therefore

$$\pounds_{\bar{X}/f}(\Psi_f^* \bar{\Omega}) = \Psi_f^* \pounds_{\bar{Y}} \bar{\Omega} = 0,$$

and thus

$$\pounds_{\bar{X}/f}[\underbrace{(\Psi_f^* \bar{\Omega}) \wedge \cdots \wedge (\Psi_f^* \bar{\Omega})}_{\bar{n}}] = 0.$$

However, by Lemma 4.3.4, we have

$$\pounds_{\bar{X}/f}(f^{\bar{n}}\underbrace{\bar{\Omega}\wedge\cdots\wedge\bar{\Omega}}_{\bar{n}}) = 0,$$

and hence $\pounds_{\bar{X}/f}(f^{\bar{n}}\bar{\Lambda}) = 0$; this implies $\operatorname{div}_{f^{\bar{n}}\bar{\Lambda}}(\bar{X}/f) = 0$. However, the identity Eq. (4.17) gives

$$\operatorname{div}_{f^{\bar{n}-1}\bar{\Lambda}}(\bar{X}) = \operatorname{div}_{f^{\bar{n}}\bar{\Lambda}}(\bar{X}/f) = 0,$$

which implies $\pounds_{\bar{X}}(f^{\bar{n}-1}\bar{\Lambda}) = 0$. □

4.3.2 The Chaplygin Hamiltonization

Here we discuss the so-called *Chaplygin Hamiltonization* of the reduced system Eq. (4.13). Let us first find the equation that is satisfied by the vector field $\bar{Y}$ defined in Eq. (4.16).

Lemma 4.3.6. *The vector field $\bar{Y} \in \mathfrak{X}(T^*\bar{Q})$ satisfies the following equation:*

$$i_{\bar{Y}}\bar{\Omega} = d\bar{H}_{\mathrm{C}} - \Psi_{1/f}^* i_{\bar{X}}\big(d(\ln f)\wedge\bar{\Theta} - \Xi\big)\,, \tag{4.18}$$

where $\bar{H}_{\mathrm{C}} : T^*\bar{Q} \to \mathbb{R}$ is defined by

$$\bar{H}_{\mathrm{C}} := \bar{H}\circ\Psi_{1/f}. \tag{4.19}$$

Proof. As shown in Eq. (4.3.1), the vector fields $\bar{X}/f$ and $\bar{Y}$ are Ψ_f-related. Therefore $i_{\bar{X}/f}\Psi_f^*\alpha = \Psi_f^* i_{\bar{Y}}\alpha$ for any differential form α [see, e.g., 1, Proposition 2.4.14]; in particular, for $\alpha = \bar{\Omega}$, we have

$$i_{\bar{X}/f}\Psi_f^*\bar{\Omega} = \Psi_f^* i_{\bar{Y}}\bar{\Omega}.$$

However, using Lemma 4.3.4 and Eq. (4.13) on the left-hand side, we have

$$\begin{aligned}
i_{\bar{X}/f}\Psi_f^*\bar{\Omega} &= i_{\bar{X}/f}(f\bar{\Omega} - df\wedge\bar{\Theta}) \\
&= i_{\bar{X}}\bar{\Omega} - i_{\bar{X}}\left(\frac{1}{f}\,df\wedge\bar{\Theta}\right) \\
&= d\bar{H} + i_{\bar{X}}\Xi - i_{\bar{X}}\big(d(\ln f)\wedge\bar{\Theta}\big) \\
&= d\bar{H} - i_{\bar{X}}\big(d(\ln f)\wedge\bar{\Theta} - \Xi\big)\,.
\end{aligned}$$

Therefore

$$\Psi_f^* i_{\bar{Y}} \bar{\Omega} = d\bar{H} - i_{\bar{X}} \big(d(\ln f) \wedge \bar{\Theta} - \Xi \big) ,$$

and then applying $\Psi_{1/f}^*$ to both sides gives Eq. (4.18). $\square$

Proposition 4.3.7 (Necessary and Sufficient Condition for Symplectization)**.** *Then the vector field $\bar{Y} \in \mathfrak{X}(T^*\bar{Q})$ is symplectic if and only if the one-form*

$$i_{\bar{X}} \big(d(\ln f) \wedge \bar{\Theta} - \Xi \big)$$

is closed.

Proof. $\bar{Y}$ is symplectic if and only if $\pounds_{\bar{Y}} \bar{\Omega} = 0$. However, using Cartan's formula and Eq. (4.18),

$$\begin{aligned} \pounds_{\bar{Y}} \bar{\Omega} &= d i_{\bar{Y}} \bar{\Omega} + i_{\bar{Y}} d\bar{\Omega} \\ &= -\Psi_{1/f}^* d i_{\bar{X}} \big(d(\ln f) \wedge \bar{\Theta} - \Xi \big) . \end{aligned}$$

Thus $\bar{Y}$ is symplectic if and only if the last term in the above equation vanishes, which is equivalent to $d i_{\bar{X}} \big(d(\ln f) \wedge \bar{\Theta} - \Xi \big) = 0$ since $\Psi_{1/f}$ is a diffeomorphism. $\square$

Combining this result with Theorem (4.3.1), we have

Corollary 4.3.8. *Suppose there exists a fiber-wise constant function $F : T^*\bar{Q} \to \mathbb{R}$ such that $i_{\bar{X}} \big(dF \wedge \bar{\Theta} - \Xi \big)$ is closed. Then, by setting $f := \exp F$, the $2\bar{n}$-form $f^{\bar{n}-1} \bar{\Lambda}$ is an invariant measure of the reduced system Eq.* (4.13)*.*

We now state the main result of this section. The following theorem will be used in the next section in relation to nonholonomic Hamilton–Jacobi theory:

Theorem 4.3.9 (A Sufficient Condition for Chaplygin Hamiltonization)**.** *Suppose there exists a fiber-wise constant function $F : T^*\bar{Q} \to \mathbb{R}$ that satisfies the equation*

$$dF \wedge \bar{\Theta} = \Xi. \tag{4.20}$$

Then, by setting $f := \exp F$, the vector field $\bar{Y} \in \mathfrak{X}(T^*\bar{Q})$ satisfies the following Hamilton's equations:

$$i_{\bar{Y}} \bar{\Omega} = d\bar{H}_{\mathrm{C}}, \tag{4.21}$$

and, as a result, the reduced nonholonomic dynamics Eq. (4.13) *has the invariant measure $f^{\bar{n}-1} \bar{\Lambda}$.*

Proof. Straightforward from Lemma 4.3.6 and Corollary 4.3.2. $\square$

Remark 4.3.10. As shown by Stanchenko [54] (see also Cantrijn et al. [12]), Eq. (4.20) is also a sufficient condition for the two-form $\bar{\Omega}_f := f(\bar{\Omega} - \Xi)$ to be closed, so that Eq. (4.13) becomes

$$i_{\bar{X}/f}\bar{\Omega}_f = d\bar{H}$$

and so the dynamics of $\bar{X}/f$ is Hamiltonian with the non-standard symplectic form $\bar{\Omega}_f$.

4.4 Nonholonomic Hamilton–Jacobi Theory via Chaplygin Hamiltonization

4.4.1 The Chaplygin Hamilton–Jacobi Equation

Theorem 4.3.9 shows that the dynamics of $\bar{Y}$ on $T^*\bar{Q}$ is, under a certain condition, Hamiltonian with the standard symplectic form $\bar{\Omega}$ on $T^*\bar{Q}$ and the Chaplygin Hamiltonian $\bar{H}_{\mathrm{C}} : T^*\bar{Q} \to \mathbb{R}$. Therefore the conventional Hamilton–Jacobi theory applies directly to this case. Specifically, the (time-independent) Hamilton–Jacobi equation for this dynamics is written as follows:

$$\bar{H}_{\mathrm{C}} \circ d\bar{W} = E, \tag{4.22}$$

with an unknown function $\bar{W} : \bar{Q} \to \mathbb{R}$ and a constant E (the total energy). We would like to call Eq. (4.22) the *Chaplygin Hamilton–Jacobi equation*.

Now that we have a conventional Hamilton–Jacobi equation related to the Hamiltonized dynamics of a nonholonomic system, a natural question to ask is: What is the relationship between the Chaplygin Hamilton–Jacobi equation (4.22) and the nonholonomic Hamilton–Jacobi equation developed in the previous chapter? In this section, we would like to establish a link between the Chaplygin Hamilton–Jacobi equation and the nonholonomic Hamilton–Jacobi equation of the previous chapter.

4.4.2 Relationship between the Chaplygin H–J and Nonholonomic H–J Equations

First recall from the previous chapter that the nonholonomic Hamilton–Jacobi equation is an equation for a one-form γ on the original configuration manifold Q (not the reduced one $\bar{Q} := Q/G$):

$$H \circ \gamma = E, \tag{4.23}$$

along with the conditions that γ, seen as a map from Q to T^*Q, takes values in the constrained momentum space $\mathcal{M} \subset T^*Q$, i.e., $\gamma : Q \to \mathcal{M}$, and also that

$$d\gamma|_{\mathcal{D}\times\mathcal{D}} = 0, \ \text{ i.e., } d\gamma(v,w) = 0 \text{ for any } v, w \in \mathcal{D}. \tag{4.24}$$

In relating the Chaplygin Hamilton–Jacobi equation (4.22) with the nonholonomic Hamilton–Jacobi equation (4.23), a natural starting point is to look into the relation between the Chaplygin Hamiltonian $\bar{H}_{\mathrm{C}}$ and the original one H: Recall from Eqs. (4.12) and (4.19) that they are related through the Hamiltonian $\bar{H}$; the upper half of the following commutative diagram shows their relations.

$$\begin{tikzcd}
& \mathbb{R} & \\
\mathcal{M} \arrow[ur, "H"] & T^*\bar{Q} \arrow[l, "\mathrm{hl}^{\mathcal{M}}"'] \arrow[u, "\bar{H}"] & T^*\bar{Q} \arrow[l, "\Psi_{1/f}"'] \arrow[ul, "\bar{H}_{\mathrm{C}}"'] \\
Q \arrow[u, dashed, "\gamma"] \arrow[rr, "\pi"'] & & \bar{Q} \arrow[u, "d\bar{W}"']
\end{tikzcd} \tag{4.25}$$

Now suppose that a function $\bar{W} : \bar{Q} \to \mathbb{R}$ satisfies the Chaplygin Hamilton–Jacobi equation (4.22). This means that the one-form $d\bar{W}$, seen as a map from $\bar{Q}$ to $T^*\bar{Q}$, satisfies $\bar{H}_{\mathrm{C}} \circ d\bar{W}(\bar{q}) = E$ for any $\bar{q} \in \bar{Q}$ with some constant E; equivalently, $\bar{H}_{\mathrm{C}} \circ d\bar{W} \circ \pi(q) = E$ for any $q \in Q$. The lower-half of the above diagram (4.25) incorporates this view, and also leads us to the following:

Theorem 4.4.1. *Suppose that there exists a fiber-wise constant function $F : T^*\bar{Q} \to \mathbb{R}$ that satisfies Eq.* (4.20)*, and hence by Theorem 4.3.9, we have Hamilton's equations* (4.21) *for the vector field $\bar{Y}$ with $f := \exp F$. Let $\bar{W} : \bar{Q} \to \mathbb{R}$ be a solution of the*

Chaplygin Hamilton–Jacobi equation (4.22)*, and define* $\gamma : Q \to \mathcal{M}$ *by*

$$\begin{aligned}\gamma(q) &:= \mathrm{hl}^{\mathcal{M}}_q \circ \Psi_{1/f} \circ d\bar{W} \circ \pi(q) \\ &= \mathrm{hl}^{\mathcal{M}}_q \left(\frac{1}{f(\bar{q})} d\bar{W}(\bar{q}) \right),\end{aligned} \tag{4.26}$$

where $\bar{q} := \pi(q)$*. Then* γ *satisfies the nonholonomic Hamilton–Jacobi equation* (4.23) *as well as Eq.* (4.24)*.*

Proof. That the γ defined by Eq. (4.26) satisfies the nonholonomic Hamilton–Jacobi equation (4.23) follows from the diagram (4.25). To show that it also satisfies Eq. (4.24), we need to perform the following lengthy calculations: Let $X, Y \in \mathfrak{X}(Q)$ be arbitrary horizontal vector fields, i.e., $X_q, Y_q \in \mathcal{D}_q$ for any $q \in Q$. We start from the following identity:

$$d\gamma(X, Y) = X[\gamma(Y)] - Y[\gamma(X)] - \gamma([X, Y]). \tag{4.27}$$

Our goal is to show that the right-hand side vanishes. Let us first evaluate the first two terms on the right-hand side of the above identity at an arbitrary point $q \in Q$: Since $Y_q \in \mathcal{D}_q$, there exists $\bar{Y}_{\bar{q}} \in T_{\bar{q}}\bar{Q}$ such that $Y_q = \mathrm{hl}^{\mathcal{D}}_q(\bar{Y}_{\bar{q}})$. Thus[1], using Lemma 4.A.2,

$$\begin{aligned}\gamma(Y)(q) &= \left\langle \mathrm{hl}^{\mathcal{M}}_q \circ \Psi_{1/f} \circ d\bar{W}(\bar{q}), \mathrm{hl}^{\mathcal{D}}_q(\bar{Y}_{\bar{q}}) \right\rangle \\ &= \left\langle \Psi_{1/f} \circ d\bar{W}(\bar{q}), \bar{Y}_{\bar{q}} \right\rangle \\ &= \frac{1}{f(\bar{q})}\, d\bar{W}(\bar{Y})(\bar{q}).\end{aligned}$$

Hence, defining a function $\gamma_Y : \bar{Q} \to \mathbb{R}$ by

$$\gamma_Y(\bar{q}) := \frac{1}{f(\bar{q})}\, d\bar{W}(\bar{Y})(\bar{q}),$$

[1] Recall that $f : T^*\bar{Q} \to \mathbb{R}$ is fiber-wise constant and thus, with a slight abuse of notation, we may write $f(\alpha_{\bar{q}}) = f(\bar{q})$ for any $\alpha_{\bar{q}} \in T^*_{\bar{q}}\bar{Q}$; therefore f is seen as a function on $\bar{Q}$ as well.

we have $\gamma(Y) = \gamma_Y \circ \pi$. Therefore

$$\begin{aligned}
X[\gamma(Y)](q) &= X[\gamma_Y \circ \pi](q) \\
&= \langle d(\gamma_Y \circ \pi)_q, X_q \rangle \\
&= \langle (\pi^* d\gamma_Y)_q, X_q \rangle \\
&= \langle d\gamma_Y(\bar{q}), T_q\pi(X_q) \rangle \\
&= \big\langle d\gamma_Y(\bar{q}), \bar{X}_{\bar{q}}) \big\rangle \\
&= \bar{X}[\gamma_Y](\bar{q}) \\
&= \bar{X}\left[\frac{1}{f}\, d\bar{W}(\bar{Y})\right](\bar{q}) \\
&= \left(\frac{1}{f}\, \bar{X}\big[\bar{Y}\big[\bar{W}\big]\big] - \frac{1}{f^2}\, df(\bar{X})\, d\bar{W}(\bar{Y})\right)(\bar{q}),
\end{aligned}$$

where $\bar{X}_{\bar{q}} := T_q\pi(X_q)$, i.e., $X_q = \mathrm{hl}^{\mathcal{D}}_q(\bar{X}_{\bar{q}})$ because X is horizontal, and f may be seen as a function on $\bar{Q}$. Hence we have

$$\begin{aligned}
X[\gamma(Y)] - Y[\gamma(X)] &= \frac{1}{f}\big(\bar{X}\big[\bar{Y}\big[\bar{W}\big]\big] - \bar{Y}\big[\bar{X}\big[\bar{W}\big]\big]\big) \\
&\quad - \frac{1}{f^2}\big(df(\bar{X})\, d\bar{W}(\bar{Y}) - df(\bar{Y})\, d\bar{W}(\bar{X})\big) \\
&= \frac{1}{f}\, \big\langle d\bar{W}, [\bar{X}, \bar{Y}] \big\rangle - \frac{1}{f^2}\, df \wedge d\bar{W}(\bar{X}, \bar{Y}),
\end{aligned} \tag{4.28}$$

where we did not write down q and $\bar{q}$ for simplicity.

Now let us evaluate the last term on the right-hand side of Eq. (4.27): First we would like to decompose $[X, Y]_q$ into the horizontal and vertical part. Since both X and Y are horizontal, we have[2]

$$\mathrm{hor}([X, Y]_q) = \mathrm{hl}^{\mathcal{D}}_q([\bar{X}, \bar{Y}]_{\bar{q}}),$$

whereas the vertical part is

$$\mathrm{ver}([X, Y]_q) = (\mathcal{A}_q([X, Y]_q))_Q\,(q) = -(\mathcal{B}_q(X_q, Y_q))_Q\,(q)$$

where we used the following relation between the connection $\mathcal{A}$ and its curvature $\mathcal{B}$

[2] See, e.g., Kobayashi and Nomizu [35, Proposition 1.3 (3), p. 65].

that hold for horizontal vector fields X and Y:

$$\begin{aligned}\mathcal{B}_q(X_q, Y_q) &= d\mathcal{A}_q(X_q, Y_q) \\ &= X[\mathcal{A}(Y)](q) - X[\mathcal{A}(Y)](q) - \mathcal{A}([X, Y])(q) \\ &= -\mathcal{A}([X, Y])(q).\end{aligned}$$

As a result, we have the decomposition

$$[X, Y]_q = \mathrm{hl}_q^{\mathcal{D}}([\bar{X}, \bar{Y}]_{\bar{q}}) - (\mathcal{B}_q(X_q, Y_q))_Q\,(q),$$

and therefore

$$\begin{aligned}\gamma([X, Y])(q) &= \left\langle \mathrm{hl}_q^{\mathcal{M}} \circ \Psi_{1/f} \circ d\bar{W} \circ \pi(q), \mathrm{hl}_q^{\mathcal{D}}([\bar{X}, \bar{Y}]_{\bar{q}}) \right\rangle \\ &\quad - \left\langle \mathrm{hl}_q^{\mathcal{M}} \circ \Psi_{1/f} \circ d\bar{W} \circ \pi(q), (\mathcal{B}_q(X_q, Y_q))_Q\,(q) \right\rangle \\ &= \left\langle \Psi_{1/f} \circ d\bar{W}(\bar{q}), [\bar{X}, \bar{Y}]_{\bar{q}} \right\rangle - \left\langle \mathbf{J}\big(\mathrm{hl}_q^{\mathcal{M}} \circ \Psi_{1/f} \circ d\bar{W}(\bar{q})\big), \mathcal{B}_q(X_q, Y_q) \right\rangle \\ &= \frac{1}{f(\bar{q})} \left\langle d\bar{W}(\bar{q}), [\bar{X}, \bar{Y}]_{\bar{q}} \right\rangle - \left\langle \mathbf{J} \circ \mathrm{hl}_q^{\mathcal{M}}\big(d\bar{W}(\bar{q})/f(\bar{q})\big), \mathcal{B}_q(X_q, Y_q) \right\rangle \\ &= \frac{1}{f(\bar{q})} \left\langle d\bar{W}, [\bar{X}, \bar{Y}] \right\rangle(\bar{q}) \\ &\quad - \frac{1}{f(\bar{q})} \left\langle \mathbf{J} \circ \mathrm{hl}_q^{\mathcal{M}}\big(d\bar{W}(\bar{q})\big), \mathcal{B}_q\big(\mathrm{hl}_q^{\mathcal{D}}(\bar{X}_{\bar{q}}), \mathrm{hl}_q^{\mathcal{D}}(\bar{Y}_{\bar{q}})\big) \right\rangle \\ &= \frac{1}{f(\bar{q})} \left\langle d\bar{W}, [\bar{X}, \bar{Y}] \right\rangle(\bar{q}) - \frac{1}{f(\bar{q})} (d\bar{W})^* \Xi(\bar{X}, \bar{Y})(\bar{q}), \end{aligned} \tag{4.29}$$

where the second equality follows from Lemma 4.A.2 and the definition of the momentum map $\mathbf{J}$; the fourth one follows from the linearity of $\mathrm{hl}^{\mathcal{M}}$ and also of $\mathbf{J}$ in the fiber variables; the last one follows from the definition of Ξ in Eq. (4.14): Since $\pi_{\bar{Q}} \circ d\bar{W} = \mathrm{id}_{\bar{Q}}$ and thus $T\pi_{\bar{Q}} \circ Td\bar{W} = \mathrm{id}_{T\bar{Q}}$, we have

$$\begin{aligned}(d\bar{W})^* \Xi(\bar{X}, \bar{Y})(\bar{q}) &= \Xi_{d\bar{W}(\bar{q})}\big(Td\bar{W}(\bar{X}_{\bar{q}}), Td\bar{W}(\bar{Y}_{\bar{q}})\big) \\ &= \left\langle \mathbf{J} \circ \mathrm{hl}_q^{\mathcal{M}}\big(d\bar{W}(\bar{q})\big), \mathcal{B}_q\big(\mathrm{hl}_q^{\mathcal{D}}(\bar{X}_{\bar{q}}), \mathrm{hl}_q^{\mathcal{D}}(\bar{Y}_{\bar{q}})\big) \right\rangle.\end{aligned}$$

Substituting Eqs. (4.28) and (4.29) into Eq. (4.27), we obtain

$$\begin{aligned}
d\gamma(X,Y) &= -\frac{1}{f^2}\, df \wedge d\bar{W}(\bar{X},\bar{Y}) + \frac{1}{f}(d\bar{W})^*\Xi(\bar{X},\bar{Y}) \\
&= -\frac{1}{f}\big(d(\ln f) \wedge d\bar{W} - (d\bar{W})^*\Xi\big)\,(\bar{X},\bar{Y}) \\
&= -\frac{1}{f}(d\bar{W})^*\big(d(\ln f) \wedge \bar{\Theta} - \Xi\big)\,(\bar{X},\bar{Y}) \\
&= -\frac{1}{f}(d\bar{W})^*(dF \wedge \bar{\Theta} - \Xi)(\bar{X},\bar{Y}) \\
&= 0,
\end{aligned}$$

where the third line follows since[3] $(d\bar{W})^* f(\bar{q}) = f\big(d\bar{W}(\bar{q})\big) = f(\bar{q})$ and also that $(d\bar{W})^*\bar{\Theta} = d\bar{W}$ [see, e.g., 1, Proposition 3.2.11 on p. 179]; the last line follows from Eq. (4.20), which is assumed to be satisfied. □

4.5 Examples

Example 4.5.1. Consider the motion of the vertical rolling disk treated in Example 3.3.1.

The Lagrangian $L : TQ \to \mathbb{R}$ and the Hamiltonian $H : T^*Q \to \mathbb{R}$ are given by

$$L = \frac{1}{2}m\big(\dot{x}^2 + \dot{y}^2\big) + \frac{1}{2}J\dot{\varphi}^2 + \frac{1}{2}I\dot{\psi}^2$$

and

$$H = \frac{1}{2}\left(\frac{p_x^2 + p_y^2}{m} + \frac{p_\varphi^2}{J} + \frac{p_\psi^2}{I}\right), \tag{4.30}$$

respectively. The velocity constraints are

$$\dot{x} = R\cos\varphi\,\dot{\psi}, \qquad \dot{y} = R\sin\varphi\,\dot{\psi},$$

or in terms of constraint one-forms,

$$\omega^1 = dx - R\cos\varphi\, d\psi, \qquad \omega^2 = dy - R\sin\varphi\, d\psi.$$

So the constraint distribution $\mathcal{D} \subset TQ$ and the constrained momentum space

[3] Again recall that $f : T^*\bar{Q} \to \mathbb{R}$ may be seen as a function on $\bar{Q}$ as well.

$\mathcal{M} \subset T^*Q$ are given by

$$\mathcal{D} = \left\{(\dot{x}, \dot{y}, \dot{\varphi}, \dot{\psi}) \in TQ \mid \dot{x} = R\cos\varphi\,\dot{\psi},\ \dot{y} = R\sin\varphi\,\dot{\psi}\right\}$$

and

$$\mathcal{M} = \left\{(p_x, p_y, p_\varphi, p_\psi) \in T^*Q \mid p_x = \frac{mR}{I}\cos\varphi\,p_\psi,\ p_y = \frac{mR}{I}\sin\varphi\,p_\psi\right\},$$

respectively.

Let $G = \mathbb{R}^2$ and consider the action of G on Q defined by

$$G \times Q \to Q; \quad ((a,b),(x,y,\varphi,\psi)) \mapsto (x+a, y+b, \varphi, \psi).$$

Then Eq. (4.2) is satisfied, and hence it is a Chaplygin system. The Lie algebra $\mathfrak{g}$ is identified with $\mathbb{R}^2$ in this case. Let us use (ξ, η) as the coordinates for $\mathfrak{g}$. Then we may write the connection $\mathcal{A} : TQ \to \mathfrak{g}$ as

$$\mathcal{A} = (dx - R\cos\varphi\,d\psi) \otimes \frac{\partial}{\partial\xi} + (dy - R\sin\varphi\,d\psi) \otimes \frac{\partial}{\partial\eta}, \tag{4.31}$$

and hence its curvature as

$$\mathcal{B} = R\left(\sin\varphi\,d\varphi \wedge d\psi \otimes \frac{\partial}{\partial\xi} - \cos\varphi\,d\varphi \wedge d\psi \otimes \frac{\partial}{\partial\eta}\right). \tag{4.32}$$

Furthermore, the momentum map $\mathbf{J} : T^*Q \to \mathfrak{g}^*$ is given by

$$\mathbf{J}(p_q) = p_x d\xi + p_y d\eta. \tag{4.33}$$

The quotient space is $\bar{Q} := Q/G = \{(\varphi, \psi)\}$. The Hamiltonian $\bar{H} : T^*\bar{Q} \to \mathbb{R}$ is

$$\bar{H} = \frac{1}{2}\left(\frac{1}{J}\,p_\varphi^2 + \frac{I + mR^2}{I^2}\,p_\psi^2\right). \tag{4.34}$$

A simple calculation shows that the horizontal lift $\mathrm{hl}^{\mathcal{M}} : T^*\bar{Q} \to \mathcal{M}$ is given by

$$\mathrm{hl}^{\mathcal{M}}(p_\varphi, p_\psi) = \left(\frac{mR}{I}\cos\varphi\,p_\psi,\ \frac{mR}{I}\sin\varphi\,p_\psi,\ p_\varphi,\ p_\psi\right), \tag{4.35}$$

Then we find from Eq. (4.14) along with Eqs. (4.31), (4.32), (4.33), and (4.35) that $\Xi = 0$. Therefore the sufficient condition for Chaplygin Hamiltonization Eq. (4.20) reduces to $dF \wedge \Theta = 0$, and hence we may choose $F = 0$ and thus $f = 1$; then the Chaplygin Hamiltonian $\bar{H}_{\mathrm{C}} : T^*\bar{Q} \to \mathbb{R}$ is identical to $\bar{H}$. So the Chaplygin Hamilton–Jacobi equation (4.22) becomes

$$\frac{1}{2}\left[\frac{1}{J}\left(\frac{\partial \bar{W}}{\partial \varphi}\right)^2 + \frac{I + mR^2}{I^2}\left(\frac{\partial \bar{W}}{\partial \psi}\right)^2\right] = E, \tag{4.36}$$

Now we employ the conventional approach of separation of variables, i.e., assume that $\bar{W} : \bar{Q} \to \mathbb{R}$ takes the following form:

$$W(\varphi, \psi) = W_\varphi(\varphi) + W_\psi(\psi).$$

Then Eq. (4.36) becomes

$$\frac{1}{2}\left[\frac{1}{J}\left(\frac{d\bar{W}_\varphi}{d\varphi}\right)^2 + \frac{I + mR^2}{I^2}\left(\frac{d\bar{W}_\psi}{d\psi}\right)^2\right] = E.$$

Since the first term on the left-hand side depends only on φ and the second only on ψ, we obtain the solution

$$\frac{d\bar{W}_\varphi}{d\varphi} = \gamma_\varphi^0, \qquad \frac{d\bar{W}_\psi}{d\psi} = \gamma_\psi^0, \tag{4.37}$$

where γ_φ^0 and γ_ψ^0 are the constants determined by the initial condition such that

$$\frac{1}{2}\left[\frac{1}{J}(\gamma_\varphi^0)^2 + \frac{I + mR^2}{I^2}(\gamma_\psi^0)^2\right] = E.$$

Then Eq. (4.26) gives the solution obtained in Example 3.3.1:

$$\gamma(x, y, \varphi, \psi) = \frac{mR}{I}\cos\varphi\,\gamma_\psi^0\,dx + \frac{mR}{I}\sin\varphi\,\gamma_\psi^0\,dy + \gamma_\varphi^0\,d\varphi + \gamma_\psi^0\,d\psi. \tag{4.38}$$

Example 4.5.2. Consider the knife edge problem treated in Example 3.3.2.

The Lagrangian $L : TQ \to \mathbb{R}$ and the Hamiltonian $H : T^*Q \to \mathbb{R}$ are given by

$$L = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) + \frac{1}{2}J\dot{\varphi}^2 + mgx\sin\alpha$$

and

$$H = \frac{1}{2}\left(\frac{p_x^2 + p_y^2}{m} + \frac{p_\varphi^2}{J}\right) - mgx\sin\alpha, \tag{4.39}$$

respectively. The velocity constraint is

$$\sin\varphi\,\dot{x} - \cos\varphi\,\dot{y} = 0$$

and so the constraint one-form is

$$\omega^1 = \sin\varphi\,dx - \cos\varphi\,dy.$$

The constraint distribution $\mathcal{D} \subset TQ$ and the constrained momentum space $\mathcal{M} \subset T^*Q$ are given by

$$\mathcal{D} = \{(\dot{x}, \dot{y}, \dot{\varphi}) \in TQ \mid \sin\varphi\,\dot{x} - \cos\varphi\,\dot{y} = 0\}$$

and

$$\mathcal{M} = \{(p_x, p_y, p_\varphi) \in T^*Q \mid \sin\varphi\,p_x = \cos\varphi\,p_y\},$$

respectively.

Let $G = \mathbb{R}$ and consider the action of G on Q defined by

$$G \times Q \to Q; \quad (a, (x, y, \varphi)) \mapsto (x, y + a, \varphi).$$

Then Eq. (4.2) is satisfied, and hence it is a Chaplygin system. The Lie algebra $\mathfrak{g}$ is identified with $\mathbb{R}$ in this case. Let us use η as the coordinate for $\mathfrak{g}$. Then we may write the connection $\mathcal{A} : TQ \to \mathfrak{g}$ as

$$\mathcal{A} = (dy - \tan\varphi\,dx) \otimes \frac{\partial}{\partial\eta}, \tag{4.40}$$

and hence its curvature as

$$\mathcal{B} = \frac{1}{\cos^2\varphi}\,dx \wedge d\varphi \otimes \frac{\partial}{\partial\eta}. \tag{4.41}$$

Furthermore, the momentum map $\mathbf{J} : T^*Q \to \mathfrak{g}^*$ is given by

$$\mathbf{J}(p_q) = p_y d\eta. \tag{4.42}$$

The quotient space is $\bar{Q} := Q/G = \{(x, \varphi)\}$. The Hamiltonian $\bar{H} : T^*\bar{Q} \to \mathbb{R}$ is

$$\bar{H} = \frac{1}{2}\left(\frac{\cos^2\varphi}{m}\,p_x^2 + \frac{1}{J}\,p_\varphi^2\right) - mgx\sin\alpha.$$

A simple calculation shows that the horizontal lift $\mathrm{hl}^{\mathcal{M}} : T^*\bar{Q} \to \mathcal{M}$ is given by

$$\mathrm{hl}^{\mathcal{M}}(p_x, p_\varphi) = \left(\cos^2\varphi\, p_x,\ \sin\varphi\cos\varphi\, p_x,\ p_\varphi\right), \tag{4.43}$$

Then we find from Eq. (4.14) along with Eqs. (4.40), (4.41), (4.42), and (4.43) that

$$\Xi = p_x \tan\varphi\, dx \wedge d\varphi.$$

Therefore the sufficient condition for Chaplygin Hamiltonization Eq. (4.20) becomes

$$p_\varphi \frac{\partial F}{\partial x} - p_x \frac{\partial F}{\partial \varphi} = p_x \tan\varphi.$$

It is easy to find the solution $F = \ln(\cos\varphi)$ and hence

$$f = \cos\varphi, \tag{4.44}$$

where we restrict φ to be in the range $(0, \pi/2)$ so that $f > 0$.

Then Eq. (4.19) gives the following Chaplygin Hamiltonian:

$$\begin{aligned}\bar{H}_{\mathrm{C}}(x, \varphi, p_x, p_\varphi) &= \bar{H}\left(x, \varphi, \frac{p_x}{\cos\varphi}, \frac{p_\varphi}{\cos\varphi}\right) \\ &= \frac{1}{2}\left(\frac{1}{m}\, p_x^2 + \frac{1}{J\cos^2\varphi}\, p_\varphi^2\right) - mgx\sin\alpha.\end{aligned}$$

So the Chaplygin Hamilton–Jacobi equation (4.22) becomes

$$\frac{1}{2}\left[\frac{1}{m}\left(\frac{\partial\bar{W}}{\partial x}\right)^2 + \frac{1}{J\cos^2\varphi}\left(\frac{\partial\bar{W}}{\partial\varphi}\right)^2\right] - mgx\sin\alpha = E, \tag{4.45}$$

Now we employ the conventional approach of separation of variables, i.e., assume that $\bar{W} : \bar{Q} \to \mathbb{R}$ takes the following form:

$$W(x, \varphi) = W_x(x) + W_\varphi(\varphi).$$

Then Eq. (4.45) becomes

$$\frac{1}{2}\left[\frac{1}{m}\left(\frac{d\bar{W}_x}{dx}\right)^2 - (2mg\sin\alpha)\, x + \frac{1}{J\cos^2\varphi}\left(\frac{d\bar{W}_\varphi}{d\varphi}\right)^2\right] = E.$$

The first two terms in the brackets depend only on x, whereas the third depends only on φ, and thus

$$\frac{1}{m}\left(\frac{d\bar{W}_x}{dx}\right)^2 - (2mg\sin\alpha)\,x = 2E - C^2, \qquad \frac{1}{J\cos^2\varphi}\left(\frac{d\bar{W}_\varphi}{d\varphi}\right)^2 = C^2,$$

with some constant C. Hence, assuming $d\bar{W}_x/dx \geq 0$, we have

$$\frac{d\bar{W}_x}{dx} = \sqrt{m(2E - C^2) + (2m^2 g\sin\alpha)\,x}, \qquad \frac{d\bar{W}_\varphi}{d\varphi} = C\sqrt{J}\cos\varphi.$$

Then Eq. (4.26) gives

$$\gamma(x, y, \varphi) = \sqrt{m(2E - C^2) + (2m^2 g\sin\alpha)\,x}\,(\cos\varphi\,dx + \sin\varphi\,dy) + C\sqrt{J}\,d\varphi,$$

This is the solution obtained in Example 3.3.2 with $C = \gamma_\varphi^0/\sqrt{J}$.

4.A Some Lemmas on the Horizontal Lift $\mathrm{hl}^{\mathcal{M}}$

Lemma 4.A.1. *The horizontal lift $\mathrm{hl}^{\mathcal{M}}$ is invariant under the action of the cotangent lift of the group action $\Phi : G \times Q \to Q$. Specifically, for any $h \in G$, we have*

$$\mathrm{hl}^{\mathcal{M}}_{hq} = T_q^*\Phi_{h^{-1}} \circ \mathrm{hl}^{\mathcal{M}}_q, \tag{4.46}$$

where $hq = \Phi_h(q)$; or equivalently, for any $\bar{\alpha}_{\bar{q}} \in T^*_{\bar{q}}\bar{Q}$,

$$\alpha^{\mathrm{h}}_{hq} = T_q^*\Phi_{h^{-1}}(\alpha^{\mathrm{h}}_q);$$

or the commutative diagram below commutes.

$$\begin{tikzcd}
\mathcal{M}_q \arrow[rr, "T_q^*\Phi_{h^{-1}}"] & & \mathcal{M}_{hq} \\
& T^*_{\bar{q}}\bar{Q} \arrow[ul, "\mathrm{hl}^{\mathcal{M}}_q"] \arrow[ur, "\mathrm{hl}^{\mathcal{M}}_{hq}"'] &
\end{tikzcd}
\qquad
\begin{tikzcd}
\alpha_q \arrow[rr, mapsto] & & \alpha_{hq} \\
& \bar{\alpha}_{\bar{q}} \arrow[ul, mapsto] \arrow[ur, mapsto] &
\end{tikzcd}$$

Proof. From the definition of $\mathrm{hl}^{\mathcal{M}}_q$,

$$\begin{aligned}
\mathrm{hl}^{\mathcal{M}}_{hq} &= \mathbb{F}L_{hq} \circ \mathrm{hl}^{\mathcal{D}}_{hq} \circ (\mathbb{F}\bar{L})^{-1}_{\bar{q}} \\
&= \mathbb{F}L_{hq} \circ T_q\Phi_h \circ \mathrm{hl}^{\mathcal{D}}_q \circ (\mathbb{F}\bar{L})^{-1}_{\bar{q}},
\end{aligned}$$

where we used the G-invariance of $\mathrm{hl}^{\mathcal{D}}$. Now, for any $v_q \in T_qQ$ and $w_{hq} \in T_{hq}Q$, using the G-invariance of the Lagrangian L,

$$\begin{aligned}
\langle \mathbb{F}L_{hq} \circ T_q\Phi_h(v_q), w_{hq}\rangle &= \frac{d}{d\varepsilon} L(T_q\Phi_h(v_q) + \varepsilon\, w_{hq})\Big|_{\varepsilon=0} \\
&= \frac{d}{d\varepsilon} L \circ T_q\Phi_h(v_q + \varepsilon\, T_{hq}\Phi_{h^{-1}}(w_{hq}))\Big|_{\varepsilon=0} \\
&= \frac{d}{d\varepsilon} L(v_q + \varepsilon\, T_{hq}\Phi_{h^{-1}}(w_{hq}))\Big|_{\varepsilon=0} \\
&= \langle \mathbb{F}L_q(v_q), T_{hq}\Phi_{h^{-1}}(w_{hq})\rangle \\
&= \left\langle T_q^*\Phi_{h^{-1}}(\mathbb{F}L_q(v_q)), w_{hq}\right\rangle ,
\end{aligned}$$

and thus we have $\mathbb{F}L_{hq} \circ T\Phi_h = T^*\Phi_{h^{-1}} \circ \mathbb{F}L_q$. Hence we obtain

$$\begin{aligned}
\mathrm{hl}^{\mathcal{M}}_{hq} &= T_q^*\Phi_{h^{-1}} \circ \mathbb{F}L_q \circ \mathrm{hl}^{\mathcal{D}}_q \circ (\mathbb{F}\bar{L})^{-1}_{\bar{q}} \\
&= T_q^*\Phi_{h^{-1}} \circ \mathrm{hl}^{\mathcal{M}}_q .
\end{aligned}$$

□

Lemma 4.A.2. *Let q be an arbitrary point in Q and $\bar{q} = \pi(q) \in \bar{Q}$. For any $\alpha_{\bar{q}} \in T^*_{\bar{q}}\bar{Q}$ and $v_{\bar{q}} \in T_{\bar{q}}\bar{Q}$, the following identity holds:*

$$\left\langle \mathrm{hl}^{\mathcal{M}}_q(\alpha_{\bar{q}}), \mathrm{hl}^{\mathcal{D}}_q(v_{\bar{q}})\right\rangle = \langle \alpha_{\bar{q}}, v_{\bar{q}}\rangle . \tag{4.47}$$

Proof. Follows from the definitions of $\bar{g}$ and $\mathrm{hl}^{\mathcal{M}}_q$ (see Eqs. (4.8) and (4.9), respectively):

$$\begin{aligned}
\left\langle \mathrm{hl}^{\mathcal{M}}_q(\alpha_{\bar{q}}), \mathrm{hl}^{\mathcal{D}}_q(v_{\bar{q}})\right\rangle &= \left\langle g^\flat_q \circ \mathrm{hl}^{\mathcal{D}}_q \circ (\bar{g}^\flat)^{-1}_{\bar{q}}(\alpha_{\bar{q}}), \mathrm{hl}^{\mathcal{D}}_q(v_{\bar{q}})\right\rangle \\
&= g_q\big(\mathrm{hl}^{\mathcal{D}}_q \circ (\bar{g}^\flat)^{-1}_{\bar{q}}(\alpha_{\bar{q}}), \mathrm{hl}^{\mathcal{D}}_q(v_{\bar{q}})\big) \\
&= \bar{g}_{\bar{q}}\big((\bar{g}^\flat)^{-1}_{\bar{q}}(\alpha_{\bar{q}}), v_{\bar{q}}\big) \\
&= \left\langle \bar{g}^\flat_{\bar{q}} \circ (\bar{g}^\flat)^{-1}_{\bar{q}}(\alpha_{\bar{q}}), v_{\bar{q}}\right\rangle \\
&= \langle \alpha_{\bar{q}}, v_{\bar{q}}\rangle .
\end{aligned}$$

□

Chapter 5

Discrete Hamilton–Jacobi Theory

5.1 Introduction

The main objective of this chapter is to present a discrete analogue of Hamilton–Jacobi theory within the framework of discrete Hamiltonian mechanics [41].

There are some previous works on discrete-time analogues of the Hamilton–Jacobi equation, such as Elnatanov and Schiff [22] and Lall and West [41]. Specifically, Elnatanov and Schiff [22] 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., 42, Chapter VIII]. Lall and West [41] formulated a discrete Lagrangian analogue of the Hamilton–Jacobi equation as a separable optimization problem.

5.1.1 Main Results

Our work was inspired by the result of Elnatanov and Schiff [22] and starts from a reinterpretation of their result in the language of discrete mechanics. This chapter further extends the result by developing discrete analogues of results in (continuous-time) Hamilton–Jacobi theory. Namely, we formulate a discrete analogue of Jacobi's solution, which relates the discrete action sum with 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 chapter 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 (discrete Hamilton–Jacobi–

Bellman equation). (See Fig. 5.1.) Specifically, we show that the discrete Hamilton–Jacobi equation applied to linear discrete Hamiltonian systems reduces to the discrete Riccati equation. This is again a discrete analogue of the well-known result that the Hamilton–Jacobi equation applied to linear Hamiltonian systems reduces to the Riccati equation [see, e.g., 34, p. 421]. Also, we establish a link with discrete-time optimal

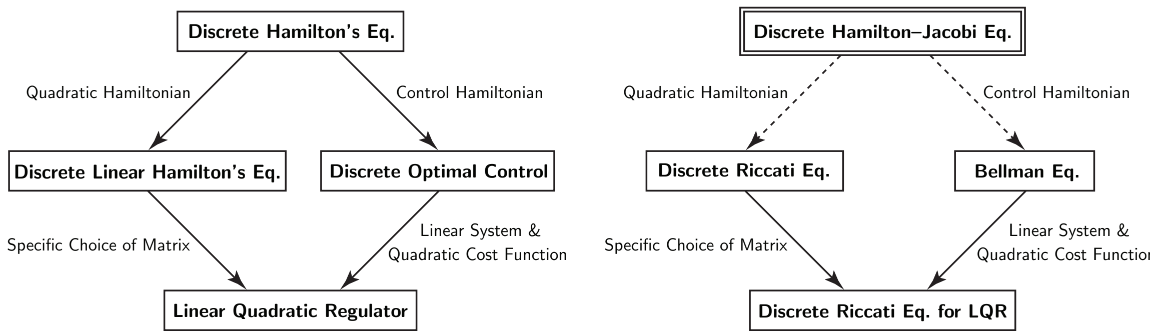

Figure 5.1: Discrete evolution equations (left) and corresponding discrete Hamilton–Jacobi-type equations (right). Dashed lines are the links established in this work.

control theory, and show that the Bellman equation of dynamic programming follows. This link makes it possible to interpret discrete analogues of Jacobi's solution and the Hamilton–Jacobi theorem in the optimal control setting. Namely we show that these results reduce to two well-known results in optimal control theory that relate the Bellman equation with the optimal solution.

5.1.2 Outline

A brief review of discrete Lagrangian and Hamiltonian mechanics is in Section 2.5. In Section 5.2 we describe a reinterpretation of the result of Elnatanov and Schiff [22] in the language of discrete mechanics and a discrete analogue of Jacobi's solution to the discrete Hamilton–Jacobi equation. The remainder of Section 5.2 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 5.3 we prove a discrete version of the Hamilton–Jacobi theorem. Section 5.4 establishes the link with discrete-time optimal control and interprets the results of the preceding sections in this setting. In Section 5.5 we apply the theory to linear discrete Hamiltonian systems, and show that the discrete Riccati equation follows from the discrete Hamilton–Jacobi equation. We then take a harmonic oscillator as a simple physical example, and solve the discrete Hamilton–Jacobi equation explicitly. Finally, Section 5.6 discusses the continuous-time limit of the theory.

5.2 Discrete Hamilton–Jacobi Equation

5.2.1 Derivation by Elnatanov and Schiff

Elnatanov and Schiff [22] 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 in Section 2.5.

Theorem 5.2.1. *Suppose that the discrete dynamics* $\{(q_k, p_k)\}_{k=0}^{N}$ *is governed by the right discrete Hamilton's equations* (2.48). *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-1 generating function[1] $S^k : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$*:*

$$\begin{aligned} \hat{p}_k &= -D_1 S^k(\hat{q}_k, q_k), \\ p_k &= D_2 S^k(\hat{q}_k, q_k); \end{aligned} \tag{5.1}$$

(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*:*

$$S^{k+1}(\hat{q}_0, q_{k+1}) - S^k(\hat{q}_0, q_k) - D_2 S^{k+1}(\hat{q}_0, q_{k+1}) \cdot q_{k+1} + H_{\mathrm{d}}^{+}\big(q_k, D_2 S^{k+1}(\hat{q}_0, q_{k+1})\big) = 0, \tag{5.2}$$

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

$$S_{\mathrm{d}}^{k+1}(q_{k+1}) - S_{\mathrm{d}}^k(q_k) - DS_{\mathrm{d}}^{k+1}(q_{k+1}) \cdot q_{k+1} + H_{\mathrm{d}}^{+}\big(q_k, DS_{\mathrm{d}}^{k+1}(q_{k+1})\big) = 0. \tag{5.3}$$

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

$$\begin{aligned} \hat{q}_{k+1} &= D_2 \hat{H}_{\mathrm{d}}^{+}(\hat{q}_k, \hat{p}_{k+1}), \\ \hat{p}_k &= D_1 \hat{H}_{\mathrm{d}}^{+}(\hat{q}_k, \hat{p}_{k+1}), \end{aligned} \tag{5.4}$$

[1] This is essentially the same as Eq. (2.41) 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: Eq. (2.41) describes the dynamics or time evolution whereas Eq. (5.1) is a change of coordinates.

or equivalently,

$$\hat{p}_k \, d\hat{q}_k + \hat{q}_{k+1} \, d\hat{p}_{k+1} = d\hat{H}_{\mathrm{d}}^{+}(\hat{q}_k, \hat{p}_{k+1}). \tag{5.5}$$

Let us first write $\hat{H}_{\mathrm{d}}^{+}$ in terms of the original right discrete Hamiltonian H_{d}^{+} and the generating function S^k. For that purpose, first rewrite Eqs. (2.48) and (5.1) 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

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

Thus in view of Eq. (5.5), we obtain

$$\hat{H}_{\mathrm{d}}^{+}(\hat{q}_k, \hat{p}_{k+1}) = H_{\mathrm{d}}^{+}(q_k, p_{k+1}) + \hat{p}_{k+1} \cdot \hat{q}_{k+1} - p_{k+1} \cdot q_{k+1} + S^{k+1}(\hat{q}_{k+1}, q_{k+1}) - S^k(\hat{q}_k, q_k). \tag{5.6}$$

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 Eq. (5.4) that we can set

$$\hat{H}_{\mathrm{d}}^{+}(\hat{q}_k, \hat{p}_{k+1}) = \hat{p}_{k+1} \cdot \hat{q}_k. \tag{5.7}$$

Then Eq. (5.6) 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 Eq. (5.1), we obtain Eq. (5.2). $\square$

Remark 5.2.2. What Elnatanov and Schiff [22] refer to the *Hamilton–Jacobi difference equation* is the following:

$$S^{k+1}(\hat{q}_0, q_{k+1}) - S^k(\hat{q}_0, q_k) - D_2 S^{k+1}(\hat{q}_0, q_{k+1}) \cdot D_2 H_{\mathrm{d}}^+(q_k, p_{k+1}) + H_{\mathrm{d}}^+(q_k, p_{k+1}) = 0. \tag{5.8}$$

It is clear that this is equivalent to Eq. (5.2) in view of Eq. (2.48)

5.2.2 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 5.2.3. *Consider the action sums Eq.* (2.31) *written in terms of the right discrete Hamiltonian, Eq.* (2.47)*:*

$$S_{\mathrm{d}}^k(q_k) := \sum_{l=0}^{k-1} \left[p_{l+1} \cdot q_{l+1} - H_{\mathrm{d}}^+(q_l, p_{l+1})\right] \tag{5.9}$$

evaluated along a solution of the right discrete Hamilton's equations (2.48)*; each $S_{\mathrm{d}}^k(q_k)$ is seen as a function of the end point coordinates q_k and the discrete end time k. Then these action sums satisfy the discrete Hamilton–Jacobi equation* (5.3)*.*

Proof. From Eq. (5.9), we have

$$S_{\mathrm{d}}^{k+1}(q_{k+1}) - S_{\mathrm{d}}^k(q_k) = p_{k+1} \cdot q_{k+1} - H_{\mathrm{d}}^+(q_k, p_{k+1}), \tag{5.10}$$

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

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

However, the term in the brackets vanishes because the right discrete Hamilton's equations (2.48) are assumed to be satisfied. Thus we have

$$p_{k+1} = DS_{\mathrm{d}}^{k+1}(q_{k+1}). \tag{5.11}$$

Substituting this into Eq. (5.10) gives Eq. (5.3). $\square$

Remark 5.2.4. Recall that, in the derivation of the continuous Hamilton–Jacobi equation [see, e.g., 26, Section 23], we consider the variation of the action integral Eq. (1.2) with respect to the end point (q, t) and find

$$dS = p\,dq - H(q, p)\,dt. \tag{5.12}$$

This gives

$$\frac{\partial S}{\partial t} = -H(q, p), \qquad p = \frac{\partial S}{\partial q}, \tag{5.13}$$

and hence the Hamilton–Jacobi equation

$$\frac{\partial S}{\partial t} + H\left(q, \frac{\partial S}{\partial q}\right) = 0. \tag{5.14}$$

In the above derivation of the discrete Hamilton–Jacobi equation (5.3), the difference in two action sums Eq. (5.10) is a natural discrete counterpart to the variation dS in Eq. (5.12). Notice also that Eq. (5.10) plays the same essential role as Eq. (5.12) does in deriving the Hamilton–Jacobi equation. Table 5.1 summarizes the correspondence between the ingredients in the continuous and discrete theories (see also Remark 5.2.4).

Table 5.1: 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 non-negative integers.

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

5.2.3 The Right and Left Discrete Hamilton–Jacobi Equations

Recall that, in Eq. (5.9), we wrote the action sum Eq. (2.31) in terms of the right discrete Hamiltonian Eq. (2.47). We can also write it in terms of the left discrete Hamiltonian Eq. (2.50) as follows:

$$S_{\mathrm{d}}^{k}(q_k) = \sum_{l=0}^{k-1} \left[-p_l \cdot q_l - H_{\mathrm{d}}^{-}(p_l, q_{l+1}) \right]. \tag{5.15}$$

Then we can proceed as in the proof of Theorem 5.2.3: First we have

$$S_{\mathrm{d}}^{k+1}(q_{k+1}) - S_{\mathrm{d}}^{k}(q_k) = -p_k \cdot q_k - H_{\mathrm{d}}^{-}(p_k, q_{k+1}). \tag{5.16}$$

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

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

However, the term in the brackets vanish because the left discrete Hamilton's equations (2.51) are assumed to be satisfied. Thus we have

$$p_k = DS_{\mathrm{d}}^{k}(q_k). \tag{5.17}$$

Substituting this into Eq. (5.16) gives the discrete Hamilton–Jacobi equation with the left discrete Hamiltonian:

$$S_{\mathrm{d}}^{k+1}(q_{k+1}) - S_{\mathrm{d}}^{k}(q_k) + DS_{\mathrm{d}}^{k}(q_k) \cdot q_k + H_{\mathrm{d}}^{-}\left(DS_{\mathrm{d}}^{k}(q_k), q_{k+1}\right) = 0. \tag{5.18}$$

We refer to Eqs. (5.3) and (5.18) as the *right and left discrete Hamilton–Jacobi equations*, respectively.

As mentioned above, Eqs. (5.9) and (5.15) are the same action sum Eq.(2.31) expressed in different ways. Therefore we may summarize the above argument as follows:

Proposition 5.2.5. *The action sums, Eq.* (5.9) *or equivalently Eq.* (5.15)*, satisfy both the right and left discrete Hamilton–Jacobi equations* (5.3) *and* (5.18)*.*

5.2.4 Explicit Forms of the Discrete Hamilton–Jacobi Equations

The expressions for the right and left discrete Hamilton–Jacobi equations in Eqs. (5.3) and (5.18) are implicit in the sense that they contain two spatial variables q_k and q_{k+1}. However Theorem 5.2.3 suggests that q_k and q_{k+1} may be considered to be related by the dynamics defined by either Eq. (2.48) or (2.51), or equivalently, the discrete Hamiltonian map $\tilde{F}_{L_\mathrm{d}} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$ defined in Eq. (2.42). 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 5.4 that the explicit form is compatible with the formulation of the well-known Bellman equation.

For the right discrete Hamilton–Jacobi equation (5.3), we first define the map $f_k^+ : Q \to Q$ as follows: Replace p_{k+1} in Eq. (2.48a) by $DS_\mathrm{d}^{k+1}(q_{k+1})$ as suggested by Eq. (5.11):

$$q_{k+1} = D_2 H_\mathrm{d}^+\big(q_k, DS_\mathrm{d}^{k+1}(q_{k+1})\big)\,. \tag{5.19}$$

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

$$f_k^+(q_k) = D_2 H_\mathrm{d}^+\big(q_k, DS_\mathrm{d}^{k+1}(f_k^+(q_k))\big)\,. \tag{5.20}$$

We may now identify q_{k+1} with $f_k^+(q_k)$ in the implicit form of the right Hamilton–Jacobi equation (5.3):

$$S_\mathrm{d}^{k+1}(f_k^+(q)) - S_\mathrm{d}^k(q) - DS_\mathrm{d}^{k+1}(f_k^+(q)) \cdot f_k^+(q) + H_\mathrm{d}^+\big(q, DS_\mathrm{d}^{k+1}(f_k^+(q))\big) = 0, \tag{5.21}$$

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* Eq. (5.21) to be the *right discrete Hamilton–Jacobi equation.* Notice that these are differential-difference equations defined on $Q \times \mathbb{N}$, with the spatial variable q and the discrete time k.

For the left discrete Hamilton–Jacobi equation (5.18), we define the map $f_k^- : Q \to Q$ as follows:

$$f_k^-(q_k) := \pi_Q \circ \tilde{F}_{L_\mathrm{d}}\big(dS_\mathrm{d}^k(q_k)\big)\,, \tag{5.22}$$

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

$$
\begin{tikzcd}
T^*Q \arrow[r, "\tilde{F}_{L_\mathrm{d}}"] & T^*Q \arrow[d, "\pi_Q"] \\
Q \arrow[u, "dS_\mathrm{d}^k"] \arrow[r, dashed, "f_k^-"'] & Q
\end{tikzcd}
\qquad
\begin{tikzcd}
dS_\mathrm{d}^k(q_k) \arrow[r, mapsto] & \tilde{F}_{L_\mathrm{d}}\big(dS_\mathrm{d}^k(q_k)\big) \arrow[d, mapsto] \\
q_k \arrow[u, mapsto] \arrow[r, dashed, mapsto] & f_k^-(q_k)
\end{tikzcd}
\tag{5.23}
$$

Notice also that, since the map $\tilde{F}_{L_\mathrm{d}} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$ is defined by Eq. (2.51), f_k^- is defined implicitly by

$$
q_k = -D_1 H_\mathrm{d}^-\big(DS_\mathrm{d}^k(q_k), f_k^-(q_k)\big)\,. \tag{5.24}
$$

In other words, replace p_k in Eq. (2.51a) by $DS_\mathrm{d}^k(q_k)$ as suggested by Eq. (5.17), 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 Eq. (5.18):

$$
S_\mathrm{d}^{k+1}(f_k^-(q)) - S_\mathrm{d}^k(q) + DS_\mathrm{d}^k(q)\cdot q + H_\mathrm{d}^-\big(DS_\mathrm{d}^k(q), f_k^-(q)\big) = 0, \tag{5.25}
$$

where we again suppressed the subscript k of q_k. We *define* Eqs. (5.21) and (5.25) to be the *right and left discrete Hamilton–Jacobi equations*, respectively. Notice that these are differential-difference equations defined on $Q \times \mathbb{N}$, with the spatial variable q and the discrete time k.

Remark 5.2.6. That the discrete Hamilton–Jacobi equation is a differential-difference equation defined on $Q\times\mathbb{N}$ corresponds to the fact that the continuous-time Hamilton–Jacobi equation (5.14) is a partial differential equation defined on $Q \times \mathbb{R}_+$.

Remark 5.2.7. Notice that the right discrete Hamilton–Jacobi equation (5.21) is more complicated than the left one (5.25), particularly because the map f_k^+ appears more often than f_k^- does in the latter; notice here that, as shown in Eq. (5.22), the maps $f_k^\pm$ in the discrete Hamilton–Jacobi equations (5.21) and (5.25) 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

$$
g_k(q_k) := \pi_Q \circ \tilde{F}_{L_\mathrm{d}}^{-1}\big(dS_\mathrm{d}^k(q_k)\big)\,, \tag{5.26}
$$

or so that the diagram below commutes.

$$
\begin{tikzcd}
T^*Q \arrow[d, "\pi_Q"'] & T^*Q \arrow[l, "\tilde{F}_{L_\mathrm{d}}^{-1}"'] \\
Q & Q \arrow[l, dashed, "g_k"] \arrow[u, "dS_\mathrm{d}^k"']
\end{tikzcd}
\qquad
\begin{tikzcd}
\tilde{F}_{L_\mathrm{d}}^{-1}\big(dS_\mathrm{d}^k(q_k)\big) \arrow[d, mapsto] & dS_\mathrm{d}^k(q_k) \arrow[l, mapsto] \\
g_k(q_k) & q_k \arrow[l, dashed, mapsto] \arrow[u, mapsto]
\end{tikzcd}
\tag{5.27}
$$

Now, in Eq. (5.3), change the indices from $(k, k+1)$ to $(k-1, k)$ and identify q_{k-1} with $g_k(q_k)$ to obtain

$$S_\mathrm{d}^k(q) - S_\mathrm{d}^{k-1}(g_k(q)) - DS_\mathrm{d}^k(q) \cdot q + H_\mathrm{d}^+\big(g_k(q), DS_\mathrm{d}^k(q)\big) = 0, \tag{5.28}$$

where we again suppressed the subscript k of q_k. This is as simple as the left discrete Hamilton–Jacobi equation (5.25). However the map g_k is, being backward in time, rather unnatural compared to f_k. Furthermore, as we shall see in Section 5.4, in the discrete optimal control setting, the map f_k is defined by a given function and thus the formulation with f_k will turn out to be more convenient.

5.2.5 The discrete Hamilton–Jacobi Equation on the Lagrangian Side

First notice that Eq. (2.31) gives

$$S_\mathrm{d}^{k+1}(q_{k+1}) - S_\mathrm{d}^k(q_k) = L_\mathrm{d}(q_k, q_{k+1}). \tag{5.29}$$

This is essentially the Lagrangian equivalent of the discrete Hamilton–Jacobi equation (5.21) as Lall and West [41] suggest. Let us apply the same argument as above to obtain the explicit form for Eq. (5.29). 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 Eq. (2.37b),

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

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

$$(q_k, q_{k+1}) = (\mathbb{F}L_{\rm d}^{-})^{-1}\big(dS_{\rm d}^{k}(q_k)\big) =: (q_k, f_k^L(q_k)), \tag{5.30}$$

where we defined the map $f_k^L : Q \to Q$ as follows:

$$f_k^L(q_k) := pr_2 \circ (\mathbb{F}L_{\rm d}^{-})^{-1}\big(dS_{\rm d}^{k}(q_k)\big)\,, \tag{5.31}$$

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 Eq. (5.29), and then replacing q_k by q, we obtain the discrete Hamilton–Jacobi equation on the Lagrangian side:

$$S_{\rm d}^{k+1}(f_k^L(q)) - S_{\rm d}^{k}(q) = L_{\rm d}\big(q, f_k^L(q)\big)\,. \tag{5.32}$$

The map f_k^L defined in Eq. (5.31) is identical to f_k^- defined above in Eq. (5.22) as the commutative diagram below demonstrates:

$$
\begin{tikzcd}
T^*Q \arrow[rr, "\tilde{F}_{L_{\rm d}}"] \arrow[dr, "(\mathbb{F}L_{\rm d}^{-})^{-1}"'] & & T^*Q \arrow[dd, "\pi_Q"] \\
& Q \times Q \arrow[ur, "\mathbb{F}L_{\rm d}^{+}"'] \arrow[dl, "pr_1"] \arrow[dr, "pr_2"'] & \\
Q \arrow[uu, "dS_{\rm d}^{k}"] \arrow[rr, dashed, "{f_k^L, f_k^-}"'] & & Q
\end{tikzcd}
\qquad
\begin{tikzcd}
dS_{\rm d}^{k}(q_k) \arrow[rr, mapsto] \arrow[dr, mapsto] & & \tilde{F}_{L_{\rm d}}\big(dS_{\rm d}^{k}(q_k)\big) \arrow[dd, mapsto] \\
& (q_k, f_k^L(q_k)) \arrow[ur, mapsto] \arrow[dl, mapsto] \arrow[dr, mapsto] & \\
q_k \arrow[uu, mapsto] \arrow[rr, dashed, mapsto] & & f_k^L(q_k)
\end{tikzcd}
\tag{5.33}
$$

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 Eq. (5.31); note the relation $\tilde{F}_{L_{\rm d}} = \mathbb{F}L_{\rm d}^{+} \circ (\mathbb{F}L_{\rm d}^{-})^{-1}$ from Eq. (2.43). Now f_k^L being identical to f_k^- implies that the discrete Hamilton–Jacobi equations on the Hamiltonian and Lagrangian sides, Eqs. (5.25) and (5.32), are equivalent.

5.3 Discrete Hamilton–Jacobi Theorem

The following gives a discrete analogue of the geometric Hamilton–Jacobi theorem (Theorem 5.2.4) by Abraham and Marsden [1]:

Theorem 5.3.1 (Discrete Hamilton–Jacobi)**.** *Suppose that $S_{\rm d}^{k}$ satisfies the right discrete Hamilton–Jacobi equation* (5.21)*, and let $\{c_k\}_{k=0}^{N} \subset Q$ be a set of points such* *that*

$$c_{k+1} = f_k^+(c_k) \quad \textit{for} \quad k = 0, 1, \ldots, N-1. \tag{5.34}$$

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

$$p_k := DS_{\mathrm{d}}^k(c_k) \tag{5.35}$$

is a solution of the right discrete Hamilton's equations (2.48)*.*

Similarly, suppose that S_{d}^k *satisfies the left discrete Hamilton–Jacobi equation* (5.25)*, and let* $\{c_k\}_{k=0}^{N} \subset Q$ *be a set of points that satisfy*

$$c_{k+1} = f_k^-(c_k) \quad \textit{for} \quad k = 0, 1, \ldots, N-1. \tag{5.36}$$

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*

$$p_k := DS_{\mathrm{d}}^k(c_k) \tag{5.37}$$

is a solution of the left discrete Hamilton's equations (2.51)*.*

Proof. To prove the first assertion, first recall the implicit definition of f_k^+ in Eq. (5.20):

$$f_k^+(q) = D_2 H_{\mathrm{d}}^+\big(q, DS_{\mathrm{d}}^{k+1}(f_k^+(q))\big). \tag{5.38}$$

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

$$c_{k+1} = D_2 H_{\mathrm{d}}^+(c_k, p_k), \tag{5.39}$$

where we used Eq. (5.34) and (5.35). On the other hand, taking the derivative of Eq. (5.21) with respect to q,

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

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 Eq. (5.38). Then substitution $q = c_k$ 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 Eqs. (5.34) and (5.35), we obtain

$$p_k = D_1 H_{\mathrm{d}}^+(c_k, p_{k+1})\,. \tag{5.40}$$

Eqs. (5.39) and (5.40) show that the sequence (c_k, p_k) satisfies the right discrete Hamilton's equations (2.48).

Now let us prove the latter assertion. First recall the implicit definition of f_k^- in Eq. (5.24):

$$q = -D_1 H_{\mathrm{d}}^-\left(DS_{\mathrm{d}}^k(q), f_k^-(q)\right) \tag{5.41}$$

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

$$c_k = -D_1 H_{\mathrm{d}}^-(p_k, c_{k+1})\,, \tag{5.42}$$

where we used Eq. (5.36) and (5.37). On the other hand, taking the derivative of Eq. (5.21) with respect to q,

$$\begin{aligned} &DS_{\mathrm{d}}^{k+1}(f_k^-(q)) \cdot Df_k^-(q) - DS_{\mathrm{d}}^k(q) + D^2 S_{\mathrm{d}}^k(q) \cdot q + DS_{\mathrm{d}}^k(q) \\ &\qquad + D_1 H_{\mathrm{d}}^-\left(DS_{\mathrm{d}}^k(q), f_k^-(q)\right) \cdot D^2 S_{\mathrm{d}}^k(q) + D_2 H_{\mathrm{d}}^-\left(DS_{\mathrm{d}}^k(q), f_k^-(q)\right) \cdot Df_k^-(q) = 0, \end{aligned}$$

which reduces to

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

due to Eq. (5.41). Then substitution $q = c_k$ gives

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

since $Df_k^-(c_k)$ is invertible by assumption. Then using Eqs. (5.36) and (5.37), we obtain

$$p_{k+1} = -D_2 H_{\mathrm{d}}^-(p_k, c_{k+1})\,. \tag{5.43}$$

Eqs. (5.42) and (5.43) show that the sequence (c_k, p_k) satisfies the left discrete Hamilton's equations (2.51). $\square$

5.4 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 (5.21) gives the Bellman equation (discrete-time Hamilton–Jacobi–Bellman equation) as a special case.

5.4.1 Discrete Optimal Control Problem

Let $\{q_k\}_{k=0}^{N}$ be the state variables in a vector space $V \cong \mathbb{R}^n$ with q_0 and q_N fixed and $u_{\rm d} := \{u_k\}_{k=0}^{N}$ be controls in the set $U \subset \mathbb{R}^m$. With a given function $C_{\rm d} : V \times U \to \mathbb{R}$, define the cost functional

$$J_{\rm d} := \sum_{k=0}^{N-1} C_{\rm d}(q_k, u_k). \tag{5.44}$$

Then a typical *discrete optimal control problem* is formulated as follows [see, e.g., 5; 11; 28; 32]:

Problem 5.4.1. Minimize the cost functional, i.e.,

$$\min_{u_{\rm d}} J_{\rm d} = \min_{u_{\rm d}} \sum_{k=0}^{N-1} C_{\rm d}(q_k, u_k) \tag{5.45}$$

subject to the constraint

$$q_{k+1} = f(q_k, u_k). \tag{5.46}$$

5.4.2 Necessary Condition for Optimality and the Discrete-Time HJB Equation

We would like to formulate the necessary condition for optimality. First introduce the augmented cost functional:

$$\begin{aligned}
\hat{J}_{\rm d}^k(q_{\rm d}, p_{\rm d}, u_{\rm d}) &:= \sum_{l=0}^{k-1} \{C_{\rm d}(q_l, u_l) - p_{l+1} \cdot [q_{l+1} - f(q_l, u_l)]\} \\
&= -\sum_{l=0}^{k-1} \Big[p_{l+1} \cdot q_{l+1} - \hat{H}_{\rm d}^{+}(q_l, p_{l+1}, u_l)\Big] \\
&= -\hat{S}_{\rm d}^k(q_{\rm d}, p_{\rm d}, u_{\rm d}),
\end{aligned}$$

where we defined the Hamiltonian

$$\hat{H}_{\mathrm{d}}^{+}(q_l, p_{l+1}, u_l) := p_{l+1} \cdot f(q_l, u_l) - C_{\mathrm{d}}(q_l, u_l), \tag{5.47}$$

and the action sum

$$\hat{S}_{\mathrm{d}}^{k}(q_{\mathrm{d}}, p_{\mathrm{d}}, u_{\mathrm{d}}) := \sum_{l=0}^{k-1} \Big[p_{l+1} \cdot q_{l+1} - \hat{H}_{\mathrm{d}}^{+}(q_l, p_{l+1}, u_l) \Big], \tag{5.48}$$

with the shorthand notation $q_{\mathrm{d}} := \{q_l\}_{l=0}^{k}$, $p_{\mathrm{d}} := \{p_l\}_{l=1}^{k}$, and $u_{\mathrm{d}} := \{u_l\}_{l=0}^{k-1}$. Then the optimality condition Eq. (5.45) is restated as

$$\min_{q_{\mathrm{d}}, p_{\mathrm{d}}, u_{\mathrm{d}}} \hat{J}_{\mathrm{d}}^{k}(q_{\mathrm{d}}, p_{\mathrm{d}}, u_{\mathrm{d}}) = \min_{q_{\mathrm{d}}, p_{\mathrm{d}}, u_{\mathrm{d}}} \left\{ -\sum_{l=0}^{k-1} \Big[p_{l+1} \cdot q_{l+1} - \hat{H}_{\mathrm{d}}^{+}(q_l, p_{l+1}, u_l) \Big] \right\}, \tag{5.49}$$

which is equivalent to

$$\max_{q_{\mathrm{d}}, p_{\mathrm{d}}, u_{\mathrm{d}}} \hat{S}_{\mathrm{d}}^{k}(q_{\mathrm{d}}, p_{\mathrm{d}}, u_{\mathrm{d}}) = \max_{q_{\mathrm{d}}, p_{\mathrm{d}}, u_{\mathrm{d}}} \sum_{l=0}^{k-1} \Big[p_{l+1} \cdot q_{l+1} - \hat{H}_{\mathrm{d}}^{+}(q_l, p_{l+1}, u_l) \Big]. \tag{5.50}$$

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

$$D_3 \hat{H}_{\mathrm{d}}^{+}(q_l, p_{l+1}, u_l) = 0, \quad l = 0, 1, \ldots, k-1. \tag{5.51}$$

Now we assume that $\hat{H}_{\mathrm{d}}^{+}$ is sufficiently regular so that the optimal control $u_{\mathrm{d}}^{*} := \{u_l^{*}\}_{l=0}^{k-1}$ is determined by

$$D_3 \hat{H}_{\mathrm{d}}^{+}(q_l, p_{l+1}, u_l^{*}) = 0, \quad l = 0, 1, \ldots, k-1. \tag{5.52}$$

Therefore u_l^{*} is a function of q_l and p_{l+1}, i.e., $u_l^{*} = u_l^{*}(q_l, p_{l+1})$.

Then we can eliminate u_{d} in the maximization problem Eq. (5.50):

$$\max_{q_{\mathrm{d}}, p_{\mathrm{d}}} S_{\mathrm{d}}(q_{\mathrm{d}}, p_{\mathrm{d}}) = \max_{q_{\mathrm{d}}, p_{\mathrm{d}}} \sum_{l=0}^{k-1} \big[p_{l+1} \cdot q_{l+1} - H_{\mathrm{d}}^{+}(q_l, p_{l+1}) \big], \tag{5.53}$$

where we defined

$$H_{\mathrm{d}}^{+}(q_l, p_{l+1}) := \hat{H}_{\mathrm{d}}^{+}(q_l, p_{l+1}, u_l^{*}) = p_{l+1} \cdot f(q_l, u_l^{*}) - C_{\mathrm{d}}(q_l, u_l^{*}), \tag{5.54}$$

and

$$S_{\mathrm{d}}^{k}(q_{\mathrm{d}}, p_{\mathrm{d}}) := \hat{S}_{\mathrm{d}}^{k}(q_{\mathrm{d}}, p_{\mathrm{d}}, u_{\mathrm{d}}^{*}) = \sum_{l=0}^{k-1} \left[p_{l+1} \cdot q_{l+1} - H_{\mathrm{d}}^{+}(q_l, p_{l+1}) \right]. \tag{5.55}$$

So now the problem is reduced to maximizing the action sum Eq. (5.55) that has exactly the same form as the one in Eq. (5.9) formulated in the framework of discrete Hamiltonian mechanics.

The corresponding right discrete Hamilton's equations are, using the expression for the Hamiltonian in Eq. (5.54),

$$\begin{aligned} q_{k+1} &= f(q_k, u_k^*), \\ p_k &= p_{k+1} \cdot D_1 f(q_k, u_k^*) - D_1 C_{\mathrm{d}}(q_k, u_k^*). \end{aligned} \tag{5.56}$$

Therefore Eq. (5.20) gives the implicit definition of f_k^+ as follows:

$$f_k^+(q_k) = f\big(q_k, u_k^*\big(q_k, DS_{\mathrm{d}}^{k+1}(f_k^+(q_k))\big)\big). \tag{5.57}$$

Hence the (right) discrete Hamilton–Jacobi equation (5.21) applied to this case gives

$$S_{\mathrm{d}}^{k+1}(f(q_k, u_k^*)) - S_{\mathrm{d}}^{k}(q_k) - DS_{\mathrm{d}}^{k+1}(f(q_k, u_k^*)) \cdot f(q_k, u_k^*) + H_{\mathrm{d}}^{+}\big(q_k, DS_{\mathrm{d}}^{k+1}(f(q_k, u_k^*))\big) = 0, \tag{5.58}$$

and again using the expression for the Hamiltonian in Eq. (5.54), we obtain

$$S_{\mathrm{d}}^{k+1}(f(q_k, u_k^*)) - S_{\mathrm{d}}^{k}(q_k) - C_{\mathrm{d}}(q_k, u_k^*) = 0, \tag{5.59}$$

or equivalently

$$\max_{u_k} \left[S_{\mathrm{d}}^{k+1}(f(q_k, u_k)) - C_{\mathrm{d}}(q_k, u_k) \right] - S_{\mathrm{d}}^{k}(q_k) = 0, \tag{5.60}$$

which is the *discrete-time Hamilton–Jacobi–Bellman (HJB) equation* or, in short, the *Bellman equation* [see, e.g., 5].

Remark 5.4.2. Notice that the discrete HJB equation (5.60) is much simpler than the discrete Hamilton–Jacobi equations (5.21) and (5.25) because of the special form of the control Hamiltonian Eq. (5.54). Also notice that, as shown in Eq. (5.57), the term $f_k^+(q_k)$ is written in terms of the given function f. See Remark 5.2.7 for comparison.

5.4.3 Relation between the Discrete HJ and HJB Equations and its Consequences

Summarizing the observation made above, we have

Proposition 5.4.3. *The right discrete Hamilton–Jacobi equation* (5.21) *applied to the Hamiltonian formulation of the discrete optimal control problem 5.4.1 gives the discrete-time Hamilton–Jacobi–Bellman equation* (5.60)*.*

This observation leads to the following well-known facts:

Proposition 5.4.4. *The optimal cost function satisfies the discrete-time Hamilton–Jacobi–Bellman equation* (5.60)*.*

Proof. This follows from a reinterpretation of Theorem 5.2.3 through Proposition 5.4.3. □

Proposition 5.4.5. *Let $S_{\mathrm{d}}^k(q_k)$ be a solution to the discrete Hamilton–Jacobi–Bellman equation* (5.60)*. Then the costate p_k in the discrete maximum principle is given as follows:*

$$p_k = DS_{\mathrm{d}}^k(c_k), \tag{5.61}$$

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

Proof. This follows from a reinterpretation of Theorem 5.3.1 through Proposition 5.4.3. □

5.5 Application To Discrete Linear Hamiltonian Systems

5.5.1 Discrete Linear Hamiltonian Systems and Matrix Riccati Equation

Example 5.5.1 (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

$$H_{\mathrm{d}}^-(p_k, q_{k+1}) = \frac{1}{2} p_k^T M^{-1} p_k + p_k^T L q_{k+1} + \frac{1}{2} q_{k+1}^T K q_{k+1}, \tag{5.62}$$

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 (2.51) are

$$\begin{aligned} q_k &= -(M^{-1} p_k + L q_{k+1}), \\ p_{k+1} &= -(L^T p_k + K q_{k+1}), \end{aligned} \tag{5.63}$$

or

$$\begin{pmatrix} q_{k+1} \\ p_{k+1} \end{pmatrix} = \begin{pmatrix} -L^{-1} & -L^{-1} M^{-1} \\ K L^{-1} & K L^{-1} M^{-1} - L^T \end{pmatrix} \begin{pmatrix} q_k \\ p_k \end{pmatrix}, \tag{5.64}$$

and hence are a discrete linear Hamiltonian system (see Section 5.A.1).

Now let us solve the left discrete Hamilton–Jacobi equation (5.25) 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 5.A.2) which contains the point $z_0 := (q_0, p_0)$. Then the dynamics is formally written as, for any discrete time $k \in \mathbb{N}$,

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

where $\tilde{F}_{L_{\mathrm{d}}} : T^*Q \to T^*Q$ is the discrete Hamiltonian map defined in Eq. (2.42). Since $\tilde{F}_{L_{\mathrm{d}}}$ is a symplectic map, Proposition 5.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 5.A.6 implies that there exists a set of functions S_{d}^k of the form

$$S_{\mathrm{d}}^k(q) = \frac{1}{2} q^T A_k q + b_k^T q + c_k \tag{5.65}$$

such that $\tilde{\mathcal{L}}_k = \operatorname{graph} dS_{\mathrm{d}}^k$; 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 Eq. (5.64) with p_k replaced by $DS_{\mathrm{d}}^k(q)$:

$$\begin{aligned} f_k^-(q) &= -L^{-1}\big(q + M^{-1} DS_{\mathrm{d}}^k(q)\big) \\ &= -L^{-1}(I + M^{-1} A_k) q - L^{-1} M^{-1} b_k. \end{aligned} \tag{5.66}$$

Then substituting Eq. (5.65) into the left-hand side of the left discrete Hamilton–

Jacobi equation (5.25) yields the following recurrence relations for A_k, b_k, and c_k:

$$A_{k+1} = L^T(I + A_k M^{-1})^{-1} A_k L - K, \tag{5.67a}$$

$$b_{k+1} = -L^T(I + A_k M^{-1})^{-1} b_k, \tag{5.67b}$$

$$c_{k+1} = c_k - \frac{1}{2} b_k^T (M + A_k)^{-1} b_k, \tag{5.67c}$$

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

Remark 5.5.2. For the A_{k+1} defined by Eq. (5.67a) to be symmetric, it is sufficient that A_k is invertible; for if it is, then Eq. (5.67a) 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 5.5.3. We can rewrite Eq. (5.67a) as follows:

$$A_{k+1} = \left[KL^{-1} + (KL^{-1}M^{-1} - L^T)A_k\right](-L^{-1} - L^{-1}M^{-1}A_k)^{-1}. \tag{5.68}$$

Notice the exact correspondence between the coefficients in the above equation and the matrix entries in the discrete linear Hamiltonian equations (5.64). In fact, this is the discrete Riccati equation that corresponds to the iteration defined by Eq. (5.64). See Ammar and Martin [2] for details on this correspondence.

To summarize the above observation, we have

Proposition 5.5.4. *The discrete Hamilton–Jacobi equation* (5.25) *applied to the discrete linear Hamiltonian system* (5.64) *yields the discrete Riccati equation* (5.68)*.*

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

A simple physical example that is described as a discrete linear Hamiltonian system is the following:

Example 5.5.5 (Harmonic oscillator)**.** Consider the one-dimensional harmonic oscillator with mass m and spring constant k. The configuration space is a real line, i.e., $Q = \mathbb{R}$, and the Lagrangian $L : T\mathbb{R} \to \mathbb{R}$ of the system is

$$L(q, \dot{q}) = \frac{m}{2}\dot{q}^2 + \frac{k}{2}q^2.$$

Introducing the angular frequency $\omega := \sqrt{k/m}$, we have

$$L(q, \dot{q}) = \frac{m}{2}\left(\dot{q}^2 + \omega^2 q^2\right).$$

It is easy to solve the (continuous) Euler–Lagrange equation and calculate Jacobi's solution explicitly:

$$S(q, t; q_0) := \int_0^t L(q(s), \dot{q}(s))\, ds = \frac{1}{2} m\omega \left[(q_0^2 + q^2)\cot(\omega t) - 2 q_0 q \csc(\omega t)\right], \tag{5.69}$$

where q_0 is the initial position: $q(0) = q_0$. This gives the exact discrete Lagrangian [48] with step size h as follows:

$$L_{\mathrm{d}}^{\mathrm{ex}}(q_k, q_{k+1}) = S(q_{k+1}, h; q_k) = \frac{1}{2} m\omega \left[(q_k^2 + q_{k+1}^2)\cot(\omega h) - 2 q_k q_{k+1} \csc(\omega h)\right]. \tag{5.70}$$

The corresponding left discrete Hamiltonian (see Eq. (2.50)), which we shall call the *exact left discrete Hamiltonian*, is then

$$H_{\mathrm{d,ex}}^{-}(p_k, q_{k+1}) = \frac{1}{2}\left[\frac{p_k^2}{m\omega}\tan(\omega h) - 2 p_k q_{k+1} \sec(\omega h) + m\omega\, q_{k+1}^2 \tan(\omega h)\right]. \tag{5.71}$$

Comparing this with the general form of the quadratic Hamiltonian Eq. (5.62), we see that this is a special case with $n = 1$ and

$$M^{-1} = \frac{\tan(\omega h)}{m\omega}, \qquad L = -\sec(\omega h), \qquad K = m\omega \tan(\omega h).$$

Note that M, L, and K are also scalars now. Thus Eq. (5.66) gives

$$f_k^{-}(q) := \pi_{\mathbb{R}} \circ \tilde{F}_{L_{\mathrm{d}}}\left(dS_{\mathrm{d}}^k(q)\right) = \left(\cos(\omega h) + \frac{\sin(\omega h)}{m\omega} A_k\right) q + \frac{\sin(\omega h)}{m\omega} b_k. \tag{5.72}$$

Now the recurrence relations Eq. (5.67) reduce to

$$\begin{aligned} A_{k+1} &= \frac{m\omega[A_k \cos(\omega h) - m\omega \sin(\omega h)]}{m\omega \cos(\omega h) + A_k \sin(\omega h)}, \\ b_{k+1} &= \frac{m\omega}{m\omega \cos(\omega h) + A_k \sin(\omega h)}\, b_k, \\ c_{k+1} &= c_k - \frac{b_k^2}{A_k + m\omega \cot(\omega h)}. \end{aligned} \tag{5.73}$$

We impose the "initial condition" $S_{\mathrm{d}}^1(q_1) = L_{\mathrm{d}}^{\mathrm{ex}}(q_0, q_1)$, which follows from Eq. (5.9) or (5.15) for $k = 1$. This gives

$$A_1 = m\omega \cot(\omega h), \qquad b_1 = -m\omega q_0 \csc(\omega h), \qquad c_1 = m\omega q_0^2 \cot(\omega h). \tag{5.74}$$

Solving the above recurrence relations using *Mathematica*, we obtain

$$A_k = m\omega \cot(\omega k h), \qquad b_k = -m\omega q_0 \csc(\omega k h), \qquad c_k = m\omega q_0^2 \cot(\omega k h), \tag{5.75}$$

and hence the solution of the left discrete Hamilton–Jacobi equation

$$S_{\mathrm{d}}^k(q) = \frac{1}{2} m\omega \left[(q_0^2 + q^2) \cot(\omega k h) - 2 q_0 q \csc(\omega k h) \right]. \tag{5.76}$$

Remark 5.5.6. Notice that, in the above example, we have $S_{\mathrm{d}}^k(q) = S(q, kh; q_0)$ from the explicit expression for Jacobi's solution Eq. (5.69) under the assumption that $q = q_k$. This is because we started with the exact discrete Lagrangian and hence the corresponding discrete dynamics is exact. Specifically, the exact discrete Lagrangian satisfies, by definition,

$$L_{\mathrm{d}}^{\mathrm{ex}}(q_l, q_{l+1}) = \int_{lh}^{(l+1)h} L(q(t), \dot{q}(t))\, dt, \qquad l \in \{0, 1, \dots, k-1\} \tag{5.77}$$

where $q(t)$ satisfies the continuous dynamics and the boundary conditions $q(lh) = q_l$ and $q((l+1)h) = q_{l+1}$. Hence

$$S_{\mathrm{d}}^k(q) := \sum_{l=0}^{k-1} L_{\mathrm{d}}^{\mathrm{ex}}(q_l, q_{l+1}) = \int_0^{kh} L(q(t), \dot{q}(t))\, dt =: S(q, kh; q_0), \tag{5.78}$$

which says that the discrete analogue of Jacobi's solution Eq. (5.9) is identical to Jacobi's solution Eq. (5.69) calculated using the continuous dynamics.

5.5.2 Application of the Hamilton–Jacobi Theorem

We illustrate how Theorem 5.3.1 works using the same example. Here we would like to see if we can "generate" the dynamics using the solution of the discrete Hamilton–Jacobi equations as in Theorem 5.3.1.

Example 5.5.7 (Harmonic oscillator)**.** Let us start from the solution obtained in

Example 5.5.5:

$$S_{\mathrm{d}}^{k}(q)=\frac{1}{2}m\omega\left[(q_0^2+q^2)\cot(\omega kh)-2q_0q\csc(\omega kh)\right]. \tag{5.79}$$

Notice that the expression for the right-hand side of Eq. (5.36) was already given in Eq. (5.72):

$$\pi_Q\circ\tilde{F}_{L_{\mathrm{d}}}\big(dS_{\mathrm{d}}^{k}(q_k)\big)=q_k\cos(\omega h)+\frac{1}{m\omega}\,DS_{\mathrm{d}}^{k}(q_k)\sin(\omega h).$$

Hence substituting Eq. (5.79) into Eq. (5.36) yields

$$q_{k+1}=\csc(\omega kh)\{q_k\sin[\omega(k+1)h]-q_0\sin(\omega h)\}. \tag{5.80}$$

Then Eq. (5.37) gives

$$p_k=DS_{\mathrm{d}}^{k}(q_k)=m\omega\csc(\omega kh)[q_k\cos(\omega kh)-q_0]. \tag{5.81}$$

It is easy to check these equations satisfy the left discrete Hamilton's equations (2.48) as Theorem 5.3.1 claims.

5.6 Continuous Limit

This section shows that the right and left discrete Hamilton–Jacobi equations (5.21) and (5.25) recover the original Hamilton–Jacobi equation (5.14) in the continuous-time limit. We reproduce the result of Elnatanov and Schiff [22] on the continuous limit of the right discrete Hamilton–Jacobi equation, applying the same argument simultaneously to the left discrete Hamilton–Jacobi equation. The main purpose of doing so here is to make it clear how the discrete ingredients are related to the corresponding continuous ones in our notation.

5.6.1 Continuous Limit of Discrete Hamilton's Equations

Let us first look at the continuous-time limit of the right and left discrete Hamilton's equations (2.48) and (2.51). This makes it clear how the discrete and continuous Hamiltonians are related in the limit. First recall from Section 2.3 of Marsden and

West [48] that the discrete Lagrangian $L_{\mathrm{d}}(q_k, q_{k+1})$ is consistent if it satisfies

$$\begin{aligned} L_{\mathrm{d}}(q_k, q_{k+1}) &= \int_{t_k}^{t_{k+1}} L(q(t), \dot{q}(t))\, dt + O(h^2) \\ &= \int_{t_k}^{t_{k+1}} [p(t) \cdot \dot{q}(t) - H(q(t), p(t))]\, dt + O(h^2). \end{aligned} \tag{5.82}$$

where $t_k = kh$, and the $(q(t), p(t))$ in the integrand is the flow defined by the continuous Lagrangian or Hamiltonian with $q(t_k) = q_k$ and $q(t_{k+1}) = q_k$. Consistency of a discrete Lagrangian implies that of the corresponding discrete flow, hence the terminology.

Lemma 5.6.1. *The right and left discrete Hamiltonians $H_{\mathrm{d}}^{\pm}$ defined in Eq.* (2.47) *and* (2.50) *with a consistent discrete Lagrangian satisfies the following relations with the continuous Hamiltonian:*

$$H(q_k, p_k) = \lim_{h \to 0} \frac{1}{h} \left[H_{\mathrm{d}}^{+}(q_k, p_{k+1}) - p_{k+1} \cdot q_k \right] = \lim_{h \to 0} \frac{1}{h} \left[H_{\mathrm{d}}^{-}(p_k, q_{k+1}) + p_k \cdot q_{k+1} \right]. \tag{5.83}$$

Proof. Simple calculations with Eqs. (2.47) and (2.50) with Eq. (5.82) show

$$\begin{aligned} \frac{1}{h} \left[H_{\mathrm{d}}^{+}(q_k, p_{k+1}) - p_{k+1} \cdot q_k \right] &= p_{k+1} \cdot \frac{q_{k+1} - q_k}{h} \\ &\quad - \frac{1}{h} \int_{t_k}^{t_k + h} [p(t) \cdot \dot{q}(t) - H(q(t), p(t))]\, dt + O(h) \end{aligned}$$

and

$$\begin{aligned} \frac{1}{h} \left[H_{\mathrm{d}}^{-}(p_k, q_{k+1}) + p_k \cdot q_{k+1} \right] &= p_k \cdot \frac{q_{k+1} - q_k}{h} \\ &\quad - \frac{1}{h} \int_{t_k}^{t_k + h} [p(t) \cdot \dot{q}(t) - H(q(t), p(t))]\, dt + O(h). \end{aligned}$$

Taking the limit as $h \to 0$ on both sides in each of the above equations gives the result. □

Definition 5.6.2. We shall say that a right/left discrete Hamiltonian $H_{\mathrm{d}}^{\pm}$ is *consistent* if it satisfies Eq. (5.83).

Proposition 5.6.3. *With consistent discrete Hamiltonians, the right and left discrete Hamilton's equations* (2.48) *and* (2.51) *recover the continuous-time Hamilton's equations in the continuous limit.*

Proof. Simple calculations with Eqs. (2.48) and (2.51) show

$$\begin{aligned}
\frac{q_{k+1}-q_k}{h} &= \frac{\partial}{\partial p_k}\left\{\frac{1}{h}\big[H_{\mathrm{d}}^{+}(q_k,p_{k+1}) - p_{k+1}\cdot q_k\big]\right\},\\
\frac{p_{k+1}-p_k}{h} &= -\frac{\partial}{\partial q_{k+1}}\left\{\frac{1}{h}\big[H_{\mathrm{d}}^{+}(q_k,p_{k+1}) - p_{k+1}\cdot q_k\big]\right\}
\end{aligned}$$

and

$$\begin{aligned}
\frac{q_{k+1}-q_k}{h} &= \frac{\partial}{\partial p_{k+1}}\left\{\frac{1}{h}\big[H_{\mathrm{d}}^{-}(p_k,q_{k+1}) + p_k\cdot q_{k+1}\big]\right\},\\
\frac{p_{k+1}-p_k}{h} &= -\frac{\partial}{\partial q_k}\left\{\frac{1}{h}\big[H_{\mathrm{d}}^{-}(p_k,q_{k+1}) + p_k\cdot q_{k+1}\big]\right\}.
\end{aligned}$$

Taking the limit as $h \to 0$ on both sides in each of the above equations gives, with Eq. 5.83,

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

5.6.2 Continuous Limit of Discrete Hamilton–Jacobi Equations

Now we are ready to discuss the continuous limit of the right and left discrete Hamilton–Jacobi equations.

Proposition 5.6.4. *With consistent discrete Hamiltonians, the right and left discrete Hamilton–Jacobi equations* (5.3) *and* (5.18) *recover the continuous-time Hamilton–Jacobi equation.*

Proof. First define $S : Q \times \mathbb{R} \to \mathbb{R}$ that satisfies $S(q_k,t_k) = S_{\mathrm{d}}^{k}(q_k)$. Simple calculations with (5.3) and (5.18) yield

$$\begin{aligned}
&\frac{1}{h}\left[S(q_{k+1},t_{k+1}) - S(q_k,t_k) - \frac{\partial S}{\partial q}(q_{k+1},t_{k+1})\cdot(q_{k+1}-q_k)\right]\\
&\qquad + \frac{1}{h}\left[H_{\mathrm{d}}^{+}\left(q_k, \frac{\partial S}{\partial q}(q_{k+1},t_{k+1})\right) - \frac{\partial S}{\partial q}(q_{k+1},t_{k+1})\cdot q_k\right] = 0
\end{aligned} \tag{5.84}$$

and

$$\begin{aligned}
&\frac{1}{h}\left[S(q_{k+1},t_{k+1}) - S(q_k,t_k) - \frac{\partial S}{\partial q}(q_{k+1},t_{k+1})\cdot(q_{k+1}-q_k)\right]\\
&\qquad + \frac{1}{h}\left[H_{\mathrm{d}}^{-}\left(\frac{\partial S}{\partial q}(q_k,t_k), q_{k+1}\right) + \frac{\partial S}{\partial q}(q_k,t_k)\cdot q_{k+1}\right] = 0.
\end{aligned} \tag{5.85}$$

The first group of the terms in brackets is common to both of the above equations. Taylor expansion of the terms gives

$$\begin{aligned}&\frac{1}{h}\left[S(q_{k+1},t_{k+1}) - S(q_k,t_k) - \frac{\partial S}{\partial q}(q_{k+1},t_{k+1})\cdot(q_{k+1}-q_k)\right]\\&\quad= \frac{\partial S}{\partial t}(q_k,t_k) + \left[\frac{\partial S}{\partial q}(q_k,t_k) - \frac{\partial S}{\partial q}(q_{k+1},t_{k+1})\right]\cdot\frac{q_{k+1}-q_k}{h} + O(h) \to \frac{\partial S}{\partial t}(q_k,t_k)\end{aligned}$$

as $h \to 0$. On the other hand, by Lemma 5.6.1, the limit as $h \to 0$ of the second group of the terms in each of Eqs. (5.84) and (5.85) is

$$\lim_{h\to 0}\frac{1}{h}\left[H_{\rm d}^{+}\left(q_k, \frac{\partial S}{\partial q}(q_{k+1},t_{k+1})\right) - \frac{\partial S}{\partial q}(q_{k+1},t_{k+1})\cdot q_k\right] = H\left(q_k, \frac{\partial S}{\partial q}(q_k,t_k)\right),$$

and

$$\lim_{h\to 0}\frac{1}{h}\left[H_{\rm d}^{-}\left(\frac{\partial S}{\partial q}(q_k,t_k), q_{k+1}\right) + \frac{\partial S}{\partial q}(q_k,t_k)\cdot q_{k+1}\right] = H\left(q_k, \frac{\partial S}{\partial q}(q_k,t_k)\right).$$

As a result, both the right and left discrete Hamilton–Jacobi equations give, in the limit as $h \to 0$,

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

which is the continuous-time Hamilton–Jacobi equation. □

5.A Discrete Linear Hamiltonian Systems

5.A.1 Discrete Linear Hamiltonian Systems

Suppose that the configuration space Q is an n-dimensional vector space, and that the discrete Hamiltonian $H_{\rm d}^{+}$ or $H_{\rm d}^{-}$ is quadratic as in Eq. (5.62). Also assume that the corresponding discrete Hamiltonian map $\tilde{F}_{L_{\rm d}} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$ is invertible. Then the discrete Hamilton's equations (2.48) or (2.51) reduce to the discrete linear Hamiltonian system

$$z_{k+1} = A_{L_{\rm d}} z_k, \tag{5.86}$$

where $z_k \in \mathbb{R}^{2n}$ is a coordinate expression for $(q_k, p_k) \in Q \times Q^*$ and $A_{L_{\rm d}} : Q \times Q^* \to Q \times Q^*$ is the matrix representation of the map $\tilde{F}_{L_{\rm d}}$ under the same basis. Since $\tilde{F}_{L_{\rm d}}$

is symplectic, A_{L_d} is an $2n \times 2n$ symplectic matrix, i.e.,

$$A_{L_\mathrm{d}}^T \mathbb{J} A_{L_\mathrm{d}} = \mathbb{J}, \tag{5.87}$$

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.

5.A.2 Lagrangian Subspaces and Lagrangian Affine Spaces

First recall the definition of a Lagrangian subspace:

Definition 5.A.1. Let V be a symplectic vector space with the symplectic form Ω. 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 5.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., 34, Theorem 6 on p. 417]:

Proposition 5.A.3. *Let $\mathcal{L}$ be a Lagrangian subspace of V and $A : V \to V$ be a symplectic transformation. Then $A^k(\mathcal{L})$ is also a Lagrangian subspace of V for any $k \in \mathbb{N}$.*

A similar result holds for Lagrangian affine spaces:

Proposition 5.A.4. *Let $\tilde{\mathcal{L}}(b) = b + \mathcal{L}$ be a Lagrangian affine space of V and $A : V \to V$ be a symplectic transformation. Then $A^k\left(\tilde{\mathcal{L}}(b)\right)$ is also a Lagrangian affine space of V for any $k \in \mathbb{N}$. More explicitly, we have*

$$A^k\left(\tilde{\mathcal{L}}(b)\right) = A^k b + A^k(\mathcal{L}).$$

Proof. Follows from a straightforward calculation. $\square$

5.A.3 Generating Functions

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

$$S(q) = \frac{1}{2} \langle Aq, q \rangle + C \tag{5.88}$$

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-forms on Q [See 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, and so 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 [34, 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 $\mathcal{L} = \operatorname{graph} A$. Hence $\operatorname{graph} dS = \operatorname{graph} A$, 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. $\square$

Corollary 5.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*

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

with an arbitrary real scalar constant C.

Proof. From the above proposition, there exists a function $S : Q \to \mathbb{R}$ of the form Eq. (5.88) such that $\mathcal{L} = \operatorname{graph} dS$. Let $\tilde{S} : Q \to R$ be defined by $\tilde{S}(q) := S(q - q_0) + \langle p_0, q \rangle$. Then

$$d\tilde{S}(q) = A(q - q_0) + p_0. \tag{5.89}$$

and thus

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

The form Eq. (5.89) follows from a direct calculation. □

Chapter 6

Conclusion and Future Work

6.1 Nonholonomic Hamilton–Jacobi Theory

We formulated a nonholonomic Hamilton–Jacobi theorem building on the work by Iglesias-Ponte et al. [30] with a particular interest in the application to exactly integrating the equations of motion of nonholonomic mechanical systems. In particular we formulated the theorem so that the technique of separation of variables applies as in the unconstrained theory. We illustrated how this works for the vertical rolling disk and knife edge. Furthermore, we proposed another way of exactly integrating the equations of motion without using separation of variables.

We also applied the conventional Hamilton–Jacobi equation to the Chaplygin-Hamiltonized nonholonomic system and obtained the Chaplygin Hamilton–Jacobi equation. We obtained an explicit formula that provides a link between the solutions of the Chaplygin Hamilton–Jacobi and nonholonomic Hamilton–Jacobi equations. This result relates the two seemingly distinct approaches to extending Hamilton–Jacobi theory to nonholonomic systems.

The following topics are interesting to consider for future work:

- *Role of symmetry in nonholonomic Hamilton–Jacobi Theory.* Many nonholonomic systems possess symmetry, and there are theories on nonholonomic reduction by symmetry [4; 8; 36; 38; 39]. Introducing the ideas of symmetry and reduction to nonholonomic Hamilton–Jacobi theory is certainly appealing. Iglesias-Ponte et al. [30] applied their Hamilton–Jacobi theorem to the so-called Chaplygin case to prove a reduced version of the theorem. Our preliminary calculations with simple examples showed that symmetry consideration leads to the assumptions made in constructing the ansatz, e.g., Eq. (3.20) for the vertical rolling disk. We are interested in exploring this idea to gain insights into integrability of nonholonomic systems.

- *Relation between measure-preservation and applicability of separations of variables.* The integrability conditions of nonholonomic systems formulated by Kozlov [40] include measure-preservation. As mentioned above, applicability of separation of variables implies the existence of conserved quantities other than the Hamiltonian. Therefore it is interesting to see how these ideas, i.e., measure-preservation, applicability of separation of variables, and existence of conserved quantities, are related to each other.

- *"Right" coordinates in nonholonomic Hamilton–Jacobi theory and relation to quasivelocities.* In the unconstrained Hamilton–Jacobi theory, there are examples which are solvable by separation of variables only after a certain coordinate transformation. As a matter of fact, Lanczos [42, p. 243] says "The separable nature of a problem constitutes no inherent feature of the physical properties of a mechanical system, but is entirely a matter of the *right system of coordinates*." It is reasonable to expect the same situation in nonholonomic Hamilton–Jacobi theory. In fact the equations of nonholonomic mechanics take simpler forms with the use of quasivelocities [10; 17]. Relating the "right" coordinates, if any, to the quasivelocities is an interesting question to consider.

- *Extension to Dirac mechanics.* Implicit Lagrangian/Hamiltonian systems defined with Dirac structures [57; 60; 61] can incorporate more general constraints than nonholonomic constraints including those from degenerate Lagrangians and Hamiltonians, and give nonholonomic mechanics as a special case. A generalization of Hamilton–Jacobi theory to such systems is in progress [45].

6.2 Discrete Hamilton–Jacobi Theory

We developed a discrete-time analogue of Hamilton–Jacobi theory starting from the discrete variational Hamilton equations formulated by Lall and West [41]. We reinterpreted and extended the discrete Hamilton–Jacobi equation given by Elnatanov and Schiff [22] to show that it possesses theoretical significance in discrete mechanics that is equivalent to that of the (continuous-time) Hamilton–Jacobi equation in Hamiltonian 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 discrete optimal control problems. This again gives discrete analogues of the corresponding known results in the continuous-time theory. Application to discrete optimal control also revealed that Theorems 5.2.3 and 5.3.1 specialize to two well-known results in discrete optimal control theory.

We are interested in the following topics for future work:

- *Application to integrable discrete systems.* Theorem 5.3.1 gives a discrete analogue of the theory behind the technique of solution by separation of variables in the sense that 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 [50] and various others discussed by Suris [55, 56].

- *Development of numerical methods based on the discrete Hamilton–Jacobi equation.* Hamilton–Jacobi equation has been used to develop structured integrators for Hamiltonian systems [see, e,g., 15, and also references therein]. The present theory, being intrinsically 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 nonholonomic Hamilton–Jacobi theory [19; 30; 51] and Dirac Hamilton–Jacobi theory [45], are another direction of 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 5.5.1 and 5.A.2 is a special case of such links. On the other hand, Proposition 5.5.4 suggests that the discrete Hamilton–Jacobi equation is a generalization of the Riccati equation. Interpreted in the context of the result by Ammar and Martin [2], the discrete Hamilton–Jacobi equation defines an iteration of Lagrangian submanifolds. We are interested in seeing possible further links provided by the generalization, such as the relationship between discrete Hamiltonian dynamical systems and iterations of Lagrangian submanifolds, and its applications to numerical methods.

Bibliography

[1] R. Abraham and J. E. Marsden. *Foundations of Mechanics*. Addison–Wesley, 2nd edition, 1978.

[2] G. Ammar and C. Martin. The geometry of matrix eigenvalue methods. *Acta Applicandae Mathematicae: An International Survey Journal on Applying Mathematics and Mathematical Applications*, 5(3):239–278, 1986.

[3] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer, 1989.

[4] L. Bates and J. Sniatycki. Nonholonomic reduction. *Reports on Mathematical Physics*, 32(1):99–115, 1993.

[5] R. Bellman. *Introduction to the Mathematical Theory of Control Processes*, volume 2. Academic Press, 1971.

[6] A. M. Bloch. Asymptotic Hamiltonian dynamics: the Toda lattice, the three-wave interaction and the non-holonomic Chaplygin sleigh. *Physica D: Nonlinear Phenomena*, 141(3-4):297–315, 2000.

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

[8] A. M. Bloch, P. S. Krishnaprasad, J. E. Marsden, and R. M. Murray. Nonholonomic mechanical systems with symmetry. *Archive for Rational Mechanics and Analysis*, 136:21–99, 1996.

[9] A. M. Bloch, O. E. Fernandez, and T. Mestdag. Hamiltonization of nonholonomic systems and the inverse problem of the calculus of variations. *Preprint*, 2008.

[10] A. M. Bloch, J. E. Marsden, and D. V. Zenkov. Quasivelocities and symmetries in non-holonomic systems. *Dynamical Systems: An International Journal*, 24 (2):187–222, 2009.

[11] J. A. Cadzow. Discrete calculus of variations. *International Journal of Control*, 11(3):393–407, 1970.

[12] F. Cantrijn, J. Cortés, M. de León, and D. Martín de Diego. On the geometry of generalized Chaplygin systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 132(02):323–351, 2002.

[13] J. F. Cariñena and M. F. Rañada. Lagrangian systems with constraints: a geometric approach to the method of Lagrange multipliers. *Journal of Physics A: Mathematical and General*, 26(6):1335–1351, 1993.

[14] J. F. Cariñena, X. Gracia, G. Marmo, E. Martínez, M. Munõz Lecanda, and N. Román-Roy. Geometric Hamilton–Jacobi theory for nonholonomic dynamical systems. *Preprint*, (`arXiv:0908.2453`).

[15] P. J. Channell and C. Scovel. Symplectic integration of Hamiltonian systems. *Nonlinearity*, 3(2):231–259, 1990.

[16] S. Chaplygin. On the theory of motion of nonholonomic systems. the reducing-multiplier theorem. *Regular and Chaotic Dynamics*, 13(4):369–376, 2008. (English translation of Matematicheskiĭ sbornik, 1911, vol. 28, issue 1).

[17] J. Cortés, M. de León, J. C. Marrero, and E. Martínez. Nonholonomic Lagrangian systems on Lie algebroids. *Discrete and Continuous Dynamical Systems - A*, 24: 213–271, 2009.

[18] M. de León and D. Martín de Diego. On the geometry of non-holonomic Lagrangian systems. *Journal of Mathematical Physics*, 37(7):3389–3414, 1996.

[19] M. de León, J. C. Marrero, and D. Martín de Diego. Linear almost Poisson structures and Hamilton–Jacobi theory. applications to nonholonomic mechanics. *Preprint*, (`arXiv:0801.4358`).

[20] M. de León, J. C. Marrero, and D. Martín de Diego. Mechanical systems with nonlinear constraints. *International Journal of Theoretical Physics*, 36(4):979–995, 1997.

[21] K. M. Ehlers, J. Koiller, R. Montgomery, and P. M. Rios. Nonholonomic systems via moving frames: Cartan equivalence and Chaplygin Hamiltonization. In *The Breadth of Symplectic and Poisson Geometry*, pages 75–120. Birkhäuser, 2004.

[22] N. A. Elnatanov and J. Schiff. The Hamilton–Jacobi difference equation. *Functional Differential Equations*, 3(279–286), 1996.

[23] Y. Fedorov and B. Jovanović. Hamiltonization of the generalized Veselova LR system. *Regular and Chaotic Dynamics*, 14(4):495–505, 2009.

[24] Y. N. Fedorov and B. Jovanović. Nonholonomic LR systems as generalized Chaplygin systems with an invariant measure and flows on homogeneous spaces. *Journal of Nonlinear Science*, 14(4):341–381, 2004.

[25] O. Fernandez, T. Mestdag, and A. Bloch. A generalization of Chaplygin's reducibility theorem. *Regular and Chaotic Dynamics*, 14(6):635–655, 2009.

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

[27] H. Goldstein, C. P. Poole, and J. L. Safko. *Classical Mechanics*. Addison Wesley, 3rd edition, 2001.

[28] V. Guibout and A. M. Bloch. A discrete maximum principle for solving optimal control problems. In *43rd IEEE Conference on Decision and Control*, volume 2, pages 1806–1811 Vol.2, 2004.

[29] S. Hochgerner and L. García-Naranjo. G-Chaplygin systems with internal symmetries, truncation, and an (almost) symplectic view of Chaplygin's ball. *Journal of Geometric Mechanics*, 1(1):35–53, 2009.

[30] D. Iglesias-Ponte, M. de León, and D. Martín de Diego. Towards a Hamilton–Jacobi theory for nonholonomic mechanical systems. *Journal of Physics A: Mathematical and Theoretical*, 41(1), 2008.

[31] S. M. Jalnapurkar, M. Leok, J. E. Marsden, and M. West. Discrete Routh reduction. *Journal of Physics A: Mathematical and General*, 39(19):5521–5544, 2006.

[32] B. W. Jordan and E. Polak. Theory of a class of discrete optimal control systems. *Journal of Electronics and Control*, 17:694–711, 1964.

[33] J. V. José and E. J. Saletan. *Classical Dynamics: A Contemporary Approach*. Cambridge University Press, 1998.

[34] V. Jurdjevic. *Geometric control theory*. Cambridge University Press, Cambridge, 1997.

[35] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*. Interscience, New York, 1963.

[36] J. Koiller. Reduction of some classical non-holonomic systems with symmetry. *Archive for Rational Mechanics and Analysis*, 118(2):113–148, 1992.

[37] J. Koiller, P. P. M. Rios, and K. M. Ehlers. Moving frames for cotangent bundles. *Reports on Mathematical Physics*, 49(2-3):225–238, 2002.

[38] W. S. Koon and J. E. Marsden. The Hamiltonian and Lagrangian approaches to the dynamics of nonholonomic systems. *Reports on Mathematical Physics*, 40 (1):21–62, 1997.

[39] W. S. Koon and J. E. Marsden. Poisson reduction for nonholonomic mechanical systems with symmetry. *Reports on Mathematical Physics*, 42(1-2):101–134, 1998.

[40] V. V. Kozlov. On the integration theory of equations of nonholonomic mechanics. *Regular and Chaotic Dynamics*, 7(2):161–176, 2002.

[41] S. Lall and M. West. Discrete variational Hamiltonian mechanics. *Journal of Physics A: Mathematical and General*, 39(19):5509–5519, 2006.

[42] C. Lanczos. *The Variational Principles of Mechanics*. Dover, 4th edition, 1986.

[43] M. Leok. *Foundations of Computational Geometric Mechanics*. PhD thesis, California Institute of Technology, 2004.

[44] M. Leok and J. Zhang. Discrete Hamiltonian variational integrators. *Preprint (*`arXiv:1001.1408`*)*, 2010.

[45] M. Leok, T. Ohsawa, and D. Sosa. Dirac Hamilton–Jacobi theory for implicit Lagrangian and Hamiltonian systems. *in preparation*.

[46] M. Leok, J. E. Marsden, and A. Weinstein. A discrete theory of connections on principal bundles. *Preprint*, 2004.

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

[48] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numerica*, pages 357–514, 2001.

[49] R. Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Applications*. American Mathematical Society, 2002.

[50] J. Moser and A. P. Veselov. Discrete versions of some classical integrable systems and factorization of matrix polynomials. *Communications in Mathematical Physics*, 139(2):217–243, 1991.

[51] T. Ohsawa and A. M. Bloch. Nonholonomic Hamilton–Jacobi equation and integrability. *Journal of Geometric Mechanics*, 1(4):461–481, 2009.

[52] M. Pavon. Hamilton–Jacobi equations for nonholonomic dynamics. *Journal of Mathematical Physics*, 46(3):032902–8, 2005.

[53] D. Sosa. Private communication. 2009.

[54] S. V. Stanchenko. Non-holonomic Chaplygin systems. *Journal of Applied Mathematics and Mechanics*, 53(1):11–17, 1989.

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

[56] Y. B. Suris. Discrete Lagrangian models. In *Discrete Integrable Systems*, volume 644 of *Lecture Notes in Physics*, pages 111–184. Springer, 2004.

[57] A. J. van der Schaft. Implicit Hamiltonian systems with symmetry. *Reports on Mathematical Physics*, 41(2):203–221, 1998.

[58] A. J. van der Schaft and B. M. Maschke. On the Hamiltonian formulation of nonholonomic mechanical systems. *Reports on Mathematical Physics*, 34(2):225–233, 1994.

[59] A. M. Vershik and V. Ya. Gershkovich. Nonholonomic problems and the theory of distributions. *Acta Applicandae Mathematicae: An International Survey Journal on Applying Mathematics and Mathematical Applications*, 12(2):181–209, 1988.

[60] H. Yoshimura and J. E. Marsden. Dirac structures in Lagrangian mechanics Part I: Implicit Lagrangian systems. *Journal of Geometry and Physics*, 57(1): 133–156, 2006.

[61] H. Yoshimura and J. E. Marsden. Dirac structures in Lagrangian mechanics Part II: Variational structures. *Journal of Geometry and Physics*, 57(1):209–250, 2006.