Discrete variational integrators and optimal control theory

Manuel DE LEÓN * David MARTÍN DE DIEGO†

Aitor SANTAMARÍA-MERINO‡

Laboratory of Dynamical Systems, Mechanics and Control
Instituto de Matemáticas y Física Fundamental
Consejo Superior de Investigaciones Científicas
Serrano 123, 28006 Madrid, SPAIN

December 10, 2007

Abstract

A geometric derivation of numerical integrators for optimal control problems is proposed. It is based in the classical technique of generating functions adapted to the special features of optimal control problems.

1 Introduction

Optimal control has been one of the driving forces behind many of the applications of mathematics to engineering, robotics, economics... In fact, the Maximum Principle was discovered by L.S. Pontryagin in 1955 in an attempt to find a solution for a highly specific optimization problem related to the manoeuvres of an aircraft. One of its main features is the interplay among different research areas, specially control theory, classical mechanics and differential geometry. Historically, Optimal Control Theory (OCT) took place during the 1950's and its geometrization was started in the 1960's. This geometric analysis of OCT has been introduced using many fundamental tools of differential geometry: Lie groups, exterior differential systems, fiber bundles, riemannian and subriemannian geometry among others.

From other point of view, a geometric methodology has been recently shown to be very useful for simulating numerically the motion of dynamical systems. Following

*mdeleon@imaff.cfmac.csic.es
†d.martin@imaff.cfmac.csic.es
‡aitors@imaff.cfmac.csic.es

this research line, new numerical methods have been developed, called geometric integrators; usually, these integrators, in simulations, can run for longer times with lower spurious effects (for instance, bad energy behavior for conservative systems) than the traditional (non-geometrical) ones. In particular, we are interested in extensions to OCT of Discrete variational integrators. These integrators have precisely their roots in the optimal control literature in the 1960's and 1970's (Jordan and Polack [JorPol:64], Cadzow [Cadz:70], Maeda [Mae:80, Mae:81]) and in 1980's by Lee [Lee:83, Lee:87], Moser and Veselov [MosVes:91]. Although this kind of symplectic integrators have been considered for conservative systems [JarNor:97a, KaMaOr:99, MarWes:01], it has been recently shown how discrete variational mechanics can include forced or dissipative systems [KMOW:00, MarWes:01], holonomic constraints [MarWes:01], time-dependent systems [LeoMdD:2002, MarWes:01], frictional contact [PKMO:02] and nonholonomic constraints (see [Cort:02, CorMar:01, LeMDSa:02a, LeMDSa:02b]). Moreover, it has been also discussed reduction theory [BobSus:99a, BobSus:99b], extension to field theories [JarNor:97b, MaPaSh:98] and quantum mechanics [NorJar:98].

In this paper, we shall continue this work by extending to the discrete variational techniques to Optimal Control Problems and relating our results with Discrete Optimal Control Theory. Mainly, we shall give a geometrical construction of symplectic integrators for OCT, proving as a direct consequence the symplecticity of some discrete optimal control problems. As a nice consequence, an easy proof of the symplecticity of discrete Hamilton equations will be given.

Since most engineering systems are time-dependent, we shall include the time variable explicitely in our control models and some geometric tools (mainly, cosymplectic geometry) of time-dependent mechanics will be useful

2 Optimal control theory

It is well known that the dynamics of a large class of engineering and economic systems can be expressed as a set of differential equations

$$\dot{q}^A = \Gamma^A(t, q(t), u(t)) \,, \; 1 \leq A \leq n \,, \tag{1}$$

where t is the time, q^A denote the state variables and u^a, $1 \leq a \leq m$, the control inputs to the system that must be specified. Given an initial condition of the state variables and given control inputs we completely know the trajectory of the state variables $q(t)$ (all the functions are assumed to be at least C^2).

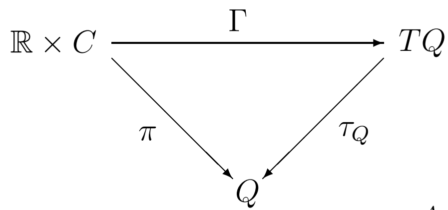
Given an initial condition, usually $q_0 = q(t_0)$, our aim is to find a C^2-piecewise smooth curve $\gamma(t) = (q(t), u(t))$, satisfying the control equations (1) and minimizing the functional

$$\mathcal{J}(\gamma) = \int_{t_0}^{T} L(t, q(t), u(t))\, dt + S(T, q(T)) \,, \tag{2}$$

for some fixed and given final time $T \in \mathbb{R}^+$. The integral $\int_{t_0}^{T} L(t, q(t), u(t))\, dt$ depends on the time history (from t_0 to T) of the state variables and the control inputs, and $S(\cdot, q(\cdot))$ is a cost function based on the final time and the final states of the system.

In a global description, one assumes a fiber bundle structure $\pi : \mathbb{R} \times C \longrightarrow Q$, where Q is the configuration manifold with local coordinates (q^A) and C is the bundle of controls, with coordinates (q^A, u^a), $1 \leq A \leq n$, $1 \leq a \leq m$.

The time-dependent ordinary differential equations (1) on Q depending on the parameters u can be seen as a vector field Γ along the projection map π, that is, Γ is a smooth map $\Gamma : \mathbb{R} \times C \longrightarrow TQ$ such that the diagram



is commutative. This vector field is locally written as $\Gamma = \Gamma^A(t, q, u)\dfrac{\partial}{\partial q^A}$.

A neccesary condition for the solutions of such problem are provided by Pontryaguin's maximum principle. If we construct the pseudo-Hamiltonian function:

$$H(t, q, p, u) = p_A \Gamma^A(t, q, u) - L(t, q, u) = p\Gamma(t, q, u) - L(t, q, u) \tag{3}$$

where p_A, $1 \leq A \leq n$, are now considered as Lagrange's multipliers, then a curve $\gamma : [t_0, T] \to C$, $\gamma(t) = (q(t), u(t))$ is an optimal trajectory if there exist functions $p_A(t)$, $1 \leq A \leq n$, such that they are solutions of the pseudo-Hamilton equations:

$$\left\{ \begin{array}{l} \dot{q}^A(t) = \dfrac{\partial H}{\partial p_A}(t, q(t), p(t), u(t)) \\ \\ \dot{p}_A(t) = -\dfrac{\partial H}{\partial q^A}(t, q(t), p(t), u(t)) \end{array} \right. \tag{4}$$

and we have

$$H(t, q(t), p(t), u(t)) = \min_v H(t, q(t), p(t), v), \quad t \in [t_0, T] \tag{5}$$

with transversality conditions

$$q(0) = q_0 \qquad \text{and} \qquad p_A(T) = -\frac{\partial S}{\partial q^A}(T, x(T))$$

Condition (5) is usually replaced by

$$\frac{\partial H}{\partial u^a} = 0, \quad 1 \leq a \leq m\ , \tag{6}$$

when we are looking for extremal trajectories.

It is well known that the Pontryaguin's necessary conditions for extremality have a geometric interpretation in terms of presymplectic (or precosymplectic) Hamiltonian systems. The total space of the system will be $\mathbb{R} \times (T^*Q \times_Q C)$, with induced coordinates (t, q^A, p_A, u^a) .

Define the Pontryaguin's Hamiltonian function $H : \mathbb{R} \times (T^*Q \times_Q C) \longrightarrow \mathbb{R}$ as follows

$$H(t, \alpha_q, u_q) = \langle \alpha_q, \Gamma(t, u_q) \rangle - L(t, u_q)$$

where $\alpha_q \in T_q^*Q$ and $(t, u_q) \in \pi^{-1}(q)$. Therefore, the coordinate expression of H is (3).

Let $\omega_Q = -d\theta_Q$ be the canonical symplectic form on T^*Q, where θ_Q is the Liouville form, and consider the canonical projection $\pi_1 : \mathbb{R} \times (T^*Q \times_Q C) \longrightarrow T^*Q$. Define the 2-form Ω_H on $\mathbb{R} \times (T^*Q \times_Q C)$ by $\Omega_H = \pi_1^*\omega_Q + dH \wedge dt$. Then, (dt, Ω_H) is a precosymplectic structure on $\mathbb{R} \times (T^*Q \times_Q C)$ (see [LeMaMD:96]).

Eqs. (4) and (6) can be intrinsically written as

$$i_X \Omega_H = 0, \quad i_X dt = 1 \tag{7}$$

Since (dt, Ω_H) is a precosymplectic structure, Eqs. (7) need not have a solution, in general.

Applying the Dirac-Bergmann-Gotay-Nester algorithmn [Dirac:64, GotNes:79] to the precosymplectic system

$$(\mathbb{R} \times (T^*Q \times_Q C), dt, \Omega_H, H)$$

(see [ChLeMa:94]) we obtain that Eqs. (6) correspond to the primary constraints for the precosymplectic system:

$$\phi^a = \frac{\partial H}{\partial u^a} = 0$$

Eqs. (7) have algebraic solution along the first constraint submanifold P_0 determined by the vanishing of the primary constraints. On the points of P_0 there is at least a pointwise solution of Eq. (7), but such solutions are not, in general, tangent to P_0. These points must be removed leaving a subset $P_1 \subset P_0$ (it is assumed than P_1 also is a submanifold). Thus, we have to restrict to a submanifold P_2 where the solutions of (7) are tangent to P_1. Proceeding further this way, we obtain a sequence of submanifolds

$$\cdots \hookrightarrow P_k \hookrightarrow \cdots \hookrightarrow P_2 \hookrightarrow P_1 \hookrightarrow P_0 \hookrightarrow \mathbb{R} \times (T^*Q \times_Q C)$$

If this algorithm stabilizes, i.e. there exists a positive integer $k \in \mathbb{N}$ such that $P_k = P_{k+1}$ and $\dim P_k \neq 0$, then we shall obtain a final submanifold $P_f = P_k$, on which a vector field X exists such that

$$(i_X \Omega_H)_{|P_f} = 0, \quad (i_X dt = 1)_{|P_f} \tag{8}$$

The constraints determining P_f are known, in the control literature, as higher order conditions for optimality.

If X is a solution of (8) then every arbitrary solution on P_f is of the form $X' = X + \xi$, where $\xi \in (\ker \Omega_H \cap \ker dt) \cap TP_f$.

Therefore, a necessary condition for optimality of the curve $\gamma : \mathbb{R} \to \mathbb{R} \times C$, $\gamma(t) = (t, q(t), u(t))$ is the existence of a lift $\tilde{\gamma}$ of γ to P_f such that $\tilde{\gamma}$ is an integral curve of a solution to Eqs. (8).

In the regular case, the final constraint manifold will be P_0 (that is, $P_0 = P_f$) and all the constraints are of the second kind following the classification of Dirac (see [LeMaMD:96]). In such case, (P_0, Ω, η) is a cosymplectic manifold, where Ω and η denote the restrictions of Ω_H and dt to the submanifold P_0. Denote also by ω and θ the restrictions of $\pi_1^*\omega_Q$ and $\pi_1^*\theta_Q$ to P_0.

The cosymplecticity of (P_0, η, Ω) is locally equivalent to the regularity of the matrix

$$\left(\frac{\partial^2 H}{\partial u^a \partial u^b} \right)_{1 \leq a, b \leq m}$$

along P_0. The dynamical equations for the optimal control problem will become

$$i_X \Omega = 0, \quad i_X \eta = 1 \tag{9}$$

Taking coordinates (t, q^A, p_A) on P_0, then (9) are equivalent to:

$$\left\{ \begin{array}{l} \dot{q}^A(t) = \dfrac{\partial H_{|P_0}}{\partial p_A}(t, q(t), p(t)) \\ \\ \dot{p}_A(t) = -\dfrac{\partial H_{|P_0}}{\partial q^A}(t, q(t), p(t)) \; , \end{array} \right. \tag{10}$$

where we have substituted in (4) the control variables u^a by its value $\bar{u}^a = f^a(t, q, p)$, applying the Implicit Function Theorem to the primary constraints $\phi^a = 0$. This also implies that we have a canonical projection from P_0 onto $\mathbb{R}$, say $\pi_0 : P_0 \to \mathbb{R}$.

In such case, there exists a unique solution X_{P_0} of Eq. (9):

$$i_{X_{P_0}}\Omega = 0, \quad i_{X_{P_0}}\eta = 1$$

and its flow preserves the cosymplectic structure given by Ω and η. That is, if we denote by F_h the flow of X_{P_0} then $F_h^*\Omega = \Omega$ and $F_h^*\eta = \eta$. In local coordinates, $F_h(t_0, q_0, p_0) = (t_0 + h, q_1, p_1)$. Denote by $F_h^{(2)}$ the mapping $F_h^{(2)}(t_0, q_0, p_0) = (q_1, p_1)$, and by $F_{t_1,t_0} : P_0^{t_0} \longrightarrow P_0^{t_1}$ the mapping defined by

$$F_{t_1,t_0}(q_0, p_0) = F^{(2)}_{t_1 - t_0}(t_0, q_0, p_0) \; ,$$

where we write $P_0^t = (\pi_0)^{-1}(t)$, with $t \in \mathbb{R}$. Obviously, $F_{t_2,t_1} \circ F_{t_1,t_0} = F_{t_2,t_0}$ in their common domain.

The submanifolds P_0^t naturally inherit a symplectic structure ω_t by taking the restriction of ω to P_0^t. Similarly, denote by θ_t the restriction of θ to P_0^t, then $\omega_t = -d\theta_t$.

It is easy to deduce that, in such case, F_{t_1,t_0} is a symplectomorphism; that is, $F^*_{t_1,t_0}\omega_{t_1} = \omega_{t_0}$, noting that

$$\Omega = \omega + dH_{|P_0} \wedge \eta$$

This last remark will be interesting for constructing geometrical integrators for explicitly time-dependent optimal control systems.

3 Generating functions

Let (M_i, ω_i), $i = 0, 1$ be two exact symplectic manifolds (i.e. ω_i is symplectic and exact, $\omega_i = -d\theta_i$, $i = 0, 1$) and suppose that $g : M_0 \to M_1$ is a diffeomorphism. Denote by Graph(g) the graph of g, $\text{Graph}(g) = \{(x_0, g(x_0)) \; / \; x_0 \in M_0\} \subset M_0 \times M_1$. Denote by $\pi_i : M_0 \times M_1 \to M_i$, $i = 0, 1$ the canonical projections, and consider the 1-form and 2-form on $M_0 \times M_1$ defined by

$$\begin{aligned} \Theta_{(1,0)} &= \pi_1^*\theta_1 - \pi_0^*\theta_0 \\ \Omega_{(1,0)} &= \pi_1^*\omega_1 - \pi_0^*\omega_0 = -d\Theta_{(1,0)} \end{aligned}$$

As it is well known $\Omega_{(1,0)}$ is a symplectic form.

Let $i_g : \text{Graph}(g) \hookrightarrow M_0 \times M_1$ be the inclusion map, then

$$i_g^*\Omega_{(1,0)} = (\pi_{0|\text{Graph}(g)})^*(g^*\omega_1 - \omega_0)$$

Using this equality, it is clear that g is a symplectomorphism if and only if $i_g^*\Omega_{(1,0)} = 0$, that is, if Graph(g) is a Lagrangian submanifold of $(M_0 \times M_1, \Omega_{(1,0)})$.

Now, if g is a symplectomorphism we have

$$i_g^*\Omega_{(1,0)} = -di_g^*\Theta_{(1,0)} = 0$$

and, therefore, at least locally, there exists a function $S : \mathrm{Graph}\,(g) \to \mathbb{R}$ such that

$$i_g^*\Theta_{(1,0)} = dS \tag{11}$$

Let (q_0, p_0) and (q_1, p_1) Darboux coordinates in M_0 and M_1, respectively. Since Graph(g) is diffeomorphic to M_0, we can take (q_0, p_0) as natural coordinates in Graph(g). Since (q_0, p_0, q_1, p_1) are coordinates in $M_0 \times M_1$, then, along Graph(g), we have $q_1 = q_1(q_0, p_0)$, $p_1 = p_1(q_0, p_0)$ and

$$p_1\, dq_1 - p_0 dq_0 = dS(q_0, p_0)$$

3.1 Generating functions of the first kind

Assume that in a neighborhood of some point $x \in$ Graph(g), we can change this system of coordinates by new independent coordinates (q_0, q_1) (the local condition is that $\det(\partial q_1/\partial p_0) \neq 0$). In such case, the function S can be expressed locally as $S = S(q_0, p_0) = S_1(q_0, q_1)$.

Definition 3.1 *The function $S_1(q_0, q_1)$ will be called a* **generating function of the first kind** *of the symplectomorphism g.*

From (11) we deduce that

$$\begin{cases} p_0 = -\dfrac{\partial S_1}{\partial q_0} \\ p_1 = \dfrac{\partial S_1}{\partial q_1} \end{cases} \tag{12}$$

(see Fig. 1).

Conversely, if $S_1(q_0, q_1)$ is a function such that $\det\left(\frac{\partial^2 S_1}{\partial q_0 \partial q_1}\right) \neq 0$ then $S_1(q_0, q_1) = (p_0, p_1)$ is a generating function of some canonical transformation g implicitly determined by Eqs. (12), $g(q_0, p_0) = (q_1, p_1)$ (see [Arn:78]).

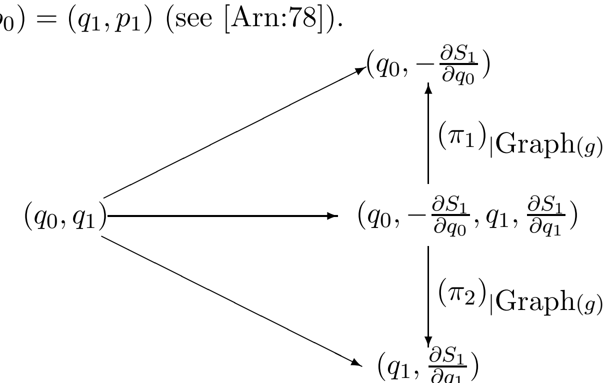


Fig.1

Now suppose that M is a fiber bundle over the real line $\mathbb{R}$, $\pi : M \to \mathbb{R}$, and $M_t = \pi^{-1}(t)$ are the fibers, where each fiber M_t is equipped with a symplectic form ω_t. Let $g_{(s,t)} : M_t \to M_s$ be a two-parameter family of symplectomorphisms satisfying

$$g_{(t_2,t_1)} \circ g_{(t_1,t_0)} = g_{(t_2,t_0)}$$

Next, we shall show how this composition law can be translated in terms of their respective generating functions. Moreover, the following results will give a geometric interpretation of the Discrete Euler-Lagrange equations [MarWes:01].

Theorem 3.2 *Let $S_1^{(t_N,t_0)}$ be a function defined by*

$$S_1^{(t_N,t_0)}(q_0, q_N) = \sum_{k=0}^{N-1} S_1^{(t_{k+1},t_k)}(q_k, q_{k+1})$$

where $q_k \in M_{t_k}$, $1 \le k \le N-1$, are stationary points of the right-hand side, that is

$$0 \;=\; D_2 S_1^{(t_k,t_{k-1})}(q_{k-1}, q_k) + D_1 S_1^{(t_{k+1},t_k)}(q_k, q_{k+1}), \quad 1 \le k \le N-1.$$

If $S_1^{(t_k,t_{k-1})}$ are generating functions of the first kind for $g_{(t_k,t_{k-1})}$, then $S_1^{t_N,t_0)}$ is a generating function of the first kind for $g_{(t_N,t_0)} : M_{t_0} \to M_{t_N}$.

Proof: Recursively, it is suffices to give the proof for $N = 2$:

$$S_1^{(t_2,t_0)}(q_0, q_2) = S_1^{(t_1,t_0)}(q_0, x) + S_1^{(t_2,t_1)}(x, q_2)$$

where x is an stationary point of the right-hand side.

From the definitions of generating functions for $g_{(t_2,t_1)}$ and $g_{(t_1,t_0)}$

$$\begin{aligned} p_1\, dq_1 - p_0\, dq_0 &= dS^{(t_1,t_0)}(q_0, q_1) \\ p_2\, dq_2 - p_1\, dq_1 &= dS_1^{(t_2,t_1)}(q_1, q_2) \end{aligned}$$

and therefore

$$p_2\, dq_2 - p_0\, dq_0 = d(S_1^{(t_2,t_1)}(q_0, q_1) + S_1^{(t_1,t_0)}(q_1, q_2))$$

It follows that

$$0 = D_2 S_1^{(t_1,t_0)}(q_0, q_1) + D_1 S_1^{(t_2,t_1)}(q_1, q_2)$$

and, obviously, for this choice of q_1 then

$$S_1^h(q_0, q_1) + S_1^h(q_1, q_2)$$

is a generating function of the first kind of $g_{(t_2,t_0)}$. ∎

Now, we are in condition to bring this procedure to the limit when the number of subintervals increases to infinity. Consider as its continuous counterpart a cosymplectic manifold (M, η, ω), where M is still a fiber bundle over $\mathbb{R}$ ($\pi_{\mathbb{R}} : M \to \mathbb{R}$) and $\eta =$

$\pi_{\mathbb{R}}^*(dt)$. Denote by $M_t = \pi_{\mathbb{R}}^{-1}(t)$, $t \in \mathbb{R}$. Take a Hamiltonian function $H : M \to \mathbb{R}$ and its Hamiltonian vector field X_H given by

$$i_{X_H}\omega = 0 \qquad \text{and} \qquad i_{X_H}\eta = 1$$

Let $F_{(t,s)} : M_s \to M_t$ be the two-parameter family of symplectomorphisms generated by X_H (see section 2) and consider as symplectic form on each fiber M_t the restriction of ω to this fiber.

We shall give a characterization of the generating functions of the first kind associated to $F_{(t,s)}$ for t close enough to s. For doing that, consider Darboux coordinates (t, q^A, p_A) on M and assume the regularity condition $\det\left(\frac{\partial^2 H}{\partial p_A \partial p_B}\right) \neq 0$. Thus,

Proposition 3.3 *A generating function of the first kind for $F_{(t,s)}$ is given by*

$$S_1^{(t_1,t_0)}(q_0, q_1) = \int_{t_0}^{t_1} \left(p(t)\dot{q}(t) - H(t, q(t), p(t))\right)\, dt$$

where $t \to (t, q(t), p(t))$ is an integral curve of the Hamilton equations such that $q(t_0) = q_0$ and $q(t_1) = q_1$.

Proof: We only use Hamilton equations and integration by parts:

$$\begin{aligned}
\frac{\partial S_1^{(t_1,t_0)}}{\partial q_0}(q_0, q_1) &= \int_{t_0}^{t_1} \left(\frac{\partial p}{\partial q_0}\dot{q} + p\frac{\partial \dot{q}}{\partial q_0} - \frac{\partial H}{\partial q}\frac{\partial q}{\partial q_0} - \frac{\partial H}{\partial p}\frac{\partial p}{\partial q_0}\right) dt \\
&= \int_{t_0}^{t_1} \left(p\frac{\partial \dot{q}}{\partial q_0} + \dot{p}\frac{\partial q}{\partial q_0}\right) dt \\
&= -p_0 + p_1\frac{\partial q_1}{\partial q_0} = -p_0
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial S_1^{(t_1,t_0)}}{\partial q_1}(q_0, q_1) &= \int_{t_0}^{t_1} \left(\frac{\partial p}{\partial q_1}\dot{q} + p\frac{\partial \dot{q}}{\partial q_1} - \frac{\partial H}{\partial q}\frac{\partial q}{\partial q_1} - \frac{\partial H}{\partial p}\frac{\partial p}{\partial q_1}\right) dt \\
&= \int_{t_0}^{t_1} \left(p\frac{\partial q}{\partial q_1} + \dot{p}\frac{\partial q}{\partial q_1}\right) dt \\
&= p_1 - p_0\frac{\partial q_0}{\partial q_1} = p_1 \quad \blacksquare
\end{aligned}$$

Remark 3.4 Suppose that $t_{i+1} - t_i = h$, for all $i = 0, \cdots N-1$, then from Theorem 3.2 we have

$$S_1^{Nh}(q_0, q_N) = \sum_{k=0}^{N-1} S_1^h(q_k, q_{k+1})$$

where

$$0 = D_2 S_1^h(q_{k-1}, q_k) + D_1 S_1^h(q_k, q_{k+1}), \quad 1 \leq k \leq N-1.$$

Now, if we take as new generating function an adequate approximation S_d^h of S_1^h then

$$0 = D_2 S_d^h(q_{k-1}, q_k) + D_1 S_d^h(q_k, q_{k+1}), \quad 1 \leq k \leq N-1.$$

are the well-known Discrete Euler-Lagrange equations (see [MarWes:01] and references therein). For instance, one can take

$$S_d^h(q_0, q_1) = h\mathcal{L}(\alpha q_0 + (1-\alpha)q_1, \frac{q_1 - q_0}{h}), \quad \alpha \in [0,1]$$

or alternatively, we could have considered more accurate approximations. Here, we are assuming that $\mathcal{L} : \mathbb{R} \times TQ \to \mathbb{R}$ is a Lagrangian function related via Legendre transformation with the Hamiltonian function H (see [Arn:78]) which is locally possible because of the regularity of H.

Denote by $S_1(q_0, q_1, t_0, t_1) = S_1^{(t_1,t_0)}(q_0, q_1)$. From Proposition (3.3), it is easy to show that:

$$\begin{aligned} D_3 S_1(q_0, q_1, t_0, t_1) = D_3 S^{(t_1,t_0)}(q_0, q_1) &= H(t_0, q_0, p_0) \\ D_4 S_1(q_0, q_1, t_0, t_1) = D_4 S^{(t_1,t_0)}(q_0, q_1) &= -H(t_1, q_1, p_1) \end{aligned}$$

(see also [MarWes:01]). As a consequence

$$D_4 S^{(t_k,t_{k-1})}(q_{k-1}, q_k) + D_3 S^{(t_{k+1},t_k)}(q_k, q_{k+1}) = 0 \tag{13}$$

It should be noticed that if we take a new function $S_d^{(t_{k+1},t_k)}$ as an adequate approximation of $S^{(t_{k+1},t_k)}$, then solutions $\{q_0, q_1, \ldots, q_N\}$ of equations

$$D_2 S_d^{(t_k,t_{k-1})}(q_{k-1}, q_k) + D_1 S_d^{(t_{k+1},t_k)}(q_k, q_{k+1}) = 0, \quad 1 \leq k \leq N-1.$$

do not satisfy (13) for arbitrary values of t_{k-1}, t_k, t_{k+1}. Therefore, we may write the system of difference equations

$$\left\{ \begin{array}{l} D_2 S_d^{(t_k,t_{k-1})}(q_{k-1}, q_k) + D_1 S_d^{(t_{k+1},t_k)}(q_k, q_{k+1}) = 0, \\ D_4 S_d^{(t_k,t_{k-1})}(q_{k-1}, q_k) + D_3 S_d^{(t_{k+1},t_k)}(q_k, q_{k+1}) = 0, \end{array} \right. \tag{14}$$

which under regularity assumptions will determine a time-dependent discrete flow

$$\Phi(q_{k-1}, q_k, t_{k-1}, t_k) = (q_k, q_{k+1}, t_k, t_{k+1})$$

with variable step size $h_k = t_{k+1} - t_k$ (see [KaMaOr:99, Lee:83, Lee:87, LeoMdD:2002, MarWes:01]).

3.2 Generating functions of the second kind

The construction of more general generating functions will be useful in next sections. For instance, suppose that (q_0, p_1) are independent local coordinates on Graph(g). Then the function S is written as $S = S(q_0, p_1)$.

We have

$$p_1\, dq_1 - p_0\, dq_0 = -q_1\, dp_1 + d(q_1 p_1) - p_0\, dq_0 = dS.$$

If we define

$$S_2(q_0, p_1) = q_1 p_1 - S(q_0, p_1),$$

where q_1 is expressed in terms of q_0 and p_1, then we deduce that

$$q_1\, dp_1 + p_0 dq_0 = dS_2(q_0, p_1)$$

Definition 3.5 *The function* $S_2(q_0, p_1)$ *will be called a* **generating function of the second kind** *of the symplectomorphism* g.

We have that

$$\begin{cases} p_0 = \dfrac{\partial S_2}{\partial q_0} \\ q_1 = \dfrac{\partial S_2}{\partial p_1} \end{cases} \tag{15}$$

Conversely, if $S_2(q_0, p_1)$ is a generating function such that $\det\left(\dfrac{\partial^2 S_2}{\partial q_0 \partial p_1}\right) \neq 0$ then S_2 is a generating function of some local symplectomorphism determined by Eqs. (15) (see [Arn:78]).

Denote by $F_{(t_2,t_1)} : M_{t_1} \to M_{t_2}$ the two-parametric group of canonical transformations generated by the Hamiltonian vector field X_H, as in the preliminaries to Proposition 3.3. We have the following.

Theorem 3.6 *Let a function* $S_2^{(t_N,t_0)}$ *be defined by*

$$S_2^{(t_N,t_0)}(q_0, p_N) = \sum_{k=0}^{N-1} S_2^{(t_{k+1},t_k)}(q_k, p_{k+1}) - \sum_{k=1}^{N-1} q_k p_k \tag{16}$$

where q_k, $1 \leq k \leq N$, *and* p_k, $0 \leq k \leq N-1$, *are stationary points of the right-hand side, that is*

$$q_k = \frac{\partial S_2^{(t_{k-1},t_k)}}{\partial p}(q_{k-1}, p_k), \quad 1 \leq k \leq N, \tag{17}$$

$$p_k = \frac{\partial S_2^{(t_k,t_{k+1})}}{\partial q}(q_k, p_{k+1}), \quad 0 \leq k \leq N-1, \tag{18}$$

then $S_2^{(t_N,t_0)}$ *is a generating function of the second kind for* $F_{(t_N,t_0)} : M_{t_0} \to M_{t_N}$.

Proof: It follows as in Theorem 3.2. ∎

As a consequence, we have that

$$S^{(t_N,t_0)}(q_0, p_N) = q_N p_N - S_2^{(t_N,t_0)}(q_0, p_N) = \sum_{k=0}^{N-1} \left[q_{k+1} p_{k+1} - S_2^{(t_{k+1},t_k)}(q_k, p_{k+1}) \right], \tag{19}$$

where the unknown coordinates are given by (17) and (18).

Proposition 3.7 *A generating function of the second kind for* $F_{(t_1,t_0)}$ *is given by*

$$S_2^{(t_1,t_0)}(q_0, p_1) = p_1 q_1 - \int_{t_0}^{t_1} \left(p(t)\dot{q}(t) - H(t, q(t), p(t))\right) dt$$

where $t \to (q(t), p(t))$ *is an integral curve of the Hamilton equations such that* $q(t_0) = q_0$ *and* $p(t_1) = p_1$.

Proof: It is proved in a similar way to Proposition 3.3. ▮

Denote by $S_2(t, q_0, p_1) = S_2^{(0,t)}(q_0, p_1)$ then it is easy to show that (see, for instance [HaLuWa:02])

Theorem 3.8 (Hamilton-Jacobi equation for S_2) *If $S_2(t, q_0, p_1)$ is a solution of the partial differential equation*

$$\frac{\partial S_2}{\partial t} = H(\frac{\partial S_2}{\partial p_1}(t, q_0, p_1), p_1), \qquad S_2(0, q_0, p_1) = q_0 p_1 \tag{20}$$

then the mapping $(q_0, p_0) \longrightarrow (q_1, p_1)$ defined by Eqs. (15) is the exact flow of the Hamiltonian system determined by H.

4 Optimal control of Discrete-time systems

In this section we shall define the general solution of an optimization problem for discrete systems and analyze its geometric behaviour, in particular, the symplecticity.

Suppose that the discrete state equations are given by the dynamical equation

$$q_{k+1}^A = f^A(k, q_k, u_k), \quad k = 0, 1, \dots, N-1, \quad A = 1, 2, \dots, m \tag{21}$$

or, shortly, $q_{k+1} = f(k, q_k, u_k)$, where q_0 is initially given.

The associate performance index or objective function is:

$$J = \bar{S}(N, q(N)) + \sum_{k=0}^{N-1} \bar{L}(k, q_k, u_k) \tag{22}$$

where $\bar{S}$ is a function of the final time and state at the final time N, and $\bar{L}$ is time-varying function of the state and control input at each intermediate discrete time k.

The optimal control problem is solved finding controls u_k^*, $k = 0, 1, \dots N-1$, that drive the system along a trajectory q_k^*, $k = 0, 1, \dots, N$, verifying the state equations such that the performance index is minimized.

4.1 Problem solution

Let us now solve the optimal control problem for the discrete optimal problem determined by (21) and (22) using the Lagrange multiplier approach. Considering the state Eqs. (21) as constraint equations, then we have $N \cdot m$ constraints, and we associate a Lagrange multiplier to each constraint. Next, we construct the augmented performance index J' by

$$J' = \sum_{k=0}^{N-1} \left[p_{k+1}(f(k, q_k, u_k) - q_{k+1}) - \bar{L}(k, q_k, u_k) \right] - \bar{S}(N, q(N)) \tag{23}$$

where $p_{k+1} = ((p_{k+1})_A)$ are considered as Lagrange multipliers with $A = 1, \dots, n$ and $k = 0, \dots, N-1$.

Taking the Hamiltonian function

$$\bar{H}(k, q_k, p_{k+1}, u_k) = p_{k+1} f(k, q_k, u_k) - \bar{L}(k, q_k, u_k)$$

we deduce that the necessary conditions for a constrained minimum are thus given by:

$$
\begin{aligned}
q_{k+1} &= \frac{\partial \bar{H}}{\partial p}(k, q_k, p_{k+1}, u_k) = f(k, q_k, u_k) && (24)\\
p_k &= \frac{\partial \bar{H}}{\partial q}(k, q_k, p_{k+1}, u_k) = p_{k+1}\frac{\partial f}{\partial q}(k, q_k, u_k) - \frac{\partial \bar{L}}{\partial q}(k, q_k, u_k) && (25)\\
0 &= \frac{\partial \bar{H}}{\partial u}(k, q_k, p_{k+1}, u_k) = p_{k+1}\frac{\partial f}{\partial u}(k, q_k, u_k) - \frac{\partial \bar{L}}{\partial u}(k, q_k, u_k) && (26)
\end{aligned}
$$

where $0 \leq k \leq N-1$, and the transversality conditions

$$
p_N = -\frac{\partial \bar{S}}{\partial q}(N, q_N) \quad \text{and} \quad q_0 \quad \text{fixed.}
$$

Observe that the recursion for the state q_k develops forward in time, but the co-state variable p_k develops backwards in time. Therefore the required boundary conditions for finding a solution are the initial state q_0 and the final co-state p_N.

Assume that

$$
\det\left(\frac{\partial^2 \bar{H}}{\partial u_a \partial u_b}\right) \neq 0
$$

then, locally, $u_k^* = h(k, q_k, p_{k+1})$. If we denote, by

$$
\tilde{H}(k, q_k, p_{k+1}) = \bar{H}(k, q_k, p_{k+1}, u_k^*)
$$

then Eqs. (24), (25) are rewritten as

$$
\begin{aligned}
q_{k+1} &= \frac{\partial \tilde{H}}{\partial p}(k, q_k, p_{k+1}) && (27)\\
p_k &= \frac{\partial \tilde{H}}{\partial q}(k, q_k, p_{k+1}) && (28)
\end{aligned}
$$

with $0 \leq k \leq N_1$.

Consider the function

$$
G_k(q_k, q_{k+1}, p_{k+1}) = \tilde{H}(k, q_k, p_{k+1}) - p_{k+1}q_{k+1}, \quad 0 \leq k \leq N-1.
$$

Then, for a fixed k:

$$
dG_k = \frac{\partial \bar{H}}{\partial q_k}(k, q_k, p_{k+1})\, dq_k + \frac{\partial \bar{H}}{\partial p_{k+1}}(k, q_k, p_{k+1})\, dp_{k+1} - p_{k+1}\, dq_{k+1} - q_{k+1}\, dp_{k+1}\ .
$$

Along solutions of Eqs. (24), (25) and (26) we have:

$$
dG_{k|\mathrm{Sol}} = p_k\, dq_k - p_{k+1}\, dq_{k+1}\ ,
$$

which implies

$$
dp_k \wedge dq_k = dp_{k+1} \wedge dq_{k+1}\ . \tag{29}
$$

along the solution of (24)-(26).

In the next subsection, we shall analyze the geometric meaning of Eq. (29), which it is obviously interpreted as symplecticity of discrete optimal control problems in terms of a natural symplectic form.

4.2 Generating functions of the second kind and discrete optimal control problems

From Proposition 3.3 the following function is a generating function of the second kind for the cosymplectic Hamiltonian system $(P_0, \eta, \Omega, H_{|P_0})$, which determines the dynamics of the optimal control problem given by (1) and (2):

$$S_2^{(t_1,t_0)}(q_0, p_1) = p_1 q_1 - \int_{t_0}^{t_1} \left(p(t)\dot{q}(t) - H_{|P_0}(t, q(t), p(t))\right) \, dt \, , \tag{30}$$

where $t \to (t, q(t), p(t))$ is the integral curve on P_0 of the vector field X_{P_0}. Here X_{P_0} is the unique solution of equation

$$i_{X_{P_0}}\Omega = dH_{|P_0} \qquad i_{X_{P_0}}\eta = 1$$

with $(q(t_0), p(t_0)) = (q_0, p_0)$ and $(q(t_1), p(t_1)) = (q_1, p_1)$.

We now focus on the construction of a numerical integrator for the Hamiltonian system $(P_0, \eta, \Omega, H_{|P_0})$ by using an approximation of the generating function. As we shall show, the obtained method also realize the integration steps by symplectomorphism transformations; then, it is a symplectic integrator.

First take a fixed time interval $h = t_{k+1} - t_k$, $k = 0, \ldots, N-1$.

Assume that we are working on vector spaces, and consider the following natural approximation:

$$\begin{aligned} \tilde{S}_2^h(k, q_k, p_{k+1}) &= p_{k+1}q_{k+1} - hp_{k+1}\left(\frac{q_{k+1} - q_k}{h}\right) - h\tilde{L}(k, q_k, p_{k+1}) \\ &\quad + hp_{k+1}\tilde{\Gamma}(k, q_k, p_{k+1}) \end{aligned}$$

where, for instance, $\tilde{L}(k, q_k, p_{k+1}) = L_{|P_0}(t_0+kh, q_k, p_{k+1})$ and $\tilde{\Gamma}(k, q_k, p_{k+1}) = \Gamma_{|P_0}(t_0+kh, q_k, p_{k+1})$.

If we denote by $\tilde{f}(k, q_k, p_{k+1})$ the function

$$\tilde{f}(k, q_k, p_{k+1}) = h\tilde{\Gamma}(k, q_k, p_{k+1}) + q_k \tag{31}$$

then,

$$\tilde{S}_2^h(k, q_k, p_{k+1}) = p_{k+1}\tilde{f}(k, q_k, p_{k+1}) - \tilde{L}(k, q_k, p_{k+1}) = \tilde{H}(k, q_k, p_{k+1}) \, .$$

Thus, equations

$$\left\{ \begin{aligned} p_k &= \frac{\partial \tilde{S}_2^h}{\partial q^k}(k, q_k, p_{k+1}) = \frac{\partial \tilde{H}}{\partial q^k}(k, q_k, p_{k+1}) \\ q_{k+1} &= \frac{\partial \tilde{S}_2^h}{\partial p_{k+1}}(k, q_k, p_{k+1}) = \frac{\partial \tilde{H}}{\partial p_{k+1}}(k, q_k, p_{k+1}) \end{aligned} \right. \tag{32}$$

are exactly (27) and (28) and the symplecticity condition (29) for discrete optimal control problems is now a trivial consequence of the generating function construction.

Remark 4.1 It is also possible to construct symplectic numerical methods of higher order; for instance, considering better approximations of the Hamilton Jacobi equation

(20) (see [ChaSco:90] and references therein). Assume for simplicity that the Hamiltonian is autonomous, that is, $H \equiv H(q,p)$. Now, first expands the generating function $S_2(t, q_0, p_1)$ as:

$$S_2(t, q_0, p_1) = q_0 p_1 + \sum_{i=1}^{\infty} t^i G_i(q_0, p_1),$$

inserts expression into Hamilton-Jacobi equation (20) and compares equal powers of t. This yields

$$\begin{aligned}
G_1(q_0,p_1) &= H(q_0,p_1)\\
G_2(q_0,p_1) &= \frac{1}{2}\left(\frac{\partial H}{\partial q_0^A}\frac{\partial H}{\partial p_{1A}}\right)\\
G_3(q_0,p_1) &= \frac{1}{6}\left(\frac{\partial^2 H}{\partial p_{1A}\partial p_{1B}}\frac{\partial H}{\partial q_0^A}\frac{\partial H}{\partial q_0^B} + \frac{\partial^2 H}{\partial p_{1A}\partial q_0^B}\frac{\partial H}{\partial q_0^A}\frac{\partial H}{\partial p_{1B}} + \frac{\partial^2 H}{\partial q_0^A\partial q_0^B}\frac{\partial H}{\partial p_{1A}}\frac{\partial H}{\partial p_{1B}}\right)\\
\cdots &= \cdots
\end{aligned}$$

Using the truncated series, we obtain an approximated generating function:

$$S_2^h(q_k, p_{k+1}) = q_k \cdot p_{k+1} + \sum_{i=1}^{r} h^r G_i(q_k, p_{k+1})$$

which defines a symplectic method of order r.

Other approaches are also admissible without using higher derivatives of the Hamiltonian H, for instance, symplectic or symplectic partitioned Runge-Kutta methods (see [HaLuWa:02, SanCal:94]).

5 Discrete Hamiltonian systems

In [ErbYan:92] Erbe and Yan have considered discrete linear Hamiltonian systems of the form:

$$\begin{aligned}
\Delta y(t) &= B(t)y(t+1) + C(t)z(t)\\
\Delta z(t) &= -A(t)y(t+1) - B^T(t)z(t)
\end{aligned}$$

where A, C are symmetric and $I - B$ is invertible. Here $\Delta y(t) = y(t+1) - y(t)$, $\Delta z(t) = z(t+1) - z(t)$ and $y, z \in \mathbb{R}^d$.

This problem is a particular case of a discrete Hamiltonian systems of the form

$$\Delta y(t) = H_z(t, y(t+1), z(t)) \qquad (33)$$
$$\Delta z(t) = -H_y(t, y(t+1), z(t)) \qquad (34)$$

where $H(t,y,z) = \frac{1}{2}(y^T, z^T)\begin{pmatrix} A(t) & B^T(t) \\ -B(t) & C(t) \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix}$. The symplecticity of the discrete linear Hamiltonian system was fully studied (see [ErbYan:92], for instance, and references therein). The existence of a corresponding symplectic structure for discrete nonlinear Hamiltonian systems given by (33) and (34) was proposed by Ahlbrandt as an open problem ([Ahlb:93] and also [Shi:02]).

From the point of view of section 3, this open problem is easily solved considering as generating function of the second kind the following one:

$$S_2^{(t+1,t)}(y(t+1),z(t)) = z(t)y(t+1) - H(t,y(t+1),z(t)) \ .$$

Then Eqs. (33) and (34) are precisely

$$\left\{ \begin{array}{rcl} y(t) & = & \frac{\partial S^{(t+1,t)}}{\partial z}(y(t+1),z(t)) \\ z(t+1) & = & \frac{\partial S^{(t+1,t)}}{\partial y}(y(t+1),z(t)) \ , \end{array} \right.$$

which guarantees the symplecticity of the discrete Hamiltonian system. In order to find the canonical transformation associated to this generating map it is only necessary to impose the local condition (see [Arn:78]):

$$\det\left(\frac{\partial^2 S_2^{(t+1,t)}(y(t+1),z(t))}{\partial y \partial z}\right) \neq 0$$

Then, in a neighbourhood of a point satisfying the above condition, there exists a symplectomorphism defined by Eqs. (33) and (34).

Acknowledgments

This work has been supported by grant BFM2001-2272 (Ministry of Science and Technology, Spain). A. Santamaría Merino wishes to thank the Programa de Formación de Investigadores of the Departamento de Educación, Universidades e Investigación of the Basque Government (Spain) for financial support.

References

[Ahlb:93] Ahlbrandt C D 1993 Equivalence of Discrete Euler Equations and Discrete Hamiltonian Systems, *J. Math. Anal. Appl.* **180**, 498-478

[Arn:78] Arnold V I 1978 *Mathematical Methods of Classical Mechanics* (Graduate Text in Mathematics 60, Springer-Verlag New York)

[BobSus:99a] Bobenko A I and Y B Suris 1999 Discrete Lagrangian reduction, discrete Euler-Poincaré equations, and semidirect products *Lett. Math. Phys.* **49**, 79-93

[BobSus:99b] Bobenko A I and Suris Y B 1999 Discrete time Lagrangian mechanics on Lie groups, with an application to the Lagrange top *Comm. Math. Phys.* **204** 147-188

[Cadz:70] Cadzow J A 1970 Discrete calculus of variations *Intern. J. Control.* **11**, 393-407

[ChaSco:90] Channell P J and Scovel C 1990 Symplectic integration of Hamiltonian Systems, *Nonlinearity* **3**, 231-259

[ChLeMa:94] Chinea D, de León M and Marrero J C 1994 The constraint algorithm for time-dependent Lagrangians, *J. Math. Phys.* **35** (7), 3410-3447

[Cort:02] Cortés J 2002 *Geometric, control and numerical aspects of nonholonomic systems* (Lecture Notes in Mathematics, vol. 1793, Springer-Verlag)

[CorMar:01] Cortés J and Martínez S 2001 Nonholonomic integrators *Nonlinearity* **14**, 1365-1392

[Dirac:64] Dirac P A M 1964 *Lecture on Quantum Mechanics* (Belfer Graduate School of Science, Yeshiva University, New York)

[ErbYan:92] Erbe L H and Yan P 1992 Disconjugancy for linear Hamiltonian difference systems, *J. Math. Anal. Appl.* **167**, 355-367

[GotNes:79] Gotay M J and Nester J M 1979 Presymplectic Lagrangian systems I: The Constraint Algorithm and the Equivalence Theorem, *Ann. Inst. Henri Poincaré*, **A30** 129-142

[HaLuWa:02] Hairer E Lubich C and Wanner G 2002 *Geometric Numerical Integration, Structure-Preserving Algorithms for Ordinary Differential Equations* (Springer Series in Computational Mathematics **31**, Springer-Verlag Berlin Heidelberg)

[JarNor:97a] Jaroszkiewicz G and Norton K 1997 Principes of discrete time mechanics I: Particle systems *J. Phys. A* **30**, 3115-3144

[JarNor:97b] Jaroszkiewicz G and Norton K 1997 Principes of discrete time mechanics II: Classical field theory *J. Phys. A* **30**, 3145-3163

[JorPol:64] Jordan B W and Polak E 1964 Theory of a class of discrete optimal control systems *J. Electron. Control* **17**, 697-711

[KaMaOr:99] Kane C Marsden J E and Ortiz M 1999 Symplectic energy-momentum integrators *J. Math. Phys.* **40**, 3353-3371

[KMOW:00] Kane C Marsden J E Ortiz M and West M 2000 Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems *Internat. J. Numer. Math. Eng.* **49**, 1295-1325.

[Lee:83] Lee T D 1983 Can time be a discrete dynamical variable? *Phys. Lett.*, **122B**, 217-220

[Lee:87] Lee T D 1987 Difference equations and conservation laws *J. Statis. Phys.*, **46**, 843-860

[LeMaMD:96] de León M Marrero J C and Martín de Diego D 1996 Time-dependent constrained Hamiltonian systems and Dirac brackets *Journal of Physics A: Math. Gen.* **29** 6843-6859

[LeMDSa:02a] de León M Martín de Diego D and Santamaría A 2002 Geometric integrators and nonholonomic mechanics, *Preprint IMAFF-CSIC*

[LeMDSa:02b] de León M, Martín de Diego D and Santamaría A 2003 Geometric numerical integration of nonholonomic systems and optimal control problems *2nd IFAC Workshop on lagrangian and Hamiltonian Methods for Nonlinear Control, Seville* 2003, 163-168.

[LeoMdD:2002] de León M and Martín de Diego D 2002 Variational integrators and time-dependent Lagrangian systems *Rep. on Math. Phys* **49** 2/3, 183-192

[Lew:86] Lewis F.L 1986 *Optimal Control* (John Wiley& Sons, New York)

[Mae:80] Maeda S 1980 Canonical structure and symmetries for discrete systems *Math. Japonica* **25**, 405-420

[Mae:81] Maeda S 1981 Extension of discrete Noether theorem *Math. Japonica* **26**, 85-90

[MaPaSh:98] Marsden J E Patrick G W and Shkoller S 1998 Multisymplectic geometry, variational integrators, and nonlinear PDEs' *Comm. Math. Phys.* **199**, 351- 395

[MarWes:01] Marsden J E and West M 2001 Discrete mechanics and variational integrators *Acta Numerica* , 357-514

[MosVes:91] Moser J and Veselov A P 1991 Discrete versions of some classical integrable systems and factorization of matrix polynomials *Comm. Math. Phys.* **139**, 217-243

[Nijvan:90] Nijmeijer H and van der Schaft A J 1990 *Nonlinear dynamical control systems* (Springer-Verlag, New York)

[NorJar:98] Norton K and Jaroszkiewicz G 1998 Principes of discrete time mechanics, III: Quantum field theory *J. Phys. A* **31**, 977-1000

[PKMO:02] Pandolfi A Kane C Marsden J E and Ortiz M 2002 Time-discretized variational formulation of nonsmooth frictional contact *Int. J. Num. Methods in Engineering* **53**, 1801-1829

[SanCal:94] Sanz-Serna J M and Calvo M P 1994 *Numerical Hamiltonian Problems* (Chapman& Hall, London)

[Shi:02] Shi Y 2002 Symplectic structure of Discrete Hamiltonian Systems, *J. Math. Anal. Appl.* **266**, 472-478