Higher Order Variational Integrators: a polynomial approach

Cédric M. Campos

HIGHER ORDER VARIATIONAL INTEGRATORS: A POLYNOMIAL APPROACH

CÉDRIC M. CAMPOS

Abstract. We reconsider the variational derivation of symplectic partitioned Runge-Kutta schemes. Such type of variational integrators are of great importance since they integrate mechanical systems with high order accuracy while preserving the structural properties of these systems, like the symplectic form, the evolution of the momentum maps or the energy behaviour. Also they are easily applicable to optimal control problems based on mechanical systems as proposed in Ober-Blöbaum et al. [2011].

Following the same approach, we develop a family of variational integrators to which we refer as symplectic Galerkin schemes in contrast to symplectic partitioned Runge-Kutta. These two families of integrators are, in principle and by construction, different one from the other. Furthermore, the symplectic Galerkin family can as easily be applied in optimal control problems, for which Campos et al. [2012b] is a particular case.

1. Introduction

In recent years, much effort in designing numerical methods for the time integration of (ordinary) differential equations has been put into schemes which are *structure preserving* in the sense that important *qualitative* features of the original dynamics are preserved in its time discretization, *cf.* the recent monograph Hairer et al. [2010]. A particularly elegant way to, *e.g.* derive symplectic integrators, is by discretizing Hamilton's principle as suggested by Suris [1990], Veselov [1988], see also Marsden and West [2001], Sanz-Serna and Calvo [1994].

However most part of the theory and examples rely on second order schemes, hence some effort must still be put into the development of accurate higher order schemes that, in long term simulations, can drastically reduce the overall computational cost. A clear example are the so called symplectic partitioned Runge-Kutta methods that integrate mechanical systems driven by a Lagrangian $L\colon (q,\dot q)\in\mathbb{R}^n\times\mathbb{R}^n\mapsto L(q,\dot q)\in\mathbb{R}$ and, possibly, by a force $f\colon (q,\dot q)\in\mathbb{R}^n\times\mathbb{R}^n\mapsto p=f(q,\dot q)\in\mathbb{R}^n$. A detailed study of such methods can be found in Hairer et al. [2010].

2010 *Mathematics Subject Classification.* Primary 65P10, Secondary 65L06, 65K10, 49Mxx.

Key words and phrases. geometric mechanics, Euler-Lagrange equation, discrete mechanics, geometric integration, variational integrator, higher-order, Runge-Kutta methods.

Subject. On the occasion of the 23rd Congress on Differential Equations and Applications, CEDYA 2013, held in Castellón, Spain.

Acknowledgments. This work has been partially supported by MEC (Spain) Grants MTM2010-21186-C02-01, MTM2011-15725E, the ICMAT Severo Ochoa project SEV-2011-0087 and the European project IRSES-project "Geomech-246981". Besides, the author wants to specially thank professors Sina Ober-Blöbaum, from the University of Paderborn, and David Martín de Diego, from the Instituto de Ciencias Matemáticas, for fruitful conversations on the subject and their constant support.

A partitioned Runge-Kutta method is an s-stage integrator $(q_0, p_0) \mapsto (q_1, p_1)$ given by the equations

$$q_1 = q_0 + h \sum_{j=1}^{s} b_j \dot{Q}_j \,, \qquad p_1 = p_0 + h \sum_{j=1}^{s} \bar{b}_j \dot{P}_j \,, \tag{1a}$$

$$Q_i = q_0 + h \sum_{j=1}^{s} a_{ij} \dot{Q}_j \,, \qquad P_i = p_0 + h \sum_{j=1}^{s} \bar{a}_{ij} \dot{P}_j \,, \tag{1b}$$

$$P_i = \frac{\partial L}{\partial \dot{q}}(Q_i, \dot{Q}_i) \,, \qquad \dot{P}_i = \frac{\partial L}{\partial \dot{q}}(Q_i, \dot{Q}_i) + f(Q_i, \dot{Q}_i) \,, \tag{1c}$$

where (b_j, a_{ij}) and $(\bar{b}_j, \bar{a}_{ij})$ are two different Runge-Kutta methods associated to collocation points $0 \leq c_1 < \ldots < c_s \leq 1$ and time step $h > 0$.

It is shown that the previous integrator is *symplectic* whenever the two sets of coefficients satisfy the relations

$$b_i \bar{a}_{ij} + \bar{b}_j a_{ji} = b_i \bar{b}_j \,, \tag{2a}$$

$$b_i = \bar{b}_i \,. \tag{2b}$$

In fact, it can be derived as a geometric variational integrator, see Marsden and West [2001], Suris [1990].

The paper is structured as follows: Section 2 is a short introduction to Discrete Mechanics and Section 3 describes the variational derivation of higher order schemes using polynomial collocation. Finally we briefly enumerate the relations and differences between symplectic partitioned Runge-Kutta schemes and symplectic Galerkin ones in Section 4 and conclude by outlining future research directions in Section 5.

2. Discrete Mechanics and Variational Integrators

One of the main subjects of Geometric Mechanics is the study of dynamical systems governed by a Lagrangian. Typically they consider mechanical systems with *configuration manifold* Q together with a *Lagrangian function* $L\colon TQ \to \mathbb{R}$, where the associated *state space* TQ describes the position and velocity of a particle moving in the system. Usually, the Lagrangian takes the form of kinetic minus potential energy, $L(q, \dot{q}) = K(q, \dot{q}) - V(q) = \frac{1}{2}\,\dot{q}^T \cdot M(q) \cdot \dot{q} - V(q)$, for some (positive definite) *mass matrix* $M(q)$.

A consequence of the *principle of least action*, also known as *Hamilton's principle*, establishes that the natural motions $q\colon [0, T] \to Q$ of the system are characterized by the celebrated *Euler-Lagrange equation* (refer to Abraham and Marsden [1978]),

$$\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} = 0 \,. \tag{3}$$

When the Lagrangian is *regular*, that is when the velocity Hessian matrix $\partial^2 L / \partial \dot{q}^2$ is non-degenerate, the Lagrangian induces a well defined map, the *Lagrangian flow*, $F_L^t\colon TQ \to TQ$ by $F_L(q_0, \dot{q}_0) := (q(t), \dot{q}(t))$, where $q \in \mathcal{C}^2([0, T], Q)$ is the unique solution of the Euler-Lagrange equation (3) with initial condition $(q_0, \dot{q}_0) \in TQ$. By means of the *Legendre transform* $\mathrm{leg}_L : (q, \dot{q}) \in TQ \mapsto (q, p = \frac{\partial L}{\partial \dot{q}}|_{(q,\dot{q})}) \in T^*Q$, where T^*Q is the *phase space* of positions plus momenta, one may transform the Lagrangian flow into the *Hamiltonian flow* $F_H^t\colon T^*Q \to T^*Q$ defined by $F_H^t(q_0, p_0) := \mathrm{leg}_L(q(t), \dot{q}(t))$.

Moreover, different preservation laws are present in these systems. For instance the Hamiltonian flow preserves the natural symplectic structure of T^*Q and the total energy of the system. Also, if the Lagrangian possess Lie group symmetries, then *Noether's theorem* asserts that some quantities are conserved, like for instance the linear momentum and/or the angular momentum.

Discrete Mechanics is, roughly speaking, a discretization of Geometric Mechanics theory. As a result, one obtains a set of discrete equations equivalent to the Euler-Lagrange equation (3) above but, instead of a direct discretization of the ODE, the latter are derived from a discretization of the base objects of the theory, the state space TQ, the Lagrangian L, etc. In fact, one seeks for a sequence $\{(t_0, q_0), (t_1, q_1), \ldots, (t_n, q_n)\}$ that approximates the actual trajectory $q(t)$ of the system ($q_k \approx q(t_k)$), for a constant time-step $h = t_{k+1} - t_k > 0$.

A *variational integrator* is an iterative rule that outputs this sequence and it is derived in an analogous manner to the continuous framework. Given a discrete Lagrangian $L_d\colon Q \times Q \to \mathbb{R}$, which is in principle thought to approximate the continuous Lagrangian action over a short time

$$L_d(q_k, q_{k+1}) \approx \int_{t_k}^{t_{k+1}} L(q(t), \dot{q}(t))\mathrm{d}t\,,$$

one applies a variational principle to derive the well-known discrete Euler-Lagrange (DEL) equation,

$$D_1 L_d(q_k, q_{k+1}) + D_2 L_d(q_{k-1}, q_k) = 0\,, \tag{4}$$

where D_i stands for the partial derivative with respect to the i-th component. The equation defines an integration rule of the type $(q_{k-1}, q_k) \mapsto (q_k, q_{k+1})$, however if we define the pre- and post-momenta

$$p_k^- := -D_1 L_d(q_k, q_{k+1}) \qquad \text{and} \qquad p_k^+ := D_2 L_d(q_{k-1}, q_k)\,, \tag{5}$$

the Euler-Lagrange equation (3) is read as the momentum matching $p_k^- = p_k^+ =: p_k$ and defines an integration rule of the type $(q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$.

The nice part of the story is that the integrators derived in this way naturally preserve (or nearly preserve) the quantities that are preserved in the continuous framework, the symplectic form, the total energy and, in presence of symmetries, the linear and/or angular momentum (for more details, see Marsden and West [2001]). Furthermore, other aspects of the continuous theory can be "easily" adapted, symmetry reduction Campos et al. [2012a], Colombo et al. [2012], Iglesias et al. [2008], constraints Johnson and Murphey [2009], Kobilarov et al. [2010], control forces Campos et al. [2012b], Ober-Blöbaum et al. [2011], etc.

3. Higher Order Variational Integrators

Higher order variational integrators for time dependent or independent systems (HOVIt) are a class of integrators that, by using a multi-stage approach, aim at a higher order accuracy on the computation of the natural trajectories of a mechanical system while preserving some intrinsic properties of such systems. In particular, symplectic-partitioned Runge-Kutta methods (spRK) and, what we call here, symplectic Galerkin methods (sG) are s-stage variational integrators of order up to $2s$.

The derivation of these methods follows a general scheme. For a fixed time step h, one considers a series of points q_k, refereed as macro-nodes. Between each couple of macro-nodes (q_k, q_{k+1}), one also considers a set of micro-data, the s stages: For the particular cases of sG and spRK methods, micro-nodes $Q_1, \ldots, Q_s$ and micro-velocities $\dot{Q}_1, \ldots, \dot{Q}_s$, respectively. Both macro-nodes and micro-data (micro-nodes or micro-velocities) are required to satisfy a variational principle, giving rise to a set of equations, which properly combined, define the final integrator.

In what follows, we will use the following notation: Let $0 \leq c_1 < \ldots < c_s \leq 1$ denote a set of collocation times and consider the associated Lagrange polynomials and nodal weights, that is,

$$l^j(t) := \prod_{i \neq j} \frac{t - c_i}{c_j - c_i} \qquad \text{and} \qquad b_j := \int_0^1 l^j(t) \mathrm{d}t \,,$$

respectively. Note that the pair of (c_i, b_i)'s define a quadrature rule and that, for appropriate c_i's, this rule may be a Gaussian-like quadrature, for instance, Gauss-Legendre, Gauss-Lobatto, Radau or Chebyschev.

Now, for the sake of simplicity and independently on the method, we will use the same notation for the nodal coefficients. We define for spRK and sG, respectively,

$$a_{ij} := \int_0^{c_i} l^j(t) \mathrm{d}t \qquad \text{and} \qquad a_{ij} := \left.\frac{\mathrm{d}l^j}{\mathrm{d}t}\right|_{c_i} .$$

Moreover, for spRK, we will also use the nodal weights and coefficients $(\bar{b}_j, \bar{a}_{ij})$ given by Equation (2) and, for sG, the source and target coefficients

$$\alpha^j := l^j(0) \qquad \text{and} \qquad \beta^j := l^j(1) \,.$$

Finally, if L denotes a Lagrangian from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$, then we define

$$P_i := \left.\frac{\partial L}{\partial \dot{q}}\right|_i = \left.\frac{\partial L}{\partial \dot{q}}\right|_{(Q_i, \dot{Q}_i)} \qquad \text{and} \qquad \dot{P}_i := \left.\frac{\partial L}{\partial q}\right|_i = \left.\frac{\partial L}{\partial q}\right|_{(Q_i, \dot{Q}_i)} ,$$

where $(Q_i, \dot{Q}_i)$ are couples of micro-nodes and micro-velocities given by each method. Besides, D_i will stand for the partial derivative with respect to the i-th component.

3.1. **Symplectic-Partitioned Runge-Kutta Methods.** Although the variational derivation of spRK methods in the framework of Geometric Mechanics is an already known fact (see Marsden and West [2001] for an "intrinsic" derivation, as the current, or Hairer et al. [2010] for a "constrained" one), we present it here again in order to ease the comprehension of and the comparison with sG methods below.

Given a point $q_0 \in \mathbb{R}^n$ and vectors $\{\dot{Q}_i\}_{i=1,\ldots,s} \subset \mathbb{R}^n$, we define the polynomial curves

$$\dot{Q}(t) := \sum_{j=1}^{s} l^j(t/h) \dot{Q}_j \qquad \text{and} \qquad Q(t) := q_0 + h \sum_{j=1}^{s} \int_0^{t/h} l^j(\tau) \mathrm{d}\tau \dot{Q}_j \,.$$

We have

$$\dot{Q}_i = \dot{Q}(h \cdot c_i) \quad \text{and} \quad Q_i := Q(h \cdot c_i) = q_0 + h \sum_{j=1}^{s} a_{ij} \dot{Q}_j \,.$$

Note that the polynomial curve Q is uniquely determined by q_0 and $\{\dot{Q}_i\}_{i=1,\dots,s}$. In fact, it is the unique polynomial curve Q of degree s such that $Q(0) = q_0$ and $\dot{Q}(h \cdot c_i) = \dot{Q}_i$. However, if we define the configuration point

$$q_1 := Q(h \cdot 1) = q_0 + h \sum_{j=1}^{s} b_j \dot{Q}_j \tag{6}$$

and consider it fixed, then Q is uniquely determined by q_0, q_1 and the $\dot{Q}_i$'s but one. Namely, take any $1 \leq i_0 \leq s$ such that $b_{i_0} \neq 0$ and fix it, then

$$\dot{Q}_{i_0} = \left(\frac{q_1 - q_0}{h} - \sum_{j \neq i_0} b_j \dot{Q}_j \right) / b_{i_0}$$

and we obtain the following relations that will be useful in what comes next

$$\frac{\partial(Q_i, Q_{i_0})}{\partial(q_0, Q_j, q_1)} = \begin{pmatrix} 0 & \delta_i^j & 0 \\ -\frac{1}{hb_{i_0}} & -\frac{b_j}{b_{i_0}} & \frac{1}{hb_{i_0}} \end{pmatrix}, \ i, j \neq i_0.$$

We now define the multi-vector discrete Lagrangian

$$L_d(\dot{Q}_1, \dots, \dot{Q}_s) := h \sum_{i=1}^{s} b_i L(Q_i, \dot{Q}_i) \,.$$

Although not explicitly stated, it also depends on q_0. The two-point discrete Lagrangian is then

$$L_d(q_0, q_1) := \underset{\mathcal{P}^s([0,h], \mathbb{R}^n, q_0, q_1)}{\text{ext}} L_d(\dot{Q}_1, \dots, \dot{Q}_s)$$

where $\mathcal{P}^s([0,h], \mathbb{R}^n, q_0, q_1)$ is the space of polynomials Q of order s from $[0,1]$ to $\mathbb{R}^n$ such that $Q(0) = q_0$ and $Q(h) = q_1$ and the vectors $\dot{Q}_i$'s determine such polynomials as discussed above. The extremal is realized by a polynomial $Q \in \mathcal{P}^s([0,h], \mathbb{R}^n, q_0, q_1)$ such that

$$\delta L_d(\dot{Q}_1, \dots, \dot{Q}_s) \cdot (\delta \dot{Q}_1, \dots, \delta \dot{Q}_s) = 0 \tag{7}$$

for any variations $(\delta \dot{Q}_1, \dots, \delta \dot{Q}_s)$, taking into account that

$$\delta q_0 = \delta q_1 = 0 \qquad \text{and} \qquad \delta \dot{Q}_{i_0} = \sum_{j \neq i_0} \frac{\partial \dot{Q}_{i_0}}{\partial \dot{Q}_j} \delta \dot{Q}_j \,.$$

For convenience, the previous equation is developed afterwards.

By the momenta-matching rule (5), we have that

$$\begin{aligned} -p_0 &= D_1 L_d(q_0, q_1) \\ &= D_{i_0} L_d(\dot{Q}_1, \dots, \dot{Q}_s) \partial \dot{Q}_{i_0} / \partial q_0 + D_{q_0} L_d(\dot{Q}_1, \dots, \dot{Q}_s) \\ &= -D_{i_0} L_d(\dot{Q}_1, \dots, \dot{Q}_s)/(hb_{i_0}) + D_{q_0} L_d(\dot{Q}_1, \dots, \dot{Q}_s) \,, \end{aligned}$$

$$\begin{aligned} p_1 &= D_2 L_d(q_0, q_1) \\ &= D_{i_0} L_d(\dot{Q}_1, \dots, \dot{Q}_s) \partial \dot{Q}_{i_0} / \partial q_1 \\ &= D_{i_0} L_d(\dot{Q}_1, \dots, \dot{Q}_s)/(hb_{i_0}) \,. \end{aligned}$$

where D_{q_0} stands for the partial derivative with respect to q_0. Combining both equations, we obtain that

$$D_{i_0}L_d(\dot{Q}_1,\dots,\dot{Q}_s) = hb_{i_0}p_1 \qquad \text{and} \qquad p_1 = p_0 + D_{q_0}L_d(\dot{Q}_1,\dots,\dot{Q}_s)\,.$$

Coming back to Equation (7), we have that

$$\begin{aligned}
&\delta L_d(\dot{Q}_1,\dots,\dot{Q}_s)\cdot(\delta\dot{Q}_1,\dots,\delta\dot{Q}_s) = \\
&\qquad = \sum_{j\neq i_0} D_jL_d(\dot{Q}_1,\dots,\dot{Q}_s)\delta\dot{Q}_j + D_{i_0}L_d(\dot{Q}_1,\dots,\dot{Q}_s)\delta\dot{Q}_{i_0} \\
&\qquad = \sum_{j\neq i_0}\left(D_jL_d(\dot{Q}_1,\dots,\dot{Q}_s) + \frac{\partial\dot{Q}_{i_0}}{\partial\dot{Q}_j}D_{i_0}L_d(\dot{Q}_1,\dots,\dot{Q}_s)\right)\delta\dot{Q}_j\,.
\end{aligned}$$

Therefore, for $j\neq i_0$, we have that

$$D_jL_d(\dot{Q}_1,\dots,\dot{Q}_s) = b_j/b_{i_0}\cdot D_{i_0}L_d(\dot{Q}_1,\dots,\dot{Q}_s)\,.$$

Thus, for any $j=1,\dots,s$, we have that

$$D_jL_d(\dot{Q}_1,\dots,\dot{Q}_s) = hb_jp_1\,. \tag{8}$$

The integrator is defined by

$$D_jL_d(\dot{Q}_1,\dots,\dot{Q}_s) = hb_jp_1\,, \tag{9a}$$

$$q_1 = q_0 + h\sum_{j=1}^{s} b_j\dot{Q}_j\,, \tag{9b}$$

$$p_1 = p_0 + D_{q_0}L_d(\dot{Q}_1,\dots,\dot{Q}_s)\,. \tag{9c}$$

Besides, using the definition of the discrete Lagrangian, we have

$$\begin{aligned}
D_jL_d(\dot{Q}_1,\dots,\dot{Q}_s) &= h\sum_{i=1}^{s} b_i\left(\frac{\partial L}{\partial q}\Big|_i\frac{\partial Q_i}{\partial\dot{Q}_j} + \frac{\partial L}{\partial\dot{q}}\Big|_i\frac{\partial\dot{Q}_i}{\partial\dot{Q}_j}\right) \\
&= h^2\sum_{i=1}^{s} b_ia_{ij}\dot{P}_i + hb_jP_j\,, \\
D_{q_0}L_d(\dot{Q}_1,\dots,\dot{Q}_s) &= h\sum_{i=1}^{s} b_i\left(\frac{\partial L}{\partial q}\Big|_i\frac{\partial Q_i}{\partial q_0} + \frac{\partial L}{\partial\dot{q}}\Big|_i\frac{\partial\dot{Q}_i}{\partial q_0}\right) \\
&= h\sum_{i=1}^{s} b_i\dot{P}_i\,.
\end{aligned}$$

Therefore, we may write

$$P_j = p_0 + h\sum_{i=1}^{s} b_i(1 - a_{ij}/b_j)\dot{P}_i = p_0 + h\sum_{i=1}^{s}\bar{a}_{ji}\dot{P}_i\,,$$

$$p_1 = p_0 + h\sum_{i=1}^{s} b_i\dot{P}_i = p_0 + h\sum_{i=1}^{s}\bar{b}_i\dot{P}_i\,,$$

were $\bar{a}_{ij}$ and $\bar{b}_i$ are given by Equation (2).

In summary, we have recovered all the equations that define the spRK integrator without forces, Equation (1) and (2), that is

$$q_1 = q_0 + h\sum_{j=1}^{s} b_j\dot{Q}_j\,, \qquad p_1 = p_0 + h\sum_{j=1}^{s} \bar{b}_j\dot{P}_j\,, \tag{10a}$$

$$Q_i = q_0 + h\sum_{j=1}^{s} a_{ij}\dot{Q}_j\,, \qquad P_i = p_0 + h\sum_{j=1}^{s} \bar{a}_{ij}\dot{P}_j\,, \tag{10b}$$

$$P_i = \frac{\partial L}{\partial \dot{q}}(Q_i,\dot{Q}_i)\,, \qquad \dot{P}_i = \frac{\partial L}{\partial \dot{q}}(Q_i,\dot{Q}_i)\,. \tag{10c}$$

3.2. **Symplectic Galerkin Methods (of 0th kind).** Galerkin methods are a class of methods to transform a problem given by a continuous operator (such as a differential operator) to a discrete problem. As such, spRK methods falls into the scope of this technique and could be also classified as "symplectic Galerkin" methods (of 1st kind). However, we want to stress here the difference between what is called spRK and what we here refer as sG. The wording should not be confused by the one used in Marsden and West [2001].

Given points $\{Q_i\}_{i=1,\ldots,s} \subset \mathbb{R}^n$, we define the polynomial curves

$$Q(t) := \sum_{j=1}^{s} l^j(t/h)Q_j \qquad \text{and} \qquad \dot{Q}(t) := \frac{1}{h}\sum_{j=1}^{s} \dot{l}^j(t/h)Q_j\,.$$

We have

$$Q_i = Q(h\cdot c_i) \quad \text{and} \quad \dot{Q}_i := \dot{Q}(h\cdot c_i) = \frac{1}{h}\sum_{j=1}^{s} a_{ij}Q_j\,.$$

Note that the polynomial curve Q is uniquely determined by the points $\{Q_i\}_{i=1,\ldots,s}$. In fact, it is the unique polynomial curve Q of degree s such that $Q(h\cdot c_i) = Q_i$. However, if we define the configuration points

$$q_0 := Q(h\cdot 0) = \sum_{j=1}^{s} \alpha^j Q_j \qquad \text{and} \qquad q_1 := Q(h\cdot 1) = \sum_{j=1}^{s} \beta^j Q_j \tag{11}$$

and consider them fixed, then Q is uniquely determined by q_0, q_1 and the Q_i's but a couple. For instance, we may consider Q_1 and Q_s as functions of the others, since the relations (11) define a system of linear equations where the coefficient matrix has determinant $\gamma := \alpha^1\beta^s - \alpha^s\beta^1 \neq 0$ (if and only if $c_1 \neq c_s$). More precisely,

$$\begin{pmatrix} Q_1 \\ Q_s \end{pmatrix} = \frac{1}{\gamma}\begin{pmatrix} \beta^s & -\alpha^s \\ -\beta^1 & \alpha^1 \end{pmatrix}\begin{pmatrix} q_0 - \sum_{j=2}^{s-1}\alpha^j Q_j \\ q_1 - \sum_{j=2}^{s-1}\beta^j Q_j \end{pmatrix},$$

from where we obtain the following relations that will be useful in what comes next

$$\frac{\partial(Q_1,Q_i,Q_s)}{\partial(q_0,Q_j,q_1)} = \frac{1}{\gamma}\begin{pmatrix} \beta^s & \alpha^s\beta^j - \alpha^j\beta^s & -\alpha^s \\ 0 & \gamma\delta_i^j & 0 \\ -\beta^1 & \alpha^j\beta^1 - \alpha^1\beta^j & \alpha^1 \end{pmatrix},\ i,j = 2,\ldots,s-1.$$

We will also take into account that

$$\frac{\partial \dot{Q}_i}{\partial(q_0,Q_j,q_1)} = \frac{1}{h}\sum_{k=1}^{s} a_{ik}\frac{\partial Q_k}{\partial(q_0,Q_j,q_1)}\,,\ i = 1,\ldots,s,\ j = 2,\ldots,s-1.$$

We now define the multi-point discrete Lagrangian

$$L_d(Q_1,\dots,Q_s) := h\sum_{i=1}^{s} b_i L(Q_i,\dot Q_i)\,.$$

The two-point discrete Lagrangian is then

$$L_d(q_0,q_1) := \operatorname*{ext}_{\mathcal{P}^s([0,h],\mathbb{R}^n,q_0,q_1)} L_d(Q_1,\dots,Q_s)$$

where $\mathcal{P}^s([0,h],\mathbb{R}^n,q_0,q_1)$ is the space of polynomials Q of order s from $[0,1]$ to $\mathbb{R}^n$ such that the points Q_i's determine such polynomials as discussed above. The extremal is realized by a polynomial $Q\in\mathcal{P}^s([0,h],\mathbb{R}^n,q_0,q_1)$ such that

$$\delta L_d(Q_1,\dots,Q_s)\cdot(\delta Q_1,\dots,\delta Q_s) = 0 \tag{12}$$

for any variations $(\delta Q_1,\dots,\delta Q_s)$, taking into account that

$$\delta q_0 = \delta q_1 = 0 \qquad\text{and}\qquad \delta Q_i = \sum_{j=2}^{s-1}\frac{\partial Q_i}{\partial Q_j}\delta Q_j\,,\ i=1,s.$$

For convenience, the previous equation is developed afterwards.

By the momenta-matching rule (5), we have that

$$\begin{aligned}
-p_0 &= D_1L_d(q_0,q_1)\\
&= \sum_{j=1}^{s} D_jL_d(Q_1,\dots,Q_s)\frac{\partial Q_j}{\partial q_0}\\
&= \beta^s/\gamma\cdot D_1L_d(Q_1,\dots,Q_s) - \beta^1/\gamma\cdot D_sL_d(Q_1,\dots,Q_s)
\end{aligned}$$

$$\begin{aligned}
p_1 &= D_2L_d(q_0,q_1)\\
&= \sum_{j=1}^{s} D_jL_d(Q_1,\dots,Q_s)\frac{\partial Q_j}{\partial q_1}\\
&= -\alpha^s/\gamma\cdot D_1L_d(Q_1,\dots,Q_s) + \alpha^1/\gamma\cdot D_sL_d(Q_1,\dots,Q_s)
\end{aligned}$$

By a linear transformation of both equations, we obtain

$$\begin{aligned}
D_1L_d(Q_1,\dots,Q_s) &= -\alpha^1p_0+\beta^1p_1 \ \text{ and}\\
D_sL_d(Q_1,\dots,Q_s) &= -\alpha^sp_0+\beta^sp_1\,.
\end{aligned}$$

Coming back to Equation (12), we have that

$$\begin{aligned}
&\delta L_d(Q_1,\dots,Q_s)\cdot(\delta Q_1,\dots,\delta Q_s) =\\
&= D_1L_d(Q_1,\dots,Q_s)\delta Q_1 + \sum_{j=2}^{s-1} D_jL_d(Q_1,\dots,Q_s)\delta Q_j + D_sL_d(Q_1,\dots,Q_s)\delta Q_s\\
&= \sum_{j=2}^{s-1}\left[D_1L_d(Q_1,\dots,Q_s)\frac{\partial Q_1}{\partial Q_j} + D_jL_d(Q_1,\dots,Q_s) + D_sL_d(Q_1,\dots,Q_s)\frac{\partial Q_s}{\partial Q_j}\right]\delta Q_j
\end{aligned}$$

Therefore, for $j = 2, \dots, s-1$, we obtain

$$\begin{aligned}\gamma D_j L_d &= (\alpha^j \beta^s - \alpha^s \beta^j) D_1 L_d + (\alpha^1 \beta^j - \alpha^j \beta^1) D_s L_d \\ &= (\alpha^j \beta^s - \alpha^s \beta^j)(-\alpha^1 p_0 + \beta^1 p_1) + (\alpha^1 \beta^j - \alpha^j \beta^1)(-\alpha^s p_0 + \beta^s p_1) \\ &= (\alpha^1 \beta^s - \alpha^s \beta^1)(-\alpha^j p_0 + \beta^j p_1)\,.\end{aligned}$$

Thus, for any $j = 1, \dots, s$, we have that

$$D_j L_d(Q_1, \dots, Q_s) = -\alpha^j p_0 + \beta^j p_1\,. \tag{13}$$

The integrator is defined by

$$D_j L_d(Q_1, \dots, Q_s) = -\alpha^j p_0 + \beta^j p_1\,,\ j = 1, \dots, s; \tag{14a}$$

$$q_0 = \sum_{j=1}^{s} \alpha^j Q_j \qquad \text{and} \qquad q_1 = \sum_{j=1}^{s} \beta^j Q_j \tag{14b}$$

Besides, using the definition of the discrete Lagrangian, we have

$$\begin{aligned}D_j L_d(Q_1, \dots, Q_s) &= h \sum_{i=1}^{s} b_i \left(\frac{\partial L}{\partial q}\Big|_i \frac{\partial Q_i}{\partial \dot{Q}_j} + \frac{\partial L}{\partial \dot{q}}\Big|_i \frac{\partial \dot{Q}_i}{\partial \dot{Q}_j} \right) \\ &= h b_j \dot{P}_j + \sum_{j=1}^{s} b_i a_{ij} P_i\,.\end{aligned}$$

Therefore, we may simply write

$$h b_j \dot{P}_j + \sum_{j=1}^{s} b_i a_{ij} P_i = -\alpha^j p_0 + \beta^j p_1\,.$$

In summary and for a proper comparison, we write the equations that define the sG integrator (without forces) in a pRK fashion, that is

$$q_0 = \sum_{j=1}^{s} \alpha^j Q_j\,, \qquad q_1 = \sum_{j=1}^{s} \beta^j Q_j\,, \tag{15a}$$

$$Q_i = \frac{1}{h} \sum_{j=1}^{s} a_{ij} Q_j\,, \qquad \dot{P}_i = \frac{\beta^i p_1 - \alpha^i p_0}{h \bar{b}_i} + \frac{1}{h} \sum_{j=1}^{s} \bar{a}_{ij} P_j\,, \tag{15b}$$

$$P_i = \frac{\partial L}{\partial \dot{q}}(Q_i, \dot{Q}_i)\,, \qquad \dot{P}_i = \frac{\partial L}{\partial \dot{q}}(Q_i, \dot{Q}_i)\,, \tag{15c}$$

where $b_i a_{ij} + \bar{b}_j \bar{a}_{ji} = 0$ and $b_i = \bar{b}_j$.

We remark that Equation (13) generalizes the ones obtained in Campos et al. [2012b], Leok [2004], where the collocation times are chosen such that $c_1 = 0$ and $c_s = 1$, which is a rather particular case.

4. Relations between spRK and sG

First of all, it is worth to say that, with a little bit of extra technicalities, one can easily include forces into both schemes. As a result, one would only need to redefine in (10) and (15)

$$\dot{P}_i = \frac{\partial L}{\partial \dot{q}}(Q_i, \dot{Q}_i) + f(Q_i, \dot{Q}_i)\,,$$

where $f\colon (q, \dot{q}) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto p = f(q, \dot{q}) \in \mathbb{R}^n$ is the external force.

Both methods can be considered of Galerkin type. In this sense, spRK and sG could be refereed as a symplectic Galerkin integrators of 1st and 0th kind, respectively, since spRK is derived from the 1st derivative of an extremal polynomial and sG from the polynomial itself. At this point, a very natural question could arise: Actually are spRK and sG two different integrator schemes? Even though the derivations of both methods are quite similar, they are in general different (although they could coincide for particular choices of the Lagrangian, the collocation times and the integral quadrature). A weak but still fair argument for this is that, at each step, spRK relies on the determination of the micro-velocities $\dot{Q}_i$, while sG does so on the micro-nodes Q_i. All the other "variables" are then computed form the determined micro-data.

With respect to the accuracy of the schemes, for any Gaussian quadrature (Gauss-Legendre, Gauss-Lobatto, Radau and Chebyschev) and any method (spRK and sG), the schemes have convergence order $2s-2$, expect for the combination of Gauss-Lobatto together with spRK which is $2s$, being s the number of internal stages.

		spRK	sG
micro-data		$\dot{Q}_i$	Q_i
polynomial degree		s	$s-1$
variational eq.'s		$s+1$	s
extra equations		1	2
quadrature	Gauss-Legendre	$2s$	$2s-2$
	Gauss-Lobatto	$2s-2$	$2s-2$
	Radau	$2s-2$	$2s-2$
	Chebyschev	$2s-2$	$2s-2$
		order method	

TABLE 1. Comparison of s-stage variational integrators.

Let's finish underlying that sG, as spRK, is inherently symplectic.

5. CONCLUSIONS

In this work, by revisiting the variational derivation of spRK methods Hairer et al. [2010], Marsden and West [2001], we have presented a new class of higher order variational integrators within the family of Galerkin schemes. These integrators are symplectic per construction and, therefore, well suited for long term simulations, where the higher order accuracy of the schemes can be exploited to reduce the overall computational cost. Also, they can be easily adapted to implement constraints or symmetry reduction Campos et al. [2012a], Iglesias et al. [2008], Marsden and West [2001] or, together with an NLP solver, to integrate optimal control problems Campos et al. [2012b], Ober-Blöbaum et al. [2011].

For the future, the sG schemes deserves a proper analysis to establish the actual differences with respect to spRK schemes and results the convergence rates. And to further take advantage of the high accuracy of the methods, we envisage to design time adaptive algorithms. Joint work with O. Junge (TUM), S. Ober-Blöbaum (UPB) and E. Trélat (CNRS-UPMC) on the control direction has already started in order to generalize Campos et al. [2012b] and clarify some aspects of Hager [2000]. As well, we have begun digging along the lines of constrained systems and higher order Lagrangians.

References

Ralph Abraham and Jerrold E. Marsden. *Foundations of mechanics.* Benjamin/Cummings Publishing Co. Inc. Advanced Book Program, Reading, Mass., 1978. ISBN 0-8053-0102-X. Second edition, revised and enlarged, With the assistance of Tudor Raţiu and Richard Cushman.

Cédric M. Campos, Hernán Cendra, Viviana Alejandra Díaz, and David Martín de Diego. Discrete Lagrange-d'Alembert-Poincaré equations for Euler's disk. *Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Math. RACSAM*, 106(1): 225–234, 2012a. ISSN 1578-7303. doi: 10.1007/s13398-011-0053-3. URL `http://dx.doi.org/10.1007/s13398-011-0053-3`.

Cédric M. Campos, Oliver Junge, and Sina Ober-Blöbaum. Higher order variational time discretization of optimal control problems. In *20th International Symposium on Mathematical Theory of Networks and Systems*, Melbourne, Australia, 2012b.

Leonardo Colombo, Fernando Jiménez, and David Martín de Diego. Discrete second-order Euler-Poincaré equations. Applications to optimal control. *Int. J. Geom. Methods Mod. Phys.*, 9(4):1250037, 20, 2012. ISSN 0219-8878. doi: 10.1142/S0219887812500375. URL `http://dx.doi.org/10.1142/S0219887812500375`.

William W. Hager. Runge-Kutta methods in optimal control and the transformed adjoint system. *Numer. Math.*, 87(2):247–282, 2000. ISSN 0029-599X. doi: 10.1007/s002110000178. URL `http://dx.doi.org/10.1007/s002110000178`.

Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2010. ISBN 978-3-642-05157-9. Structure-preserving algorithms for ordinary differential equations, Reprint of the second (2006) edition.

David Iglesias, Juan C. Marrero, David Martín de Diego, and Eduardo Martínez. Discrete nonholonomic Lagrangian systems on Lie groupoids. *J. Nonlinear Sci.*, 18(3):221–276, 2008. ISSN 0938-8974. doi: 10.1007/s00332-007-9012-8. URL `http://dx.doi.org/10.1007/s00332-007-9012-8`.

Elliot R. Johnson and Todd D. Murphey. Dangers of two-point holonomic constraints for variational integrators. In *Proceedings of the 2009 conference on American Control Conference*, ACC'09, pages 4723–4728, Piscataway, NJ, USA, 2009. IEEE Press. ISBN 978-1-4244-4523-3. URL `http://dl.acm.org/citation.cfm?id=1702715.1703110`.

Marin Kobilarov, Jerrold E. Marsden, and Gaurav S. Sukhatme. Geometric discretization of nonholonomic systems with symmetries. *Discrete Contin. Dyn. Syst. Ser. S*, 3(1):61–84, 2010. ISSN 1937-1632. doi: 10.3934/dcdss.2010.3.61. URL `http://dx.doi.org/10.3934/dcdss.2010.3.61`.

Melvin Leok. *Foundations of computational geometric mechanics.* PhD thesis, California Institute of Technology, May 2004. URL `http://resolver.caltech.edu/CaltechETD:etd-03022004-000251`. Thesis (Ph.D.)–California Institute of Technology.

J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numer.*, 10:357–514, 2001. ISSN 0962-4929. doi: 10.1017/S096249290100006X. URL `http://dx.doi.org/10.1017/S096249290100006X`.

S. Ober-Blöbaum, O. Junge, and J.E. Marsden. Discrete mechanics and optimal control: an analysis. *ESAIM Control Optim. Calc. Var.*, 17

(2):322–352, 2011. ISSN 1292-8119. doi: 10.1051/cocv/2010012. URL http://dx.doi.org/10.1051/cocv/2010012.

J. M. Sanz-Serna and M. P. Calvo. *Numerical Hamiltonian problems*, volume 7 of *Applied Mathematics and Mathematical Computation*. Chapman & Hall, London, 1994. ISBN 0-412-54290-0.

Yu. B. Suris. Hamiltonian methods of Runge-Kutta type and their variational interpretation. *Mat. Model.*, 2(4):78–87, 1990. ISSN 0234-0879.

A. P. Veselov. Integrable systems with discrete time, and difference operators. *Funktsional. Anal. i Prilozhen.*, 22(2):1–13, 96, 1988. ISSN 0374-1990. doi: 10.1007/BF01077598. URL http://dx.doi.org/10.1007/BF01077598.

Instituto de Ciencias Matemáticas, Campus de Cantoblanco, UAM, Calle Nicolás Cabrera, 15, 28049 Madrid (Spain)