NORGES TEKNISK-NATURVITENSKAPELIGE UNIVERSITET

Symplectic integrators for index one constraints

by

Robert I McLachlan[1], Klas Modin[1], Olivier Verdier[2], and Matt Wilkins[1]

[1] *Institute of Fundamental Sciences, Massey University, Private Bag 11 222, Palmerston North 4442, New Zealand*
[2] *Department of Mathematical Sciences, NTNU, 7491 Trondheim, Norway*

PREPRINT
NUMERICS NO. 3/2012

NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY
TRONDHEIM, NORWAY

This report has URL
`http://www.math.ntnu.no/preprint/numerics/2012/N3-2012.pdf`
Address: Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

Symplectic integrators for index one constraints

Robert I McLachlan[1], Klas Modin[1], Olivier Verdier[2], and Matt Wilkins[1]

[1] *Institute of Fundamental Sciences, Massey University, Private Bag 11 222, Palmerston North 4442, New Zealand*
[2] *Department of Mathematical Sciences, NTNU, 7491 Trondheim, Norway*

July 18, 2012

We show that symplectic Runge–Kutta methods provide effective symplectic integrators for Hamiltonian systems with index one constraints. These include the Hamiltonian description of variational problems subject to position and velocity constraints nondegenerate in the velocities, such as those arising in subRiemannian geometry.

1 Introduction: constrained Hamiltonian systems

We consider constrained Hamiltonian systems of the form

$$J\dot{z} = \nabla H(z), \quad z \in C \subset \mathbb{R}^m \tag{1}$$

where $z \in \mathbb{R}^m$, $\omega := \frac{1}{2}dz \wedge Jdz$ is a closed 2-form[1], $H\colon \mathbb{R}^m \to R$ is a Hamiltonian, and C is a constraint submanifold such that $i^*\omega$ (where $i\colon C \to \mathbb{R}^m$ is the inclusion of C in $\mathbb{R}^m$) is nondegenerate, i.e., such that $(C, i^*\omega)$ is a symplectic manifold. The dynamics on C depends only on the restricted Hamiltonian i^*H and restricted symplectic form $i^*\omega$. Systems with holonomic (position) constraints take this form, with $z = (q,p)$, $\omega = dq \wedge dp$, and $C = \{(q,p) : h_i(q) = 0,\ Dh_i(q)H_p(q,p) = 0,\ 1 \le i \le k\}$ consisting of primary and secondary constraints; a nondegeneracy assumption ensures that C is symplectic. The widely used RATTLE method [3, 5] provides a (class of) symplectic integrators for this case when J is constant: it integrates in coordinates z with Lagrange multipliers to enforce the constraints. However, there are no known symplectic integrators for general constrained Hamiltonian systems of the form of Eq. (1).

In this paper we describe a class of symplectic integrators for a class of Hamiltonian systems of the form (1) containing constraints that can depend on both position and velocity[2]. The class includes the Hamiltonian description of problems arising from variational problems in subRiemannian geometry, in which velocities are constrained to lie in a given (nonintegrable) distribution. We give this application first. In the following proposition,

[1] We use vector notation in wedge products, writing $dq \wedge dp$ for $\sum_{i=1}^m dq_i \wedge dp_i$ and $dz \wedge Jdz$ for $\sum_{i,j=1}^m J_{ij} dz_i \wedge dz_j$, where the dimension m is determined from the context.

[2] We do not use the term *nonholonomic* which is reserved for constrained mechanical systems satisfying the Lagrange–d'Alembert principle, whose flow is not in general symplectic.

the linear independence assumption on the constraints is equivalent to constraining the velocities to lie in a k-dimensional distribution of the tangent space of the positions.

Proposition 1. *Let M be a symmetric nonsingular $n \times n$ mass matrix, $V\colon \mathbb{R}^n \to \mathbb{R}$ a smooth potential, $g_i : \mathbb{R}^n \to \mathbb{R}^n$, $i = 1, \ldots, k$ be smooth functions whose values are linearly independent for all arguments, and q be a smooth extremal with fixed endpoints for the functional*

$$S(q) = \int_{t_0}^{t_1} L(t, q, \dot{q})dt = \int_{t_0}^{t_1} \left(\frac{1}{2}\dot{q}^T M \dot{q} - V(q) \right) dt \tag{2}$$

subject to the constraints $g_i(q) \cdot \dot{q} = 0, i = 1, \ldots, k$. Then

$$J\dot{z} = \nabla H(z) \tag{3}$$

where

$$\begin{aligned}
J &= \begin{pmatrix} 0 & -\mathbb{I}_{n\times n} & 0 \\ \mathbb{I}_{n\times n} & 0 & 0 \\ 0 & 0 & 0_{k\times k} \end{pmatrix} \\
z &= \begin{pmatrix} q \\ p \\ \lambda \end{pmatrix} \in \mathbb{R}^{2n+k} \\
p &= M\dot{q} - \sum_{i=1}^k \lambda_i g_i(q) \\
H(z) &= \frac{1}{2} \left(p + \sum_{i=1}^k \lambda_i g_i(q) \right)^T M^{-1} \left(p + \sum_{i=1}^k \lambda_i g_i(q) \right) + V(q)
\end{aligned}$$

and, furthermore, the Euler–Lagrange equations for (2) are equivalent to the generalized Hamiltonian system (3). Eq. (3) forms a constrained Hamiltonian system of the type (1) with constraint submanifold C a graph over (q, p), i.e., $C := \{(q, p, \lambda)\colon \lambda = \tilde{\lambda}(q, p)\}$ and restricted symplectic form $i^*\omega = dq \wedge dp$.

Proof. Introducing Lagrange multipliers $\lambda_1, \ldots, \lambda_k$, the Euler–Lagrange equations for (2) are

$$\frac{d}{dt}\left(\nabla_{\dot{q}} F\right) - \nabla_q F = 0, \tag{4}$$

$$g_i(q) \cdot \dot{q} = 0, \quad i = 1, \ldots, k, \tag{5}$$

where

$$F(q, p, \lambda) = \frac{1}{2}\dot{q}^T M \dot{q} - V(q) - \sum_{i=1}^k \lambda_i g_i(q) \cdot \dot{q}.$$

Expanding out equation (4) gives the Euler–Lagrange equations

$$\frac{d}{dt}\left(M\dot{q} - \sum_{i=1}^k \lambda_i g_i(q) \right) - \nabla_q F = 0$$

$$\frac{d}{dt}\left(M\dot{q} - \sum_{i=1}^k \lambda_i g_i(q) \right) + \left(\nabla V(q) + \sum_{i=1}^k \lambda_i D g_i(q) \dot{q} \right) = 0 \tag{6}$$

Define the conjugate momentum $p \in \mathbb{R}^n$ using the standard Legendre transform

$$p := \nabla_{\dot{q}} F = M\dot{q} - \sum_{i=1}^{k} \lambda_i g_i(q) \tag{7}$$

so that

$$\dot{q} = M^{-1}\left(p + \sum_{i=1}^{k} \lambda_i g_i(q)\right). \tag{8}$$

Using equations (7) and (8) in equation (4) gives the expression for $\dot{p}$

$$\dot{p} = -\nabla V(q) - \sum_{i=1}^{k} \lambda_i Dg_i(q) M^{-1}\left(p + \sum_{j=1}^{k} \lambda_j g_j(q)\right). \tag{9}$$

Define the Hamiltonian, again in the standard way, as $H(q,p,\lambda) := \dot{q}\cdot p - F(q,\dot{q},\lambda)$; explicitly,

$$\begin{aligned}
H &= \dot{q}\cdot p - F \\
&= \dot{q}\cdot p - \frac{1}{2}\dot{q}^T M\dot{q} + V(q) + \sum_{i=1}^{k} \lambda_i g_i(q)\cdot\dot{q} \\
&= \dot{q}\cdot\left(p - \frac{1}{2}M\dot{q} + \sum_{i=1}^{k} \lambda_i g_i(q)\right) + V(q) \\
&= \dot{q}\cdot\left(M\dot{q} - \sum_{i=1}^{k} \lambda_i g_i(q) - \frac{1}{2}M\dot{q} + \sum_{i=1}^{k} \lambda_i g_i(q)\right) + V(q) \\
&= \dot{q}\cdot\left(\frac{1}{2}M\dot{q}\right) + V(q) \\
&= \frac{1}{2}\left(p + \sum_{i=1}^{k} \lambda_i g_i(q)\right)^T M^{-1}\left(p + \sum_{i=1}^{k} \lambda_i g_i(q)\right) + V(q)
\end{aligned}$$

A calculation shows the equivalence of the right hand side of (8) and $\nabla_p H$; of the right hand side of (9) and $-\nabla_q H(q,p,\lambda)$; and of constraints $g_i(q)\cdot\dot{q} = 0$ and $0 = \nabla_\lambda H(q,p,\lambda)$.

The constraints $0 = \nabla_\lambda H(q,p,\lambda)$ are the following set of equations linear in λ,

$$\begin{pmatrix} g_1\cdot M^{-1}g_1 & \cdots & g_1\cdot M^{-1}g_k \\ \vdots & \ddots & \vdots \\ g_k\cdot M^{-1}g_1 & \cdots & g_k\cdot M^{-1}g_k \end{pmatrix}\begin{pmatrix}\lambda_1 \\ \vdots \\ \lambda_k\end{pmatrix} = -\begin{pmatrix} g_1\cdot M^{-1}p \\ \vdots \\ g_k\cdot M^{-1}p\end{pmatrix} \tag{10}$$

which has a unique solution for λ for all q, p because the matrix is $GM^{-1}G^T$ where G is the $k\times n$ matrix whose ith row is g_i^T. The assumption that the g_i are linearly independent means that G has full rank k and hence that $GM^{-1}G^T$ is nonsingular. The constraints therefore have a unique solution for λ that we write as $\lambda = \tilde{\lambda}(q,p)$, that is, the constraint submanifold is a graph over (q,p). Differentiating these constraints with respect to t then yields ODEs for $\dot{\lambda}$, that is, the system (3) has (differentiation) index 1. The symplectic form on C is $\frac{1}{2}dz\wedge Jdz = dq\wedge dp$. □

2 Symplectic integrators for generalized Hamiltonian systems

The Hamiltonian form (3) suggests considering generalized Hamiltonian systems of the form $J\dot{z} = \nabla H(z)$, $z \in \mathbb{R}^m$, where J is a constant antisymmetric matrix, and we do not specify the constraints. Note that many kinds of constrained Hamiltonian systems (including those with holonomic constraints) can be written in this form; the constraint manifold C is constructed as the subset of initial conditions for which the equations have a solution. In general, these equations may not have solutions for all initial conditions; in the extreme case $J = 0$, the equations are purely algebraic. However, it is easily seen that any solutions that do exist do preserve the (possibly degenerate) bilinear form $u^T Jv$, for

$$\frac{d}{dt} u^T Jv = u^T J\dot{v} + \dot{u}^T Jv = u^T H_{zz}(z)v - v^T H_{zz}(z)u = 0$$

where H_{zz} is the Hessian of H—this does not require invertibility of J.

In the particular case of Proposition 1, the generalized Hamiltonian system that is obtained is equivalent to a canonical Hamiltonian system obtained by eliminating the Lagrange multipliers λ. Let $\lambda = \tilde{\lambda}(q,p)$ be the solution to (10). Then Hamilton's equations for $\tilde{H}(q,p) := H(q,p,\lambda(q,p))$ are

$$\begin{aligned}
\dot{q}_i &= \frac{\partial \tilde{H}}{\partial p_i}(q,p) \\
&= \frac{\partial H}{\partial p_i}(q,p,\tilde{\lambda}(q,p)) + \sum_{j=1}^{k} \frac{\partial H}{\partial \lambda_j}(q,p,\tilde{\lambda}(q,p)) \frac{\partial \tilde{\lambda}_j}{\partial p_i}(q,p) \\
&= \frac{\partial H}{\partial p_i}(q,p,\tilde{\lambda}(q,p)) \\
\dot{p}_i &= -\frac{\partial \tilde{H}}{\partial q_i}(q,p) \\
&= -\frac{\partial H}{\partial q_i}(q,p,\tilde{\lambda}(q,p)) - \sum_{j=1}^{k} \frac{\partial H}{\partial \lambda_j}(q,p,\tilde{\lambda}(q,p)) \frac{\partial \tilde{\lambda}_j}{\partial q_i}(q,p) \\
&= -\frac{\partial H}{\partial q_i}(q,p,\tilde{\lambda}(q,p))
\end{aligned}$$

which, together with $\frac{\partial H}{\partial \lambda}(q,p,\tilde{\lambda}(q,p)) = 0$, are equivalent to (3). That is, the two operations of eliminating the Lagrange multipliers and mapping the Hamiltonian to its Hamiltonian vector field commute; this can also be seen abstractly by considering the symplectic manifold C with canonical coordinates (q,p), symplectic form $dq \wedge dp$, and Hamiltonian i^*H.

The midpoint rule is known to be symplectic when the structure matrix J is invertible [3]. However, as for the continuous time case, J need not be invertible.

Proposition 2. *Any solutions of the midpoint rule applied to $J\dot{z} = \nabla H$ preserve the 2-form $\frac{1}{2} dz \wedge J dz$.*

Proof. It must be shown that $dz_1 \wedge Jdz_1 = dz_0 \wedge Jdz_0$, where z_1 is the result of applying the midpoint rule to z_0. The midpoint rule is

$$\frac{Jz_1 - Jz_0}{\Delta t} = \nabla H\left(\frac{z_0 + z_1}{2}\right) := \nabla H(\bar{z}) \tag{11}$$

where Δt is the time step and the method takes a point z_0 to z_1.

Taking exterior derivatives of equation (11) yields

$$Jdz_1 = Jdz_0 + \Delta t H_{zz}(\bar{z})\left(\frac{dz_0 + dz_1}{2}\right)$$

where H_{zz} is the Hessian of H. Then

$$\begin{aligned}
J(dz_1 - dz_0) &= \frac{1}{2}\Delta t H_{zz}(\bar{z})(dz_0 + dz_1) \\
\Rightarrow (dz_0 + dz_1) \wedge J(dz_1 - dz_0) &= (dz_0 + dz_1) \wedge \frac{1}{2}\Delta t H_{zz}(\bar{z})(dz_0 + dz_1) \\
\Rightarrow (dz_0 + dz_1) \wedge J(dz_1 - dz_0) &= 0 \\
\Rightarrow dz_1 \wedge Jdz_1 - dz_0 \wedge Jdz_0 &= 0
\end{aligned}$$

□

This is an instance of the following more general result.

Proposition 3. *Any solutions of any symplectic Runge–Kutta method applied to $J\dot{z} = \nabla H$ preserve the 2-form $\frac{1}{2}dz \wedge Jdz$, where J is any constant antisymmetric matrix.*

Proof. The s stage symplectic Runge-Kutta method is

$$JZ_i = Jz_0 + \Delta t \sum_{j=1}^{s} a_{ij} JF_j \quad (12)$$

$$Jz_1 = Jz_0 + \Delta t \sum_{j=1}^{s} b_j JF_j, \quad (13)$$

where

$$JF_j = \nabla H(Z_j) \quad (14)$$

The coefficients of a symplectic Runge–Kutta method obey

$$b_i b_j - b_j a_{ji} - b_i a_{ij} = 0. \quad (15)$$

Taking the exterior product of equations (12), (13), and (14) gives

$$Jdz_0 = JdZ_i - \Delta t \sum_{j=1}^{s} a_{ij} JdF_j \quad (16)$$

$$Jdz_1 = Jdz_0 - \Delta t \sum_{j=1}^{s} b_j JdF_j \quad (17)$$

$$JdF_j = H_{zz} dZ_j \quad (18)$$

From equation (18),

$$dZ_j \wedge JdF_j = dZ_j \wedge H_{zz} dZ_j = 0 \quad (19)$$

Then

$$
\begin{aligned}
&dz_1 \wedge J dz_1 - dz_0 \wedge J dz_0 \\
&\quad = dz_1 \wedge J(dz_0 + \Delta t \sum_{j=1}^{s} b_j dF_j) - dz_0 \wedge J dz_0 \text{ (using (17))} \\
&\quad = -J dz_1 \wedge \left(dz_0 + \Delta t \sum_{j=1}^{s} b_j dF_j \right) - dz_0 \wedge J dz_0 \\
&\quad = -\left(J dz_0 + \Delta t \sum_{j=1}^{s} b_j J dF_j \right) \wedge \left(dz_0 + \Delta t \sum_{j=1}^{s} b_j dF_j \right) - dz_0 \wedge J dz_0 \\
&\quad = \Delta t dz_0 \wedge J \sum_{j=1}^{s} b_j dF_j + \Delta t \sum_{j=1}^{s} b_j dF_j \wedge J dz_0 + \Delta t^2 \sum_{j=1}^{s} b_j dF_j \wedge J \sum_{j=1}^{s} b_j dF_j \\
&\quad = -\Delta t \sum_{j=1}^{s} b_j J dz_0 \wedge dF_j + \Delta t \sum_{i=1}^{s} b_i dF_i \wedge J dz_0 + \Delta t^2 \sum_{j=1}^{s} b_j dF_j \wedge J \sum_{j=1}^{s} b_j dF_j \\
&\quad = -\Delta t \sum_{j=1}^{s} b_j J \left(dZ_j - \Delta t \sum_{i=1}^{s} a_{ji} dF_i \right) \wedge dF_j \text{ (using (16))} \\
&\qquad + \Delta t \sum_{i=1}^{s} b_i dF_i \wedge J \left(dZ_i - \Delta t \sum_{j=1}^{s} a_{ij} dF_j \right) \text{ (using (16))} \\
&\qquad + \Delta t^2 \sum_{j=1}^{s} b_j dF_j \wedge J \sum_{j=1}^{s} b_j dF_j \\
&\quad = -\Delta t^2 \left[\sum_{j,i=1}^{s,s} b_j a_{ji} dF_i \wedge J dF_j + \sum_{i,j=1}^{s,s} b_i a_{ij} dF_i \wedge J dF_j \right] \text{ (using (19))} \\
&\qquad + \Delta t^2 \sum_{j=1}^{s} b_j dF_j \wedge J \sum_{j=1}^{s} b_j dF_j \\
&\quad = \Delta t^2 \left[\sum_{i,j=1}^{s,s} (b_i b_j - b_j a_{ji} - b_i a_{ij}) \, dF_i \wedge J dF_j \right] \\
&\quad = 0 \text{ (using (15))}
\end{aligned}
$$

□

A full study of the geometry of the relations (z_0, z_1) generated in Proposition 3 remains to be undertaken.[3] Unfortunately, the relations (z_0, z_1) in Proposition 3 do not yield good integrators for arbitrary J and H. For example, holonomic constraints can be specified as generalized Hamiltonian systems with $H = \tilde{H}(q,p) + \sum_{i=1}^{k} \lambda_i h_i(q)$. In this case the midpoint rule, say, generates maps from *all* (q_0, p_0) to (q_1, p_1) with the constraints satisfied at the midpoint. Not only is the phase space 'wrong', this method is known to be not

[3] The relations generated in Proposition 3 are a generalization of the Viterbo generating functions used in symplectic topology [7]. These take the form $S\colon Q \times \mathbb{R}^k \to \mathbb{R}$; the submanifold $p = S_q(q,\lambda)$, $0 = S_\lambda(q,\lambda)$ is Lagrangian in T^*Q. The parameters λ allow the representation of larger classes of Lagrangian submanfolds than the standard generating function $S(q)$ which generates $p = S_q(q)$ which is necessarily a graph over Q.

convergent in general [2]. The situation is much better for the index one constraints of Proposition 1.

Proposition 4. *For the index 1 constrained problem of Proposition 1, Proposition 3 yields integrators that are well-defined for sufficiently small Δt, convergent of the same order as the Runge–Kutta method, preserve the constraint submanifold, and preserve the symplectic form on the constraint submanifold.*

Proof. In this case the constraint part of the Runge–Kutta equations read

$$0 = \nabla_\lambda H(Q_i, P_i, \Lambda_i),\ i = 1, \dots, k,$$
$$0\frac{\lambda_1 - \lambda_0}{\Delta t} = \sum_{i=1}^{s} b_i \nabla_\lambda H(Q_i, P_i, \Lambda_i)$$

Therefore the Lagrange multipliers Λ_i at each stage are given by the *exact* Lagrange multipliers evaluated at (Q_i, P_i), i.e. $\Lambda_i = \tilde{\lambda}(Q_i, P_i)$, and λ_1 is arbitrary. For convenience, we add the extra equations $\lambda_0 = \tilde{\lambda}(q_0, p_0)$, $\lambda_1 = \tilde{\lambda}(q_1, p_1)$ which do not affect the method at all. The resulting method is equivalent to that obtained by eliminating the Lagrange multipliers in the Hamiltonian, applying a symplectic Runge–Kutta method, and lifting back to the constraint manifold by $\lambda = \tilde{\lambda}(q, p)$. It is therefore well defined for sufficiently small Δt and convergent of the same order as the Runge–Kutta method. Because $\frac{1}{2} dz \wedge J dz = dq \wedge dp$, the symplectic form $dq \wedge dp$ is preserved on the constraint manifold. □

3 General constraints

Under certain conditions, namely that the Legendre transform that defines the conjugate momenta must be invertible to give $\dot{q}$, Proposition 1 can be generalized to allow a general Lagrangian and general constraints. A very thorough geometric treatment of this type of constraint, applying the Gotay–Nestor geometric version of the Dirac–Bergmann constraint algorithm, can be found in [4].

Proposition 5. *If the Legendre transform mapping $(\dot{q}, q, \lambda) \to (p, q, \lambda)$ given in equation (21) is invertible then the Euler–Lagrange equations for the action*

$$S(q) = \int_{t_0}^{t_1} L(t, q, \dot{q}) dt$$

subject to the constraints $g_i(q, \dot{q}) = 0, i = 1, \dots, k$ are equivalent to the generalized Hamiltonian system

$$J\dot{z} = \nabla H(z) \tag{20}$$

where

$$
\begin{aligned}
J &= \begin{pmatrix} 0 & -\mathbb{I}_{n\times n} & 0 \\ \mathbb{I}_{n\times n} & 0 & 0 \\ 0 & 0 & 0_{k\times k} \end{pmatrix} \\
z &= \begin{pmatrix} q \\ p \\ \lambda \end{pmatrix} \\
p &= \nabla_{\dot q} F(\dot q, q, \lambda) && (21) \\
H(z) &= \dot q \cdot p - F(\dot q, q, \lambda) \\
F(\dot q, q, \lambda) &= L(t, q, \dot q) - \sum_{i=1}^{k} \lambda_i g_i(q, \dot q) \\
& && (22)
\end{aligned}
$$

(Here the Lagrange multipliers in the $(q, \dot q)$ formulation are determined by (23)–(25) below.) If, in addition, the matrix $G(q, \dot q)$ given by $G_{ij} = \partial g_i(q, \dot q)/\partial \dot q_j$ has full rank k for all q, $\dot q$ then the system of Eq. (20) has index one, i.e., can be solved for $\lambda = \tilde\lambda(q, p)$.

Proof. The Euler-Lagrange equations are

$$\frac{\partial}{\partial t}\left(\nabla_{\dot q} F\right) - \nabla_q F = 0, \tag{23}$$

$$g_i(q, \dot q) = 0, \quad i = 1, \ldots, k, \tag{24}$$

where

$$F(\dot q, q, \lambda) = L(t, q, \dot q) - \sum_{i=1}^{k} \lambda_i g_i(q, \dot q). \tag{25}$$

Define the conjugate momentum p as

$$p := \nabla_{\dot q} F. \tag{26}$$

By assumption equation (26) can be rearranged to give $\dot q$

$$\dot q = f(q, p, \lambda). \tag{27}$$

Using equation (26) in equation (23) gives the expression for $\dot p$

$$\dot p = \nabla_q F. \tag{28}$$

Define the Hamiltonian as

$$
\begin{aligned}
H(q, p, \lambda) &:= \dot q \cdot p - F(\dot q, q, \lambda) && (29) \\
&= p \cdot f(q, p, \lambda) - F(\dot q, q, \lambda) \qquad \text{using (27)} && (30)
\end{aligned}
$$

Then (where f_p is the Jacobian derivative $\partial f_i(q,p,\lambda)/\partial p_j$, etc.) we have

$$\begin{aligned}
\nabla_p H &= f + f_p p - \nabla_p F \\
&= f + f_p p - f_p \nabla_{\dot q} F \\
&= f + f_p p - f_p p && \text{using (26)} \\
&= \dot q && \text{using (27)} \qquad (31) \\
\nabla_q H &= f_q p - f_q \nabla_{\dot q} F - \nabla_q F \\
&= f_q p - f_q p - \nabla_q F && \text{using (26)} \\
&= -\dot p && \text{using (28)} \qquad (32) \\
\nabla_\lambda H &= f_\lambda p - f_\lambda \nabla_{\dot q} F + g \\
&= f_\lambda p - f_\lambda p + g && \text{using (26)} \\
&= 0 && \text{using (24)} \qquad (33)
\end{aligned}$$

establishing the equivalence of the Euler–Lagrange equations to (20). The constraints in the Hamiltonian formulation are

$$0 = \nabla H(q,p,\lambda) = g(q, f(q,p,\lambda))$$

and their Jacobian derivative with respect to λ is the matrix

$$\sum_{j=1}^{n} \frac{\partial g_i}{\partial \dot q_j} \frac{\partial \dot q_j}{\partial \lambda_k}.$$

The first factor is G. The second factor is the derivative of the inverse Legendre transform (assumed invertible) with respect to λ. The forward Legendre transform is $p = \frac{\partial L}{\partial \dot q} + \sum_{i=1}^k \lambda_i \frac{\partial g_i}{\partial \dot q}$ and its derivative with respect to λ is G^T. By the chain rule, if G has rank k then the Jacobian is nonsingular for all q, $\dot q$ and the constraints have a unique solution $\lambda = \tilde\lambda(q,p)$ for all q,p, that is, the system has index one. □

The integrators we considered in Section 2 work for any index one system.

Proposition 6. *For any constant antisymmetric J, if the generalized Hamiltonian system $J\dot z = \nabla H(z)$ can be solved for $\lambda = \lambda(q,p)$ where $z = (q,p,\lambda)$ are Darboux coordinates for J, then any symplectic Runge–Kutta method applied to this systems yields constrained symplectic integrators convergent with their classical order.*

Note that the assumptions are satisfied if $|H_{\lambda\lambda}| \neq 0$. The constraints may be nonlinear in λ, and need not be solved analytically; the entire Runge–Kutta system for (Q_i, P_i, Λ_i) can be numerically solved simultaneously.

Proposition 5 can be generalized further, to any singular Lagrangian $L(q,\dot q,\lambda)$, and still further to Lagrangians $L(z,\dot z)$ where $|L_{\dot z\dot z}| = 0$, but the required nondegeneracy assumptions are not as geometrically transparent as those in Proposition 5.

4 Sub-Riemannian with holonomic constraints

Proposition 1 allowed a variational problem with subRiemannian constraints to be converted into an unconstrained Hamiltonian system that can be solved using the symplectic midpoint rule. The proposition in this section shows that if *holonomic* constraints are added to the original variational problem, then the resulting Hamiltonian system is a simple *holonomically* constrained system. This system can be solved by a symplectic method such as RATTLE [6].

Proposition 7. *Let M be a symmetric nonsingular $n \times n$ mass matrix, $V : \mathbb{R}^n \to \mathbb{R}$ a smooth potential, $g_i : \mathbb{R}^n \to \mathbb{R}^n$, $i = 1, \ldots, k$ be k smooth functions , and q be a smooth extremal with fixed endpoints for the functional*

$$S(q) = \int_{t_0}^{t_1} L(t, q, \dot{q}) dt = \int_{t_0}^{t_1} \left(\frac{1}{2} \dot{q}^T M \dot{q} - V(q) \right) dt \tag{34}$$

subject to the velocity *constraints $g_i(q) \cdot \dot{q} = 0, i = 1, \ldots, k$ and the* holonomic *constraints $h_i(q) = 0, i = 1, \ldots, l$. Then*

$$J\dot{z} = \nabla H(z) \tag{35}$$

where

$$J = \begin{pmatrix} 0 & -\mathbb{I}_{n\times n} & 0 & 0 \\ \mathbb{I}_{n\times n} & 0 & 0 & 0 \\ 0 & 0 & 0_{k\times k} & 0 \\ 0 & 0 & 0 & 0_{l\times l} \end{pmatrix}, \quad z = \begin{pmatrix} q \\ p \\ \lambda \\ \lambda^h \end{pmatrix}$$

$$p = M\dot{q} - \sum_{i=1}^{k} \lambda_i g_i(q)$$

$$H(z) = \frac{1}{2} \left(p + \sum_{i=1}^{k} \lambda_i g_i(q) \right)^T M^{-1} \left(p + \sum_{i=1}^{k} \lambda_i g_i(q) \right) + V(q) + \sum_{i=1}^{l} \lambda_i^h h_i(q)$$

and, furthermore, the Euler–Lagrange equations for (34) are equivalent to the generalized Hamiltonian system (35). If, in addition, the velocity constraints are linearly independent for all q, then Eq. (35) is equivalent to a canonical holonomically constrained Hamiltonian system.

Proof. As in Proposition 1 the extended Lagrangian F, the conjugate momenta p, and the Hamiltonian $H(q, p, \lambda, \lambda^h)$ are defined by

$$\begin{aligned} F &:= \frac{1}{2} \dot{q}^T M \dot{q} - V(q) - \sum_{i=1}^{k} \lambda_i g_i(q) \cdot \dot{q} - \sum_{i=1}^{l} \lambda_i^h h_i(q), \\ p &:= \nabla_{\dot{q}} F = M\dot{q} - \sum_{i=1}^{k} \lambda_i g_i(q), \\ H &:= \dot{q} \cdot p - F. \end{aligned}$$

The rest of the proof is a calculation along the same lines as for Proposition 1. □

Proposition 8. *Subject to standard nondegeneracy assumptions on the Hamiltonian, the following algorithm yields a convergent, second order integrator that is symplectic on the constraint manifold defined by the (primary) holonomic constraints and the secondary constraints induced by them: (i) apply* RATTLE *using the holonomic constraints; (ii) in the inner step of* RATTLE*, when a time step of the unconstrained system is required, apply the midpoint rule to the generalized Hamiltonian system with Hamiltonian $H(q, p, \lambda, 0)$.*

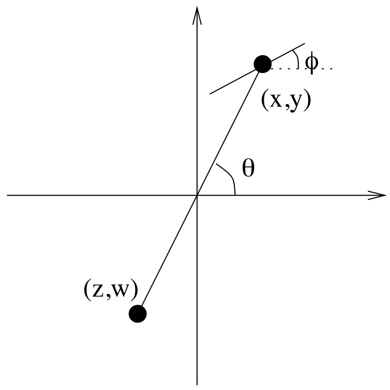

Figure 1: A two wheeled vehicle with the front wheel at an angle ϕ and the entire vehicle on an angle of θ

Proof. Eliminating the velocity constraints by solving for the Lagrange multipliers yields a standard holonomically constrained system. Applying RATTLE (with the midpoint rule in the inner step) to this system yields a convergent second order integrator on the constraint surface. Applying the midpoint rule in the inner step is equivalent to applying the midpoint rule to the generalized Hamiltonian system with Hamiltonian $H(q, p, \lambda, 0)$. □

5 Example: Sub-Riemannian geodesics

The motion of a two-wheeled vehicle with a front steering wheel and a non-steering back wheel, moving on a smooth surface, will be modelled. We consider the two wheeled vehicle shown in Fig. 1 with length L, back wheel at (z, w), and front wheel at (x, y). The front wheel is at an angle ϕ and the vehicle is at an angle θ.

If the speed of the front wheel is v, its velocity of the front wheel must obey

$$\begin{aligned} \dot{x} &= v \cos\phi \\ \dot{y} &= v \sin\phi \end{aligned}$$

Eliminating v, the velocity of the front wheel obeys the constraint

$$\dot{x} \sin\phi - \dot{y} \cos\phi = 0 \tag{36}$$

Similarly, the velocity of the back wheel obeys the constraint

$$\dot{z} \sin\theta - \dot{w} \cos\theta = 0 \tag{37}$$

We can eliminate equation (37), and thus the variables z and w, using the distance between the two wheels which relates the four variables. Notice that

$$\begin{aligned} x - z &= L \cos\theta \\ y - w &= L \sin\theta \end{aligned}$$

which after taking time derivatives gives

$$\begin{aligned} \dot{z} &= \dot{x} + L\dot{\theta} \sin\theta \\ \dot{w} &= \dot{y} - L\dot{\theta} \cos\theta \end{aligned}$$

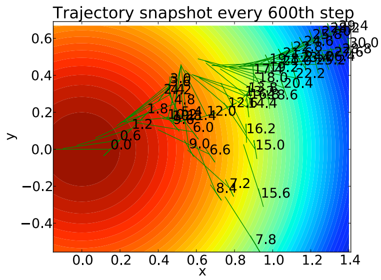

Figure 2: Snapshots at every 600^{th} step of the bicycle starting at $(x, y, \theta, \phi, p_x, p_y, p_\theta, p_\phi) = (0.15, 0, \pi/4, 0, 1, 0, 0)$, with $\Delta t = 0.001$, and travelling in a potential $V(\mathbf{q}) = 1 + \cos(r)$. At about $t = 7.8$ and $t = 15.6$ the steering wheel is aligned with the bicycle and the bicycle changes direction and retreats rapidly.

which substituted into equation (37) gives

$$\dot{x} \sin\theta - \dot{y} \cos\theta + L\dot{\theta} = 0 \tag{38}$$

The constraints given in equations (36) and (38) can be written as

$$g_i(q) \cdot \dot{q} = 0, i = 1, 2$$

where $q = (x, y, \theta, \phi)^T$ and

$$g_1(q) = (\sin\phi, -\cos\phi, 0, 0)^T \tag{39}$$

$$g_2(q) = (\sin\theta, -\cos\theta, L, 0)^T \tag{40}$$

We take the Lagrangian to be

$$L = \frac{1}{2}\left(\dot{x}^2 + \dot{y}^2 + \alpha\dot{\theta}^2 + \beta\dot{\phi}^2\right) - V(x, y) \tag{41}$$

where the potential $V(x, y)$ is the (scaled) height of the surface, giving an index one system as in Proposition 1. That is, we are calculating geodesics (in the case $V = 0$) of the subRiemannian metric defined by Eqs. (41), (36) and (38). We use the midpoint rule.

The first four tests use the potential $V(q) = -\cos r$, where r is the midpoint of the vehicle. The first test is to compute a simple trajectory of a bicycle of length 0.3, starting with the centroid of the bicycle at the origin. The bicycle is given a small push by setting the initial generalized momenta $p_y \neq 0$. The initial conditions are $(x, y, \theta, \phi, p_x, p_y, p_\theta, p_\phi) = (0.15, 0, \pi/4, 0, 1, 0, 0)$. Fig. 2 shows the bicycle at regular snapshots through time. The background colours show the potential: red is high, and blue is low. Note that that at about $t = 7.8$ and $t = 15.6$ the steering wheel is aligned with the bicycle and the bicycle changes direction and retreats rapidly.

The second test is to check the order of the method by plotting, in Fig. 3, the error of various solutions. A reference trajectory with $\Delta t = 10^{-4}$ and final time $10s$ is computed, and trajectories with bigger time steps are compared to it. The slope of the error line shows that the method is second order. A comparison to a highly accurate reference solution calculated with MATLAB's `ode15s` and tolerance 10^{-12} also shows the second order accuracy of our method (Fig. 4).

Our third test is to numerically test the symplectic condition by evaluating the symplectic bilinear form for a number of steps. See Fig. 5. The reference trajectory is the one used in

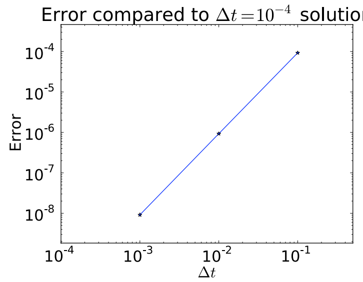

Figure 3: The error of three runs compared to a reference solution with $\Delta t = 1 \times 10^{-4}$. The initial conditions were $(x, y, \theta, \phi, p_x, p_y, p_\theta, p_\phi) = (0.15, 0, 0, \frac{\pi}{4}, 0, 1, 0, 0)$. The method is order Δt^2

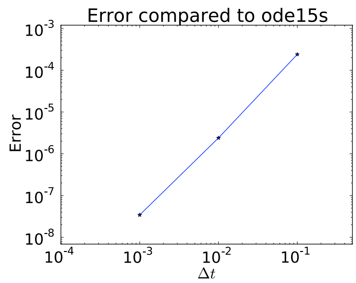

Figure 4: Error compared to MATLAB `ode15s` solution. Initial conditions were $(x, y, \theta, \phi, p_x, p_y, p_\theta, p_\phi) = (0.3, 0, 0, \pi, 0, 1.09, 1)$, with $\boldsymbol{\lambda} = (0, -0, 3)$, final time $10s$, and travelling in no potential field. The solution agrees with the MATLAB solution.

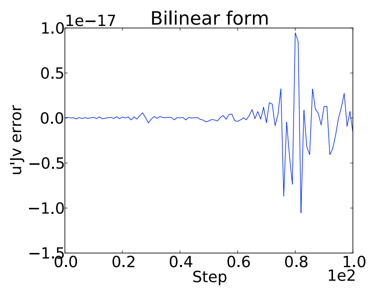

Figure 5: The symplectic bilinear form error evaluated for 100 steps with $\Delta t = 0.1$. The trajectories $\phi_{z_i}(n)$ for three nearby initial conditions, z_0, z_1, and z_2, were calculated. The change in the symplectic form was estimated as $u_n^T J v_n - u_0^T J v_0$ where $u_n = \phi_{z_1}(n) - \phi_{z_0}(n)$ and $v_n = \phi_{z_2}(n) - \phi_{z_0}(n)$. The initial points were $z_0 = (0.15, 0, 0, \frac{\pi}{4}, 0, 1, 0, 0)$, $z_1 = (0.1500000085, 1 \times 10^{-8}, -1 \times 10^{-8}, 0.785398145543, -1 \times 10^{-8}, 0.99999998, 1 \times 10^{-8}, -1 \times 10^{-8})$ and $z_2 = (0.1500000115, -1 \times 10^{-8}, 1 \times 10^{-8}, 0.785398181251, 1 \times 10^{-8}, 0.99999998, -1 \times 10^{-8}, -1 \times 10^{-8})$.

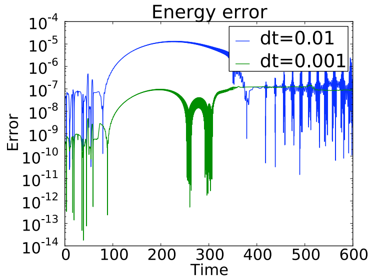

Figure 6: The energy error over time. This is the energy at each step minus the initial energy. The bicycle is trapped in the potential bowl $-V(\mathbf{q})$, and the energy error does not show a linear growth in time.

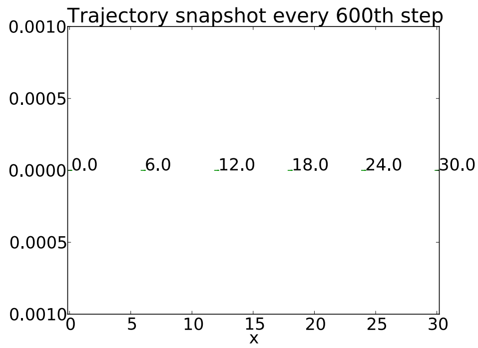

Figure 7: Snapshots at every 600$^{\text{th}}$ step of the bicycle starting at $(x, y, \theta, \phi, p_x, p_y, p_\theta, p_\phi) = (0.15, 0, 0, 1, 0, 0, 0)$, with $\Delta t = 0.01$, and travelling in no potential field. The bicycle stays straight. The trajectory is unstable and eventually wanders from a straight path.

the error plot with with $\Delta t = 0.1$. Two nearby trajectories are also calculated to allow an estimate of the change in $dq \wedge dp$. The calculated change is near roundoff indicating that the integrator is symplectic.

The fourth test is to plot, in Fig. 6, the energy error for $\Delta t = 0.01$ and $\Delta t = 0.001$. From the figure the energy errors appear to be bounded, as expected for a symplectic integrator.

The free motion case ($V(q) = 0$) has two simple solutions that are relative equilibria for the translation and rotation symmetries of the problem, namely straight line and circular motion. The first trajectory to check is a straight line. If the bicycle starts with $\theta = \phi = 0$, and there is no potential field, then bicycle should remain travelling in a straight line.

Let $\theta = \phi = 0$, $\dot{x} = 1$, and $\dot{y} = 0$. The constraints in equations (36) and (38) are satisfied. Equation (8) gives the initial generalized momenta values: all are zero except $p_x = 1$. In Fig. 7 this simple trajectory of the bicycle is confirmed.

For the circle, let $\theta = at$, $\phi = at + \frac{\pi}{2}$, $\dot{x} = -c\sin(\theta)$, and $\dot{y} = c\cos(\theta)$. There are two constants, a and c, to be determined. Equation (36) gives $\lambda = (1, -c)$, and equation (38) gives $aL = c$. Using these values in equation (8) gives the initial generalized momenta values: $(p_x, p_y, p_\theta, p_\phi) = (0, 0, a(1 + L^2), a)$. For this simple trajectory a is chosen to be 1. In Fig. 8 the circle trajectory of the bicycle is confirmed. If the trajectory is computed for larger times the bicycle leaves the circle; the solution appears to be unstable, but, interestingly, appears to repeatedly return to the circular orbit, indicating a possible relative homoclinic structure of this problem. This is also suggested from the evolution of

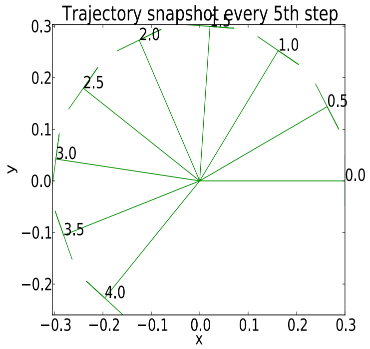

Figure 8: Snapshots every 5^{th} step of the bicycle starting at $(x, y, \theta, \phi, p_x, p_y, p_\theta, p_\phi) = (0.3, 0, 0, \pi, 0, 1.09, 1)$, with $\Delta t = 0.1$, and travelling in no potential field. The bicycle stays in a circle for many revolutions (not shown for clarity), but the trajectory is not stable, so eventually wanders.

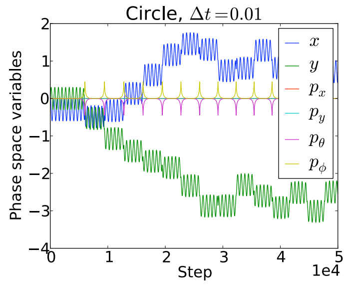

Figure 9: Some of the phase space variables for the bicycle trajectory starting in a circle. The solution for p_θ and p_ϕ suggest a relative homoclinic orbit.

the Lagrange multipliers, $\lambda - (1, -c)$ being shown in Fig. 11.

6 Example: the Heisenberg problem

A previous study of geometric integrators for subRiemannian variational problems used a discrete variational approach to obtain constrained symplectic integrators [1]. Our approach, applying symplectic integrators to the Hamiltonian formulation, yields geometric integrators with the same geometric properties, but uses standard integrators that allow any order with standard implementations, and does not require an approximation of $\dot{q}$, that is, it naturally yields first-order trajectories in (q, p) instead of second-order trajectories in q.

We repeat the numerical illustration of [1, pg. 12], the Heisenberg problem, using our approach. This is to find the extremal $q(t) = (x(t), y(t), z(t))$ of

$$S(q) = \int_{t_0}^{t_1} L(t, q, \dot{q}) dt = \int_{t_0}^{t_1} \left(\frac{1}{2} \dot{q}^T \dot{q} - V(q) \right) dt$$

subject to the constraint $g(q) \cdot \dot{q} = 0$, where $g(q) = (-y, x, 1)$.

Equation (8) gives $\dot{q}$

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} + \lambda \begin{pmatrix} -y \\ x \\ 1 \end{pmatrix} \tag{42}$$

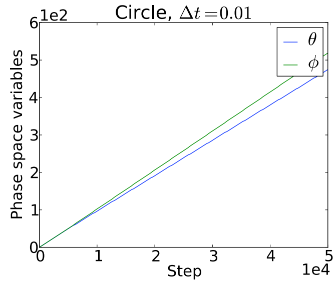

Figure 10: Some of the phase space variables for the bicycle trajectory starting in a circle. These angle variables are monotonically increasing.

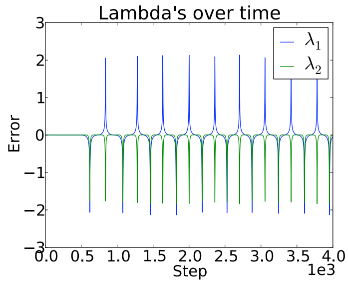

Figure 11: The error in the λ's for the circle trajectory. Note that there are many more steps than shown in Fig. 8. This shows that the λ's periodically return to their values for a circle.

Using equation (9) the $\dot{p}$ can be written

$$\begin{pmatrix} \dot{p_x} \\ \dot{p_y} \\ \dot{p_z} \end{pmatrix} = -\nabla V(q) - \lambda \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \left[\begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} + \lambda \begin{pmatrix} -y \\ x \\ 1 \end{pmatrix} \right] \tag{43}$$

and we have the constraint $g \cdot (p + \lambda g) = 0$, which gives

$$\lambda = -\frac{g \cdot p}{g \cdot g}$$

A simple trajectory starting with the same initial conditions as in [1, pg. 15] is shown in Fig. 12. Their initial conditions are $(x, y, z, \dot{x}, \dot{y}, \dot{z}, \lambda) = (0, 0, 0, 0.1, 0.3, 0, 1)$, which when converted to generalized momenta variables are $(x, y, z, p_x, p_y, p_z, \lambda) = (0, 0, 0, 0.1, 0.3, 1, 1)$. Qualitatively the results look like [1, pg.14].

7 Discussion

We have constructed symplectic integrators for a different class of constrained Hamiltonian systems than the holonomic constraints most commonly considered in the literature. The class includes important practical problems arising in subRiemannian geometry. We have restricted our attention to symplectic Runge–Kutta methods; a generalization to partitioned methods in which different Runge–Kutta coefficients are used for q, for p, and for λ is straightforward. In other work [5], we reinterpret these methods as an instance of RATTLE in an extended phase space; that point of view also suggests different generalisations. However, the nondegeneracy conditions are essential for the method, indeed, for

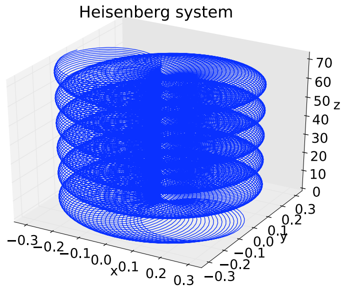

Figure 12: The Heisenberg example starting at $(x, y, z, \dot{x}, \dot{y}, \dot{z}) = (0, 0, 0, 0.1, 0.3, 0)$ or $(x, y, z, p_x, p_y, p_z, \lambda) = (0, 0, 0, 0.1, 0.3, 1, 1)$. Qualitatively the results look like [1, pg.14].

the entire approach, to work. It is not clear to what extent the approach can be extended to handle more general constraints, for example, to the system $J\dot{z} = \nabla H + \lambda \nabla g$, where the constraint submanifold $g(z) = 0$ is symplectic. No symplectic, constraint-preserving method is known for this problem. As remarked before Proposition 4, a full study of the geometry of the relations (z_0, z_1) generated in Proposition 3 remains to be undertaken. Any solutions are symplectic, so this gives access to a much larger class of symplectic maps than do traditional generating functions. Note that new variables (analogous to λ) can be added as needed to generate larger classes of maps.

Acknowledgements

O. Verdier would like to acknowledge the support of the GeNuIn Project, funded by the Research Council of Norway, as well as the hospitality of the Institute for Fundamental Sciences of Massey University, New Zealand, where some of this research was conducted. K. Modin would like to thank the Department of Mathematics at NTNU in Trondheim and the Swedish Research Council for support. This research was supported by the CRiSP Project, funded by the European Commission's Seventh Framework Programme, and by the Marsden Fund of the Royal Society of New Zealand.

References

[1] Roberto Benito and David Martin de Diego. Discrete vakonomic mechanics. *Journal of Mathematical Physics*, 46(8):083521, 2005. 16, 17, 18

[2] E. Hairer and L. O. Jay. Implicit Runge–Kutta methods for higher index differential-algebraic systems. In *Contributions in Numerical Mathematics*, volume 2, pages 213–224, River Edge, N.J., 1993. World Scientific. 8

[3] B. Leimkuhler and S. Reich. *Simulating Hamiltonian dynamics*. Cambridge monographs on applied and computational mathematics. Cambridge University Press, 2004. 2, 5

[4] Sonia Martínez, Jorge Cortés, and Manual de León. The geometrical theory of constraints applied to the dynamics of vakonomic mechanical systems: The vakonomic bracket. *Journal of Mathematical Physics*, 41(4):2090–2120, 2000. 8

[5] R. I. McLachlan, K. Modin, O. Verdier, and M. C. Wilkins. Geometric generalisations of SHAKE and RATTLE. *J. Foundations Comput. Math.*, 2012. submitted. 2, 17

[6] Robert D. Skeel, Benedict J. Leimkuhler, and Benedict J. Leimkuhler. Symplectic numerical integrators in constrained Hamiltonian systems. *Journal of Computational Physics*, 112:117–125, 1994. 10

[7] C. Viterbo. Generating functions, symplectic geometry, and applications. In *Proceedings of the International Congress of Mathematicians*, pages 537–547, Basel, 1995. Birkhauser. 7