Variational and linearly implicit integrators, with applications

MOLEI TAO∗
School of Mathematics, Georgia Institute of Technology, Atlanta, GA, USA
∗Corresponding author: mtao@gatech.edu

AND

HOUMAN OWHADI
Applied & Computational Mathematics, California Institute of Technology, Pasadena, CA, USA
and Control & Dynamical Systems, California Institute of Technology, Pasadena, CA, USA

[Received on 16 January 2013; revised on 25 August 2014]

We show that symplectic and linearly implicit integrators proposed by Zhang & Skeel (1997, Cheap implicit symplectic integrators. Appl. Numer. Math., 25, 297–302) are variational linearizations of Newmark methods. When used in conjunction with penalty methods (i.e., methods that replace constraints by stiff potentials), these integrators permit coarse time-stepping of holonomically constrained mechanical systems and bypass the resolution of nonlinear systems. Although penalty methods are widely employed, an explicit link to Lagrange multiplier approaches appears to be lacking; such a link is now provided (in the context of two-scale flow convergence (Tao, M., Owhadi, H. & Marsden, J. E. (2010) Nonintrusive and structure-preserving multiscale integration of stiff ODEs, SDEs and Hamiltonian systems with hidden slow dynamics via flow averaging. Multiscale Model. Simul., 8, 1269–1324). The variational formulation also allows efficient simulations of mechanical systems on Lie groups.

Keywords: variational integrators; symplectic and linearly implicit; coarse time-stepping; constrained dynamics; mechanical system on Lie group.

1. Introduction and main results

Integrators: Symplectic integrators are popular for simulating mechanical systems due to their structure-preserving properties (e.g., Hairer et al., 2006). Implicit methods, on the other hand, allow accurate coarse time-stepping of a class of stiff or multiscale problems (e.g., Li et al., 2008; Filbet & Jin, 2010). It is also a classical treatment to linearize implicit methods so that expensive nonlinear solves can be avoided (e.g., Beam & Warming, 1976). Although linearizations of most implicit symplectic methods are not symplectic, Zhang and Skeel found a family of symplectic and linearly implicit integrators (Zhang & Skeel, 1997), which allows efficient and structure-preserving simulations. We show that their method is not only symplectic, but in fact variational.

Specifically, consider mechanical systems governed by Newton’s equation:

\[ \dot{x} = v, \quad M\dot{v} = -\nabla V(x), \]  

where \( V \in C^2(\mathbb{R}^n) \) and \( M \) is an \( n \times n \) symmetric, positive-definite constant matrix.
If we consider the following discrete Lagrangian (see Section 2.1 for explanations):

\[
\tilde{L}_d(x_k, x_{k+1}, a_k, a_{k+1}) = h \left( \frac{1}{2} \left( \frac{x_{k+1} - x_k}{h} \right)^T M \left( \frac{x_{k+1} - x_k}{h} \right) \right)
- \frac{1}{2} \left( \beta h^2 \frac{1}{2} a_k^T M a_k + V(x_k) + \beta h^2 a_k^T \nabla V(x_k) + \frac{1}{2} \beta^2 h^4 a_k^T \text{Hess } V(x_k) a_k \right)
- \frac{1}{2} \left( \beta h^2 \frac{1}{2} a_{k+1}^T M a_{k+1} + V(x_{k+1}) + \beta h^2 a_{k+1}^T \nabla V(x_{k+1}) \right)
+ \frac{1}{2} \beta^2 h^4 a_{k+1}^T \text{Hess } V(x_{k+1}) a_{k+1} \right),
\]

(1.2)

then the Euler–Lagrange equation of the variational principle

\[
\delta \sum_{k=1}^{N} \tilde{L}_d(x_k, x_{k+1}, a_k, a_{k+1}) = 0
\]

yields the following symplectic method (originally stated in Zhang & Skeel, 1997, Section 3).

**Integrator 1.1** **Zhang and Skeel’s symplectic method (Z&S):**

\[
\begin{align*}
x_{k+1} &= x_k + hv_k + \frac{1}{2} h^2 f_k, \\
v_{k+1} &= v_k + \frac{1}{2} h(f_k + f_{k+1}), \\
a_k &= -M^{-1} \nabla V(x_k) - M^{-1} \text{Hess } V(x_k) \beta h^2 a_k, \\
f_k &= a_k - \frac{1}{2} \beta^2 h^4 M^{-1} a_k \cdot V^{(3)}(x_k) \cdot a_k, 
\end{align*}
\]

(1.3)

where \(V^{(3)}(\cdot)\) is a third-order tensor corresponding to the third-derivative of \(V\), the symbol \(\cdot\) stands for tensor contraction and therefore \(a_k \cdot V^{(3)}(x_k) \cdot a_k\) is again a vector.

For computational efficiency, \(a_k\) should be obtained by solving a symmetric linear system instead of inverting a matrix. In this sense, Z&S is linearly implicit.

**Theorem 1.2** Z&S is:

1. unconditionally linearly stable if \(\beta \geq \frac{1}{4}\);
2. variational (and thus symplectic and conserving momentum maps);
3. second-order convergent (if stable) and can be made arbitrarily high-order convergent;
4. symmetric (‘time-reversible’).
Showing Z&S is variational ensures,\(^1\) due to a discrete Noether theorem (e.g., Marsden & West, 2001; Hairer et al., 2006), that it also preserves momentum maps that correspond to system symmetries. This additional preservation property is desired in mechanical system simulations. The variational formulation also leads to a possible extension to Lie groups (Section 4.5).

**Constrained dynamics:** One of our motivations for studying Z&S originates from a need for coarse time-steppings in penalty methods for constrained dynamics.

To model constrained dynamics, let

\[
S(q(t)) := \int_a^b \frac{1}{2} \dot{q}(t)^T M \ddot{q}(t) - V(q(t)) \, dt
\]

be the action associated with system (1.1). Under a holonomic constraint \(g(q) = 0\), the system evolution coincides with the critical trajectory on the constraint manifold, i.e., the solution of

\[
\frac{\delta S}{\delta q} = 0 \quad \text{and} \quad \text{for all } t, q(t) \in g^{-1}(0). \tag{1.5}
\]

This trajectory can also be obtained by solving the differential algebraic system

\[
\begin{cases}
M \ddot{q} = -\nabla V(q) + \lambda^T \nabla g(q), \\
g(q) = 0.
\end{cases}
\tag{1.6}
\]

Penalty methods approximate rigid constraints by stiff potentials; this is a classical idea and we refer the reader to Rubin & Ungar (1957), Takens (1980), Terzopoulos et al. (1987) and Platt & Barr (1988) for a noncomprehensive list of references. More precisely, modify the potential energy \(V(q)\) to \(V(q) + \frac{1}{2} \omega^2 g(q)^T g(q)\); then, the solution of (1.5) is approximated by the solution of the following unconstrained mechanical system:

\[
M \ddot{q}^\omega = -\nabla V(q^\omega) - \omega^2 g(q^\omega)^T \nabla g(q^\omega), \tag{1.7}
\]

where \(\omega\) is large enough. Paraphrasing Platt & Barr (1988), the problem is, ‘as a result of stiffness, the numerical differential equation solver takes very small time steps, using a large amount of computing time without getting much done’. Z&S alleviates this problem because it can use coarse time steps and does not solve nonlinear systems (see Section 1.2).

On a related matter, although penalty methods are widely employed and proved convergent to constrained dynamics (see Section 1.1), a quantitative analysis of its link to the Lagrange multiplier approach (1.6) appeared to be lacking. We show that the solution of (1.7) converges to that of (1.6) as \(\omega \to \infty\) in the sense of two-scale flow convergence (see Tao et al., 2010, Definition 1.1). More precisely, we have (explained in Section 3; throughout this paper, ‘bounded’ means having a norm bounded by an \(\omega\)-independent constant) the following theorem.

**Theorem 1.3** Denote by \(q^\omega(t)\) the solution to (1.7) with \(q^\omega(0) = q_0\) and \(\dot{q}^\omega(0) = \dot{q}_0\) (where \(g(q_0) = 0\) and \((d/dt)g(q_0) = \nabla g(q_0) \cdot \dot{q}_0 = 0\)). Suppose that \(M\) is nonsingular, \(V(\cdot)\) is bounded from below, \(V(q)\) diverges towards infinity as \(|q| \to \infty\), \(V(\cdot)\) and \(g(\cdot)\) are \(C^2\) with bounded derivatives, and for all

---

\(^1\) In general, symplectic methods are only locally variational, but variational methods are symplectic; see, for instance, Marsden & West (2001).
\( q \in g^{-1}(0) \), \( \nabla g(q) \) has a constant rank equal to the codimension of the constraint manifold; then,

\[
\lambda(t) := - \lim_{T \to 0} \lim_{\omega \to \infty} \frac{1}{T} \int_t^{t+T} \omega^2 g(q(\omega(s))) \, ds
\]

exists. Also, the solution \( q(t) \) of

\[
\begin{aligned}
M \ddot{q}(t) &= -\nabla V(q(t)) + \lambda(t)^T \nabla g(q(t)), \\
g(q(t)) &= 0, \quad q(0) = q_0, \quad \dot{q}(0) = \dot{q}_0
\end{aligned}
\]

exists and satisfies

\[
q^{(\omega)} F \rightarrow q
\]

in the sense of two-scale flow convergence (Tao et al., 2010), i.e., for all bounded \( t \geq 0 \) and all bounded and uniformly Lipschitz-continuous test function \( \varphi \),

\[
\lim_{T \to 0} \lim_{\omega \to \infty} \frac{1}{T} \int_t^{t+T} \varphi(q^{(\omega)}(s)) - \varphi(q(s)) \, ds = 0.
\]

Outline of the paper: Section 2 derives Z&S from a variational principle, relates it to Newmark integrators and discusses its properties. Section 3 illustrates how penalty methods converge to the Lagrange multiplier approach. Section 4 applies the method to constrained systems (pendula and water molecular dynamics), a nonconstrained model of DNA division and a mechanical system on \( SO(3) \), illustrating the benefits of a variational formulation.

1.1 On penalty methods

The penalty strategy of replacing holonomic constraints by stiff potentials is widely used. For example, it is a common treatment in computer graphics (e.g., Terzopoulos et al., 1987; Witkin et al., 1987; Platt & Barr, 1988).

It is known that the penalized solution converges to constrained dynamics in \( C^1 \) topology, as long as its initial condition is in the tangent bundle of the constraint manifold. We refer the reader to, e.g., the pioneering work of Rubin & Ungar (1957) and Takens (1980), to Bornemann & Schütte (1997), Bornemann (1998) and Shatah & Zeng (2002) for recent progress, and to Hairer et al. (2006, Chapter XIV.3) for a review.

The reverse point of view has also been employed, particularly in molecular dynamics, where stiff oscillatory molecular bonds are replaced by rigid constraints for the purpose of allowing larger time-steps (e.g., Fixman, 1974; Schlick, 2010). If the initial velocity is not in the tangent plane, then a correction potential might also be required to account for the nonzero normal energy (e.g., Fixman, 1974; Reich, 1995; Schütte & Bornemann, 1997). The Fixman potential (Fixman, 1974) is a classical example of such a correction, in particular when investigating thermodynamic properties of molecular systems (see, e.g., Bajars et al., 2011); on the other hand, Bornemann & Schütte (1995) suggest that Fixman might not be the right correction for deterministic systems.
1.2 One constrained dynamics

Other popular constrained dynamics methods include: generalized coordinates on the constraint manifold (e.g., Jain et al., 1993) and Lagrange multipliers (e.g., SHAKE (Ryckaert et al., 1977), RATTLE (Andersen, 1983), SETTLE (Hess et al., 1997), LINCS (Miyamoto & Kollman, 1992) and M-SHAKE (Kräutler et al., 2001)). The equivalence between these two approaches is well-established (e.g., Wendlandt & Marsden, 1997). These numerical methods allow an $o(1)$ integration step, but they also require solving nonlinear systems at each step. Unfortunately, linearization of these methods are no longer symplectic, and therefore resorting to linearization for a speed-up is at the risk of losing long-time accuracy.

The advantage of using a penalty approach depends on the system: if the system has a large number of coupled constraints, then an integration of the penalized system, even with small steps, would still be faster than generalized coordinate and Lagrange multiplier methods, which require solving high-dimensional nonlinear systems.

Z&S provides a compromise by allowing large integration steps ($o(1)$, independent of $\omega$) with limited cost of a linear solve per iteration. It remains accurate when applied to penalized system (1.7), even though the $o(1)$ step does not resolve stiffness of the equation. This is because stiffness in this system results in fast oscillations nontangent to a stable slow manifold (Section 3). Although implicit methods damp high frequencies in oscillations (e.g., Hairer & Wanner, 1996), the approximation of fast oscillations by slower ones (as in Li et al., 2008) is sufficient for the approximation of slow dynamics on the constraint manifold.

We refer the reader to M-SHAKE (Kräutler et al., 2001) for an example of recent developments to the Lagrange multiplier method. While M-SHAKE is limited to systems with distance constraints, Z&S combined with the penalty method can implement arbitrary holonomic constraints.

2. Z&S: structure-preserving and stable integrators

2.1 Derivation from Newmark integrators

The Newmark family of algorithms are extensively used in structural dynamics (Newmark, 1959).

**Integrator 2.1 Newmark**:

\[
\begin{align*}
q_{k+1} &= q_k + h\dot{q}_k + \frac{h^2}{2}[(1-2\beta)a_k + 2\beta a_{k+1}], \\
\dot{q}_{k+1} &= \dot{q}_k + h[(1-\gamma)a_k + \gamma a_{k+1}], \\
a_k &= -M^{-1}\nabla V(q_k).
\end{align*}
\]  

(2.1)

Newmark is generally implicit when $\beta \neq 0$. When $\gamma = \frac{1}{2}$, it is second-order accurate and variational (Kane et al., 2000), and we restrict ourselves to this case in this paper. Integrator 2.1 does not preserve the canonical symplectic form, and it was shown in Skeel et al. (1997) and Marsden & West (2001) that if one pushes forward the update map by a coordinate transform $\eta: TQ \to TQ$ defined as

\[(x, v) := \eta(q, \dot{q}) = (q + \beta h^2 M^{-1} \nabla V(q), \dot{q}), \]  

(2.2)

then we obtain an integrator that preserves the canonical symplectic form on $T^*Q$. 

Downloaded from https://academic.oup.com/imajna/article-abstract/36/1/80/2363595/Variational-and-linearly-implicit-integrators-with by Georgia Institute of Technology user on 19 September 2017
**Integrator 2.2** Pushforward Newmark:

\[
\begin{aligned}
x_{k+1} &= x_k + hv_k + \frac{1}{2}h^2 a_k, \\
v_{k+1} &= v_k + \frac{1}{2}h(a_k + a_{k+1}), \\
a_k &= -M^{-1}\nabla V(x_k + \beta h^2 a_k).
\end{aligned}
\] (2.3)

These two methods are unconditionally linearly stable if \( \beta \geq \frac{1}{4} \) (Skeel et al., 1997; Chiba & Kako, 2002). Newmark with \( \beta \geq \frac{1}{4} \) is known to be nonlinearly stable under specific conditions (Hughes, 1977), and Pushforward Newmark is known to be stable near stable fixed points in nonresonant nonlinear settings (Skeel & Srinivas, 2000). Nevertheless, there are nonlinear cases in which Newmark is no longer stable (Kuhl & Ramm, 1996; Erlicher et al., 2002). In fact, few convergent methods are unconditionally stable for arbitrary nonlinear systems to the authors’ knowledge (see also Wood & Oduor, 1988).

Now, consider a second-order discretization of the action \( \int_{(k+1)h}^{kh} L(x, \dot{x}) \, dt \), where \( L = \frac{1}{2}\dot{x}^T M \dot{x} - V(x) \) is the Lagrangian for (1.1). This approximation, known as a discrete Lagrangian (see Marsden & West, 2001 for a review of variational integrators and Abraham & Marsden, 2008 for one of many excellent reviews of analytical mechanics), has the form

\[
\mathcal{L}_d(x_k, x_{k+1}, a_k, a_{k+1}) = h \left( \frac{1}{2} \left( \frac{x_{k+1} - x_k}{h} \right)^T M \left( \frac{x_{k+1} - x_k}{h} \right) \right) - \frac{1}{2} \left( \beta h^2 a_k^T Ma_k + V(x_k + \beta h^2 a_k) \right) - \frac{1}{2} \left( \beta h^2 a_{k+1}^T Ma_{k+1} + V(x_{k+1} + \beta h^2 a_{k+1}) \right). \] (2.4)

It can be shown that (2.3) is the Euler–Lagrange equation associated with (2.4), i.e., the equation that the critical point of the discretized action \( \sum_k \mathcal{L}_d(x_k, x_{k+1}, a_k, a_{k+1}) \) satisfies. Note that this variational formulation is explicit and distinct from the one implicitly defined in Marsden & West (2001). The novelty is the augmentation of acceleration parameters \( a_k \) s.

However, updates given by the Euler–Lagrange equation (2.3) are still implicit. Therefore, we Taylor-expand (2.4) to second order in \( a \) and derive (1.2). To obtain the corresponding discrete Euler–Lagrange equation, we compute \( \partial \mathcal{L}_d / \partial a_k = 0 \), which leads to

\[
Ma_k + \nabla V(x_k) + \beta h^2 \text{Hess} V(x_k)a_k = 0. \] (2.5)

We then compute the discrete Legendre transform (see Marsden & West, 2001 for notation and terminology), which introduces the momentum and leads to

\[
\begin{aligned}
p_k &= -D_1\mathcal{L}_d(x_k, x_{k+1}, a_k, a_{k+1}) \\
&= M\frac{x_{k+1} - x_k}{h} + h \left( -Ma_k + \frac{1}{2} \beta h^2 a_k \cdot V^{(3)}(x_k) \cdot a_k \right), \\
p_{k+1} &= D_2\mathcal{L}_d(x_k, x_{k+1}, a_k, a_{k+1}) \\
&= M\frac{x_{k+1} - x_k}{h} - \frac{h}{2} \left( -Ma_{k+1} + \frac{1}{2} \beta h^2 a_{k+1} \cdot V^{(3)}(x_{k+1}) \cdot a_{k+1} \right).
\end{aligned}
\] (2.6)

Since the velocity and momentum are related via \( v_k = M^{-1}p_k \), we obtain the Z&S update (1.3).
Because the new discrete Lagrangian $\tilde{L}_d$ is quadratic in $a$, nonlinear solves in Pushforward Newmark are replaced by linear solves in Z&S. Consequently, Z&S exhibits a speed advantage. Numerical illustrations of this advantage are in Section 4.

2.2 On Z&S

Linearizing equations: Z&S is obtained as the linearization of a Pushforward Newmark update map combined with a small correction. More precisely, the Taylor expansion of Line 3 of (2.3) leads to (2.5), and Lines 1 and 2 in (2.3) can be rewritten in terms of momentum as

\[
\begin{align*}
\frac{p_k}{h} &= -D_1L_d(x_k, x_{k+1}, a_k, a_{k+1}) \\
&= M \frac{x_{k+1} - x_k}{h} + \frac{h}{2} (-Ma_k), \\
\frac{p_{k+1}}{h} &= D_2L_d(x_k, x_{k+1}, a_k, a_{k+1}) \\
&= M \frac{x_{k+1} - x_k}{h} - \frac{h}{2} (-Ma_{k+1}).
\end{align*}
\]

The differences are two $O(h^5)$ terms (corresponding to $\frac{1}{4} \beta^2 h^2 a \cdot V^{(3)}(x) \cdot a$). To be consistent with the literature, we summarize this variant using velocity instead of momentum as follows.

**INTTEGRATOR 2.3** Zhang and Skeel’s method simplified (Z&Ss):

\[
\begin{align*}
x_{k+1} &= x_k + hv_k + \frac{1}{2} h^2 a_k, \\
v_{k+1} &= v_k + \frac{1}{2} h (a_k + a_{k+1}), \\
a_k &= -(M + \text{Hess} V(x_k) \beta h^2)^{-1} \nabla V(x_k).
\end{align*}
\]

**THEOREM 2.4** Z&Ss is:

1. unconditionally linearly stable if $\beta \geq \frac{1}{4}$;
2. symplectic if the $n \times n$ matrix $M + \text{Hess} V(x) \beta h^2$ commutes with the $n \times n$ matrix $V^{(3)}(x) \cdot (M + \text{Hess} V(x) \beta h^2)^{-1} \cdot \nabla V(x)$ (· is tensor contraction);
3. second-order convergent (if stable) and can be made arbitrarily high-order convergent;
4. symmetric (‘time-reversible’).

Z&Ss is not always symplectic due to the removal of $O(h^5)$ terms. However, it requires no high-order tensor operations, and is thus a good choice for high-dimensional problems.

Partial Newton solve: Line 3 of Z&Ss can be viewed as executing only the first step of a Newton solver for the nonlinear equation $a_k = -(M + \text{Hess} V(x_k) \beta h^2)^{-1} \nabla V(x_k)$.

Preconditioning, filtering and regularization: The factor of $(M + \text{Hess} V(x) \beta h^2)^{-1}$ in front of $\nabla V(x)$ can be thought of as an optimization preconditioner or a way to filter (Hesthaven et al., 2007; Faou & Grebert, 2011)/regularize (García-Archilla et al., 1999; Sanz-Serna, 2008) high-frequency oscillations.
2.3 Properties

(Proofs of results introduced in this paragraph are standard and available online at http://www.math.gatech.edu/~mtao/TaOw14_supplemental.pdf [last accessed 18 January 2015])

**Theorem 2.5 (Stability)** Z&S (Integrator 1.1) is unconditionally linearly stable if and only if $\beta \geq \frac{1}{4}$.

The proofs of the unconditional linear stability (for $\beta \geq \frac{1}{4}$) of Integrators 1.1 and 2.3 are similar. If the potential is of form $V(x) = V_0(x) + \epsilon^{-1}V_1(x)$, then the following modification of Z&Ss is unconditionally linearly stable$^2$ as long as $\beta > \frac{1}{4} + O(\epsilon)$.

**Integrator 2.6 Simplified Z&Ss for stiff systems ($\epsilon^{-1} \gg 1$):**

\[
\begin{aligned}
x_{k+1} &= x_k + hv_k + \frac{1}{2}h^2a_k, \\
v_{k+1} &= v_k + \frac{1}{2}h(a_k + a_{k+1}), \\
a_k &= -M^{-1}(\nabla V_0(x_k) + \epsilon^{-1}\nabla V_1(x_k)) - M^{-1}\epsilon^{-1}\text{Hess } V_1(x_k)\beta h^2a_k.
\end{aligned}
\]

**Theorem 2.7 (Consistency)** Consider an integrator for (1.1) given by

\[
\begin{aligned}
x_{k+1} &= x_k + hv_k + \frac{1}{2}h^2a_k, \\
v_{k+1} &= v_k + \frac{1}{2}h(a_k + a_{k+1} + h^4g(x_k) + h^4g(x_{k+1})), \\
a_k &= -M^{-1}\nabla V(x_k) - M^{-1}f(x_k)h^2a_k,
\end{aligned}
\]

where $f, g \in \mathcal{C}(\mathcal{Q})$ are arbitrary functions. If $V \in \mathcal{C}^3(\mathcal{Q})$, this integrator has third-order truncation error.

**Corollary 2.8** Z&S (Integrator 1.1), Z&Ss (Integrator 2.3) and simplified Z&Ss for stiff systems (Integrator 2.6) are second-order convergent, provided that they are stable.

Symmetry (i.e., time-reversibility) is one desired property of numerical integrators, because it leads to good long-time performance (see, for instance, Hairer et al., 2006 or Leimkuhler & Reich, 2004).

**Theorem 2.9 (Symmetry/time-reversibility)** Let $f \in \mathcal{C}^1(\mathcal{Q})$ be an arbitrary function. The integrator defined by

\[
\begin{aligned}
x_{k+1} &= x_k + hv_k + \frac{1}{2}h^2f_k, \\
v_{k+1} &= v_k + \frac{1}{2}h(f_k + f_{k+1}), \\
f_k &= f(x_k)
\end{aligned}
\]

is symmetric (time-reversible).

**Corollary 2.10** Z&S (Integrator 1.1) is symmetric (time-reversible).

**Remark 2.11** Arbitrary high-order Z&S can be obtained using standard splitting schemes as in Neri (1988), Yoshida (1990) and Hairer et al. (2006). A fourth-order example is provided in Supplementary Material.

---

$^2$ In the sense that the solution remains bounded for all $h$ when $V_0$ has Lipschitz-continuous first-derivative with bounded Lipschitz constant and $V_1$ is quadratic and positive definite.
Theorem 2.12 Z&S (Integrator 1.1) is symplectic.

Lemma 2.13 (Symplecticity) Consider an integrator given by (2.10). If \( f \in C^1(Q) \) is a function with symmetric Jacobian, then this integrator is symplectic.

Remark 2.14 The commutation condition in Theorem 2.4 ensures a symmetric Jacobian, and hence the symplecticity of Z&Ss. Two very special cases where this condition is satisfied are: when the system contains only 1 degree of freedom, or when \( \text{Hess}(V) \) can be diagonalized by a matrix independent of \( x \).

Remark 2.15 Fully nonlinear implicit symplectic methods (e.g., midpoint or Newmark) are not exactly symplectic due to numerical errors in nonlinear solves, which are often much larger than those in linear solves.

3. Lagrange multiplier methods as limits of penalty methods

Lagrange multiplier and penalty methods, respectively, simulate (1.6) and (1.7). It is known (Sections 1.1 and 1.2) that both are equivalent to constrained dynamics (1.5) (the latter in the \( \omega \to \infty \) limit). We now quantify the equivalence between the two.

First, observe that this equivalence is not necessarily achieved via

\[
\lambda(t) = -\lim_{\omega \to \infty} \omega^2 g(q^\omega(t)).
\] (3.1)

Consider, for instance, a 2 degrees of freedom example, in which \( V(q) = (q_2 - q_1)^2/2 \), \( g(q) = q_1 \), \( q(0) = [0; 1] \) and \( p(0) = [0; 0] \). Lagrange multiplier method (1.6) yields \( q_1(t) = 0 \), \( q_2(t) = \cos t \) and \( \lambda(t) = -\cos t \). Penalty method (1.7) leads to \( q_1^\omega(t) = (1/\sqrt{4 + \omega^2}) (\cos(\omega_1 t) - \cos(\omega_2 t)) \), where \( \omega_{1,2} = \sqrt{(2 + \omega^2 + \sqrt{4 + \omega^4})/2} \). We see (3.1) cannot hold because \( \lim_{\omega \to \infty} \omega^2 q_1^\omega(t) \) does not exist due to fast oscillations.

The appropriate notion of equivalence is provided by Theorem 1.3. The idea is as follows: energy conservation implies that \( g(q^\omega) \) is at most \( O(1/\omega) \) (see Lemma A.1). In fact, \( g(q^\omega) \) can be further shown to be \( O(1/\omega^2) \) (see Remark A.3 or Kevorkian & Cole, 1996), and constraints are satisfied with small errors that oscillate rapidly. To describe the Lagrange multiplier system (1.6) as a limit of penalized systems (1.7), the convergence of these fast oscillations should be understood in a weak sense, whereas slow dynamics on the constrained manifold converges strongly. Thus, we employ two-scale flow convergence (Equation (1.11)) for this description. Convergence is first proved for a flat constraint manifold (Lemma A.2), and then local charts are patched together (see appendix); this leads to Theorem 1.3.

Remark 3.1 If the limit in (3.1) exists, then (1.8) simplifies to (3.1) and two-scale F-convergence becomes strong convergence.

4. Application examples

Z&S (Integrator 1.1) is applied in Sections 4.1, 4.2 and 4.4; Section 4.3 employs simplified Z&Ss for a stiff system (Integrator 2.6) due to its efficiency for high-dimensional systems; Section 4.5 is based on variational formulation (1.2).

Speed comparisons are provided in terms of running times (using Matlab 7.7 on an Intel Core 2 Duo 2.4G laptop, with nonlinear solver of ‘fsolve’); however, these numbers are machine and platform dependent (e.g., Matlab is very well optimized for linear algebra), and should serve only as a qualitative illustration of efficiencies.
4.1 Double pendulum

Implementation: One way to represent planar double pendulum is to use 4 degrees of freedom and 2 nonlinear constraints. Using the notation of (1.7), we have

$$M = \begin{bmatrix} m_1 & 0 & 0 & 0 \\ 0 & m_1 & 0 & 0 \\ 0 & 0 & m_2 & 0 \\ 0 & 0 & 0 & m_2 \end{bmatrix},$$

$$V(x_1, y_1, x_2, y_2) = -g y_1 - g y_2,$$

$$g(x_1, y_1, x_2, y_2) = \begin{bmatrix} x_1^2 + y_1^2 - L_1^2 \\ (x_2 - x_1)^2 + (y_2 - y_1)^2 - L_2^2 \end{bmatrix}. $$

For simplicity, we adopt a dimensionless convention and assume $m_1 = m_2 = g = 1$.

The Z&S simulation of the penalized system (1.7) is straightforward. SHAKE (Ryckaert et al., 1977) is used as the Lagrangian multiplier method in our experiments; it is nonlinearly implicit.

Symplectic integration in generalized coordinates $\theta, \phi$ ($x_1 = L_1 \sin \theta$, $y_1 = -L_1 \cos \theta$, $x_2 = L_1 \sin \theta + L_2 \sin \phi$, $y_2 = -L_1 \cos \theta - L_2 \cos \phi$) is also implicit. This is because, after writing down the Lagrangian, one will note a position-dependent mass matrix of

$$\tilde{M}(\theta, \phi) = \begin{bmatrix} 2L_1^2 & L_1 L_2 (\cos \theta \cos \phi + \sin \theta \sin \phi) \\ L_1 L_2 (\cos \theta \cos \phi + \sin \theta \sin \phi) & L_2^2 \end{bmatrix}. \quad (4.1)$$

Consequently, even the most well-known ‘explicit’ variational integrators such as variational Euler (i.e., leapfrog) and velocity-Verlet, will be implicit.

Note that although $g$ is quadratic, the penalized ODE is cubically nonlinear.

Results: Figure 1 illustrates errors of different methods. Newmark (Integrator 2.1) with only the first step of nonlinear solve (Row 5) has a large error due to the loss of symplecticity, even though the method is still consistent. On the contrary, Z&S (Row 6) yields small errors almost identical to those of fully nonlinearly solved Newmark (Row 4).

Z&S produces larger error than SHAKE, because there is modelling error due to finite $\omega$ in addition to integration error (Row 3). We chose an intermediate $\omega$, which is sufficiently large to approximate the constraints, yet small enough to show that the penalized system is only an approximation. A larger $\omega$ leads to a more accurate approximation, but if it is too large, e.g., $\omega = 2000$ (i.e., a stiffness of $\omega^2 = 4 \times 10^6$), instability occurs in all Z&S, original Newmark and implicit midpoint due to strong nonlinearity.

If $\omega$ is finite, the approximation error is predicted to be $O(\omega^{-2})$ (Remark A.3). See Fig. 2(a) for a numerical illustration. Figure 2(b) compares the Lagrange multiplier computed by SHAKE with the one obtained from the penalized system via Theorem 1.3. There is no strong convergence, but only a two-scale F-convergence.

It is known that the double pendulum contains a chaotic region (e.g., Richter & Scholz, 1984). A variational integrator is desired for simulating such systems (Channell & Scovel, 1990; McLachlan & Atela, 1992). None of our symplectic simulations (Rows 1–4, 6) led to numerical leakage between regular and chaotic regions.

Generalized coordinate implicit VE (benchmark), SHAKE, variational Euler, Newmark with full nonlinear solve, Newmark with one-step nonlinear solve and Z&S, respectively, spent 91.3, 3.8, 1.0, 5.3, 0.2 (2.6 if Hess $V$ is not analytically provided, but approximated by the nonlinear solver), and 0.4 s on the above simulation.
4.2 A simple high-dimensional example: a chain of many pendula

Consider a chain of \( n \) pendula, which approximates a continuous rope. The system is similarly modelled by (1.7) with

\[
M = \begin{bmatrix}
1 & 0 & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 1 & 0 \\
0 & 0 & \cdots & 0 & 1
\end{bmatrix},
\]

\[
V(x_1, y_1, \ldots, x_n, y_n) = \sum_{i=1}^{n} y_i,
\]

\[
\begin{aligned}
g(x_1, y_1, \ldots, x_n, y_n) &= \begin{bmatrix}
x_1^2 + y_1^2 - L_1^2 \\
(x_2 - x_1)^2 + (y_2 - y_1)^2 - L_2^2 \\
\vdots \\
(x_n - x_{n-1})^2 + (y_n - y_{n-1})^2 - L_n^2
\end{bmatrix}.
\end{aligned}
\]
Figure 3 shows good agreement between SHAKE and Z&S (trajectories instead of errors are shown due to lack of accurate benchmark—an analytical solution is unavailable, and Lagrange multiplier formulations and generalized coordinate approaches involve solving large nonlinear systems). SHAKE with $h = 0.025$, $h = 0.05$ and Z&S with $h = 0.05$, respectively, spent 16.5, 8.1 and 1.1 s in these simulations.

4.3 Molecular dynamics of water cluster

Consider the dynamics of water molecules, each interacting with others via nonbonded interactions of electrostatic and van der Waals forces (both highly nonlinear).

Water model: Use the popular TIP3P model (e.g., Jorgensen et al., 1983). Let $q_{ai}$ and $p_{ai}$ be the position and momentum of the $a$th molecule’s $i$th atom (both 3-vectors). The Hamiltonian is

$$
\mathcal{H} = \sum_{a=1}^{N} \sum_{i=1}^{3} \frac{1}{2} p_{ai}^2 m_i^{-1} p_{ai} + \sum_{a=1}^{N-1} \sum_{b=a+1}^{N} \left( \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{K_{i} Q_{i} Q_{j}}{r_{ai,bj}} + \frac{A}{r_{ai,bj}^{12}} - \frac{C}{r_{ai,bj}^{6}} \right),
$$

where $r_{ai,bj} := \|q_{ai} - q_{bj}\|$ is the inter-atom distance, $m_1 = m_3$ are hydrogen mass and $m_2$ is oxygen mass, $K_i$ is electrostatic constant, $Q_i$ is the partial charge of atom $i$ relative to the electron charge, and the $A$ and $C$ are Lennard–Jones constants that approximate van der Waals forces.

In this TIP3P model (or many other prevailing models such as SPC, BF, TIPS2 and TIP4P, which are discussed in, e.g., Jorgensen et al., 1983; van der Spoel et al., 1998), each water molecule is considered as a rigid body, with two O–H bond lengths and H–O–H bond angle fixed as constants $r_{OH}$ and $\alpha_{HOH}$. Detailed values of these model parameters could be found in, e.g., Jorgensen et al. (1983).
Therefore, the following vectorial constraint enforces the geometry of molecules:

\[
g(q) = \begin{bmatrix}
(q_{11} - q_{12})(q_{11} - q_{12})^T - r_{\text{OH}}^2 \\
(q_{13} - q_{12})(q_{13} - q_{12})^T - r_{\text{OH}}^2 \\
(q_{11} - q_{13})(q_{11} - q_{13})^T - r_{\text{HH}}^2 \\
(q_{N1} - q_{N2})(q_{N1} - q_{N2})^T - r_{\text{OH}}^2 \\
(q_{N3} - q_{N2})(q_{N3} - q_{N2})^T - r_{\text{OH}}^2 \\
(q_{N1} - q_{N3})(q_{N1} - q_{N3})^T - r_{\text{HH}}^2
\end{bmatrix},
\]  

(4.3)

where \( r_{\text{HH}} := 2r_{\text{OH}} \sin(\alpha_{\text{HOH}}/2) \) is a constant. This leads to penalized Hamiltonian

\[
\tilde{\mathcal{H}} = \mathcal{H} + \frac{1}{2} \omega^2 \sum_{a=1}^{N} \left( (r_{a1,a2}^2 - r_{\text{OH}}^2)^2 + (r_{a3,a2}^2 - r_{\text{OH}}^2)^2 + (r_{a1,a3}^2 - r_{\text{HH}}^2)^2 \right).
\]  

(4.4)
**Constant temperature simulation:** Constant temperature simulations are of practical importance because (i) thermal fluctuations are an indispensable component of molecular dynamics; and (ii) the N-body system is chaotic and its long-time deterministic simulation has limited predictive power. We use Langevin dynamics (e.g., Schlick, 2010) as our constant temperature model.

In this model, molecules experience perturbation by noise and dissipation due to friction, and the dynamics can be expressed by the following SDEs:

\[
\begin{align*}
\frac{dq}{dt} & = \frac{\partial H}{\partial p}, \\
\frac{dp}{dt} & = -\frac{\partial H}{\partial q} dt - \gamma \frac{\partial H}{\partial p} dt + \sqrt{2\gamma \beta^{-1}} dW,
\end{align*}
\] (4.5)

where \( W \) is a \( 9N \)-dimensional Wiener process, \( \beta^{-1} > 0 \) is the constant temperature and \( \gamma > 0 \) is dissipation strength. The system admits an invariant measure of Boltzmann–Gibbs (BG; also known as the canonical ensemble), given by

\[
\pi(q,p) = Z^{-1} \exp(-\beta H),
\] (4.6)

where \( Z = \int_{\mathbb{R}^{6N}} \exp(-\beta H) \, dq \, dp \) is the partition function.

To simulate (4.5), we use the Geometric Langevin Algorithm (GLA; see Bou-Rabee & Owhadi, 2010). GLA allows for an extension of Hamiltonian integrators to Langevin integrators. It is a splitting scheme, based on composing the one-step update of a deterministic integrator with the exact flow of an Ornstein–Uhlenbeck process (given by \( dp = -\gamma M^{-1} p \, dt + \sqrt{2\gamma \beta^{-1}} dW \), i.e., driftless noise and friction). It has been shown (Bou-Rabee & Owhadi, 2010) that if the deterministic integrator is symplectic, then GLA not only provides a good approximation of trajectories, but also of BG (the invariant distribution). In this example, the deterministic building block is simplified Z&Ss (Integrator 2.6) or SHAKE.

Constant temperature molecular dynamics is a rich research field, and our investigation will only be numerical. The thermodynamic properties of a system with strong restraint may not be equivalent to those of a constrained system (e.g., Bajars et al., 2011); the Fixman potential is a classical way to correct the difference (see Perchak et al., 1985 for a debate on the validity of this correction). Proving that a numerical method samples a good approximation of the invariant distribution is nontrivial. Bou-Rabee & Owhadi (2010) combine ergodicity with the backward error analysis of symplectic integrators to show that the invariant distribution is preserved with a high order of accuracy. It is conjectured (Bou-Rabee & Owhadi, 2010, Remark 2.1) that SHAKE + GLA approximately samples a constrained BG distribution

\[
\hat{\pi}(q,p) = \hat{Z}^{-1} \exp(-\beta \hat{H}),
\] (4.7)

where \( \hat{Z} = \int_{T^*g^{-1}(0)} \exp(-\beta \hat{H}) \, dq \, dp \). Relating (4.6) and (4.7) is left as a future investigation. See Vanden-Eijnden & Ciccotti (2006), Bajars et al. (2011), Hartmann (2008) and Lelièvre et al. (2012) for more about finite temperature constrained dynamics.

**Numerical results:** One quantity of interest in a water cluster is the distribution of interatomic oxygen–oxygen distances in the thermal equilibrium limit, also known as the OO radial distribution (Jorgensen et al., 1983). To illustrate the accuracy of Z&S in sampling BG, Fig. 4 shows histograms obtained by long-time simulations of SHAKE and simplified Z&Ss (Integrator (2.6)) that approximate this distribution. We chose a system of size \( N = 7 \) (i.e., 63 degrees of freedom) so that peaks in the distribution could be clearly distinguished. SHAKE required 13472 s, including 12284 s on nonlinear
solves (with tolerances of $10^{-6}$ on variable and $10^{-10}$ on function value), whereas simplified Z&Ss used 1549 s, including 67 s on linear solves. Parameters are $\omega = 20$, $h = 0.05$ in both simulations, $\gamma = 0.01$ and $\beta = 50$.

We also provide two deterministic simulations (with noise and friction turned off; other parameters remain unchanged unless indicated otherwise): (i) Fig. 5 compares Lagrange multipliers computed by SHAKE and from the penalized system to illustrate Theorem 1.3. (ii) Figure 6 compares simplified Z&Ss, SHAKE and partially solved Newmark (nonsymplectic) in terms of energy and momentum conservations. Simplified Z&Ss lost symplecticity due to simplification, but it still exhibits improved preservation properties comparing with partially solved Newmark.

To test scalability, we increase $N$ to 100 (900 degrees of freedom) and illustrate results in Fig. 7. SHAKE spent 42234 s, including 14272 s on solving nonlinear systems and $\sim 28000$ s on computing $V$ and $\nabla V$, whereas simplified Z&Ss spent 30059 s, including 319 s on solving linear systems, and $\sim 29000$ s on $V$, $\nabla V$ and Hess $V$.

Efficiency of force evaluation: Although simplified Z&Ss accelerates updates by linearization, for large systems the computational bottleneck is likely to be on force evaluations, but not updates. Fortunately, significant progress has been made to accelerate force evaluations, such as the fast multipole method (Greengard & Rokhlin, 1987), or simply the idea of ignoring weak long-range forces. We did not employ any of them, but they can be used in adjunct to simplified Z&Ss.

Times spent on force evaluations by SHAKE and simplified Z&Ss are comparable; Hessian computations in simplified Z&Ss did not incur much overhead. This is because the potential is a function of relative distances $r_{ij} = \|x_i - x_j\|$. For such $f$,

$$
\frac{\partial^2 f(r)}{\partial x_i \partial x_j} = \frac{\partial r}{\partial x_i} \frac{\partial^2 f}{\partial r^2} \frac{\partial r}{\partial x_j} + \frac{\partial f}{\partial r} \frac{\partial^2 r}{\partial x_i \partial x_j},
$$

but $\partial r/\partial x$ and $\partial f/\partial r$ are already computed when calculating the gradient, $\partial^2 f/\partial x_i \partial x_j$ is cheap to obtain and $\partial^2 f/\partial r^2$ is the only new component of computation, but it is a scalar. In addition, nonlinear

Fig. 4. Empirical OO radial distribution in a 7-water cluster obtained by long-time ($T = 10000$) simulations of SHAKE and simplified Z&Ss.

\[\begin{align*}
\text{SHAKE } r_{OO} \text{ histogram} \\
\text{Z&Ss } r_{OO} \text{ histogram}
\end{align*}\]
Fig. 5. Lagrange multipliers from the penalized system (simplified Z&Ss) and by SHAKE. Only the first oxygen atom is shown, and illustration is terminated before chaos. Empirical average uses a time window of width 0.2; $\omega$ is temporarily enlarged to 500 for clearer visualization of details.

The linear system associated with the Hessian can also be solved in $O(N)$ time. This is because the Hessian is dominated by a block diagonal due to localized stiff penalty terms in (4.4). Simplified Z&Ss further reduces the Hessian to completely block diagonal, and linear solves are executed molecule by molecule. Similar efficiency can be obtained for polymers as long as the number of bonds is at the same order as the number of atoms.

4.4 Coarse time-stepping of a DNA model

We now show how Z&S accelerates the simulation of an unconstrained multiscale system. Consider the simple DNA model proposed in Mezić (2006) and further studied, e.g., in Toit et al. (2009) and Koon et al. (2013). The displacement angle of the $k$th base in one strand, $\theta_k$, follows:

$$\ddot{\theta}_k = \theta_{k+1} - 2\theta_k + \theta_{k-1} - \epsilon U'(\theta_k),$$

where $U(\theta) = (\exp(-a[1 - \cos(\theta) - x_0]) - 1)^2$ is a Morse potential modelling complementary base pairings between two DNA strands, and linear force models the tendency of alignment between neighbouring bases. Unitless parameters are $a = 7$, $x_0 = 0.3$, $\epsilon = \frac{1}{1400}$ and the number of base-pairs $N = 200$ (Toit et al., 2009). Two stable configurations are given by minima of $U$ and correspond to closed double strands. Nonlinearity in this system is critical, for it leads to transitions between meta-stable
Fig. 6. Energy, linear and angular momentum preservation by simplified Z&Ss, SHAKE and partially solved Newmark. $h = 0.05$ for all. For clarity, plots are drawn with a 20:1 downsample rate.

Fig. 7. Empirical OO radial distribution in a 100-water cluster obtained by simulations of SHAKE and simplified Z&Ss till $T = 1000$. 
states that correspond to the opening of double strands. We simulate such transitions with initial positions $\theta_k = 0.8 + 0.1 \xi_k$ (i.i.d. standard norm), which is near a stable configuration, and initial momenta $\dot{\theta}_k = \cos(4\pi k/N)/\sqrt{N}$, which facilitates the opening-up of double strands (Toit et al., 2009).

It is known Toit et al. (2009) that $\bar{\theta} = \sum \theta_k/N$ is a slow variable and can be used as a reaction parameter, whereas individual $\theta_k$s are fast variables. Figure 8 presents simulations by velocity-Verlet (benchmark) and Z&S ($\beta = 0.3$) in these variables. The phase portrait shows that the DNA transits between meta-stable configurations $\theta = \arccos(0.7) \approx 0.795$ and $\theta = 2\pi - \arccos(0.7) \approx 5.49$. Note that these are long-time simulations and the system is chaotic (Mezić, 2006).

The Z&S energy is lower than the benchmark because fast oscillations are damped by large time-steps. Both $h = 0.2$ in velocity-Verlet and $h = 2$ in Z&S are near stability limits. The methods, respectively, used 23.26 and 3.24 s of CPU time.

4.5 Lie group integration

Formulating Z&S as a variational principle allows us to generalize the method to mechanical systems on Lie groups.
Consider a prototypical example of magnetized 3D rigid body with identity inertia matrix immersed in a constant magnetic field. The configuration space is $Q = SO(3)$. Denote by $R(t) \in Q$ the (generalized) rigid body position; in coordinates it is a $3 \times 3$ matrix satisfying $R^T R = I$. Suppose when $R = I$ both the magnetic field and the dipole are in the $z$-direction; then, the potential energy can be written as $V(R) = \langle \mu R e_3, B e_3 \rangle = B \mu e_3^T R^{-1} e_3$, where $B$ and $\mu$ are field strength and dipole moment, and $e_3 = [0 \ 0 \ 1]^T$.

Let $\Omega(t) \in \mathbb{R}^3$ be convective angular velocity of the body; then the kinetic energy is $\frac{1}{2} \Omega^T \Omega$. Introduce an isomorphism between $\mathbb{R}^3$ and $\mathfrak{so}(3)$ (the Lie algebra of SO(3)) by

$$\Omega \mapsto \hat{\Omega} = \begin{bmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{bmatrix}.$$  

Then, $\dot{R} = R \hat{\Omega}$. It is known (Marsden & Ratiu, 2010) that dynamics of this mechanical system can be obtained from either of the following equivalent variational principles:

1. $$\delta \int_0^T L(R, \dot{R}) \, dt = 0, \quad (4.10)$$
   with arbitrary variations of $R(t) \in Q$.

2. $$\delta \int_0^T l(R, \xi) \, dt = 0, \quad (4.11)$$
   with variations in the form $\delta \xi = \dot{\eta} + \text{ad}_\xi \eta$ under $R \in Q$ and $\xi = R^{-1} \dot{R}$.

For our system, $L(R, \dot{R}) = \frac{1}{2} \text{tr}(\dot{R}^T \dot{R}) - B \mu e_3^T R^{-1} e_3$ and $l(R, \hat{\Omega}) = \frac{1}{2} \Omega^T \Omega - B \mu e_3^T R^{-1} e_3$ (note $\Omega^T \Omega = \frac{1}{2} \text{tr}(\hat{\Omega}^T \hat{\Omega})$).

We propose to simulate the system by modifying (4.10). The result is compared with a benchmark derived from (4.11) via the Hamilton–Pontryagin principle, backward Variational Euler discretization and Cayley approximation of the exponential map (see Iserles, 2001; Hairer et al., 2006 for Cayley approximation, Bou-Rabee & Marsden, 2009 for the benchmark method and Iserles et al., 2000; Lee et al., 2007 for examples of other Lie group integrators). The benchmark uses update rules:

$$\begin{cases} R_{k+1} = R_k (I - h \hat{\Omega}_{k+1} / 2)^{-1} (I + h \hat{\Omega}_{k+1} / 2), \\
\hat{\Omega}_{k+1} = \hat{\Omega}_k + \frac{h^2}{4} \left( \hat{\Omega}^T_{k+1} \hat{\Omega}_{k+1} \hat{\Omega}^T_{k+1} - \hat{\Omega}^T_k \hat{\Omega}_k \hat{\Omega}^T_k \right) + h R_k \frac{\partial l}{\partial R} (R_k, \hat{\Omega}_{k+1}) \end{cases} \quad (4.12)$$

Note $R$ is in a three-dimensional manifold, and the differential in the last term should not be computed as a partial derivative with respect to nine Cartesian coordinates of $R$; otherwise the last term will not be in $\mathfrak{so}(3)$. Instead, we follow Holm et al. (1998) and obtain

$$h R_k \frac{\partial l}{\partial R} (R_k, \hat{\Omega}_{k+1}) = hB \mu R_k \frac{-e_3^T R^{-1}_k e_3}{\partial R_k} = hB \mu (R^{-1}_k e_3 \times e_3)^\wedge. \quad (4.13)$$

Equation (4.12) is variational and thus numerically energy- and momentum preserving. Owing to Cayley approximation, it also preserves the SO(3) structure in the sense that $R_k^T \dot{R}_k = I$ up to arithmetic error.
However, variational methods of this type are intrinsically nonlinearly implicit due to curved geometry when $Q$ is noncommutative (e.g., Bou-Rabee & Marsden, 2009; Kobilarov et al., 2009).

Our goal is to avoid expensive nonlinear solves and bypass force evaluations that require geometric calculations (such as (4.13)). To do so, we first add penalization to (4.10):

$$\delta \int_0^T \frac{1}{4} \text{tr}(\dot{R}^T R) - B\mu R(3, 3) + \frac{1}{2} \omega^2 \text{tr} ((R^T R - I)^T (R^T R - I)) \, dt = 0,$$

where $R \in \text{SO}(3)$ is relaxed to $R \in \mathbb{R}^{3 \times 3}$. We then discretize the action as follows:

$$\mathcal{L}_d(R_k, R_{k+1}, a_k) = \text{tr} \left( \frac{1}{4} \left( \frac{R_{k+1} - R_k}{h} \right)^T \left( \frac{R_{k+1} - R_k}{h} \right) - \beta h^2 \frac{1}{2} a_k^T a_k \right)$$

$$- B\mu e_3^T (R_k + \beta h^2 a_k) e_3 - \text{tr} \left( \frac{1}{2} \omega^2 \left( \left( R_k + \beta h^2 a_k \right)^T \left( R_k + \beta h^2 a_k \right) - I \right)^2 \right).$$

Finally, we truncate terms that are higher than second order in $a_k$. After using trace identities $\text{tr}(AB) = \text{tr}(BA)$ and $\text{tr}(A^T) = \text{tr}(A)$, the truncated action simplifies to

$$\tilde{\mathcal{L}}_d(R_k, R_{k+1}, a_k) = \text{tr} \left( \frac{1}{4} \left( \frac{R_{k+1} - R_k}{h} \right)^T \left( \frac{R_{k+1} - R_k}{h} \right) - \beta h^2 \frac{1}{2} a_k^T a_k \right)$$

$$- B\mu e_3^T (R_k + \beta h^2 a_k) e_3 - \text{tr} \left( \frac{1}{2} \omega^2 \left( (R_k^T R_k - I)^2 + 4\beta h^2 (R_k^T R_k - R_k^T R_k - R_k^T R_k) a_k \right. \right.$$

$$+ 2\beta^2 h^4 (a_k^T R_k^T a_k + R_k^T a_k R_k^T a_k + a_k^T R_k^T a_k - a_k^T a_k)) \right).$$

Unconstrained variation of this action with respect to $a_k$ gives

$$a_k = -B\mu e_3^T e_3^T - 2\omega^2 (R_k^T R_k - R_k + \beta h^2 (a_k^T R_k^T R_k + R_k^T a_k R_k + R_k^T R_k a_k - a_k)).$$

(4.15)

Standard variational integrator construction leads to

$$p_k = -D_1 \tilde{\mathcal{L}}_d(R_k, R_{k+1}, a_k), \quad p_{k+1} = D_2 \tilde{\mathcal{L}}_d(R_k, R_{k+1}, a_k).$$

Let $f_k = a_k + 2\omega^2 \beta^2 h^4 (a_k^T R_k^T a_k + R_k^T a_k R_k^T a_k + R_k^T R_k a_k - a_k)$ and use (4.15) for simplification; then the above becomes

$$\begin{cases} p_{k+1} = p_k + hf_k, \\
R_{k+1} = R_k + 2hp_{k+1}. \end{cases}$$

These are our variational linearized SO(3) integrators. Note that (4.14) is based on a first-order quadrature; the second-order trapezoidal rule would lead to

$$\begin{cases} p_{k+1/2} = p_k + \frac{h}{2}f_k, \\
R_{k+1} = R_k + 2hp_{k+1}, \\
p_{k+1} = p_k + \frac{h}{2}f_{k+1}. \end{cases}$$
Fig. 9. Snapshots of magnetized rigid body dynamics. Red and black arrows represent magnetic field and dipole, respectively.

Fig. 10. Preservation of energy, Lie group structure and deviation from benchmark trajectory $R'$. (a) Variational Z&S ($h=0.1$); and (b) variational Euler on Lie group ($h=0.1$).

These are similar to Z&S updates (Integrator 1.1) although Z&S works in $\mathbb{R}^n$. Some may question the usefulness of a variational formulation, because one can represent $R$ by a nine-dimensional vector, view the penalized system as Newton’s equation and then use Z&S. In Z&S updates (1.3), however, $V^{(3)}$ is essentially a 6-tensor, and its brute-force calculation in coordinates, as well as its contractions with $a$ from both left and right, will be unpleasant. A variational approach minimizes the involvement of coordinates and reduces the effort.

Figure 9 demonstrates benchmark ($h=0.0001$) and variational Z&S simulations ($h=0.1$, $\omega=10$, $\beta=0.4$); their difference, regarded as our method’s error, is quantified in Fig. 10(a); the error of the benchmark method with $h=0.1$ is also provided in Fig. 10(b) as a comparison. The full simulation is available at http://youtu.be/29deMRDRsuU (last accessed 18 January 2015). Initial conditions are $R(0)=I$ and $\Omega(0)=[1;0.2;0.1]$. 
Our method and variational Euler with both \( h = 0.1 \), respectively, spent 0.03 and 1.45 s on computations. However, a variational Lie group integrator is much better at preserving the Lie group structure. Applicabilities of the two approaches are disjoint: for example, variational Z&S generally suits computer graphics better, where real time rendering requires high efficiency, while demand on accuracy is moderate (as long as the result looks good); to orient satellites (e.g., Junge & Ober-Blobaum, 2005), on the other hand, one should choose variational Lie group integrators over variational Z&S, and it is worth CPU hours to precompute trajectories with high fidelity.

Acknowledgements

We thank Mathieu Desbrun for motivation and discussions, Joel Tropp and Eitan Grinspun for discussions, Carmen Sirois for proofreading the manuscript and anonymous reviewers for comments.

Funding

This work was supported by NSF grant CMMI-092600, a generous gift from UTRC and Courant Instructorship from New York University.

References


Appendix

Lemma A.1 Consider (1.7). If $V$ is bounded from below, then there is a constant $C$, such that $\|g(q^{\omega}(s))\| \leq C/\omega$ for all $s$. Moreover, if $V(q)$ diverges to infinity as $|q| \to \infty$, then there is a constant $C$ such that $\|q^{\omega}(s)\| \leq C$.

Proof. Note that the energy $[\dot{q}^{\omega}]^T M \dot{q}^{\omega}/2 + V(q^{\omega}) + \omega^2 g(q^{\omega})^T g(q^{\omega})$ in the penalized system (1.7) is conserved and determined by the initial condition. Therefore, $V(\cdot)$ being bounded from below and $g^T g \geq 0$ imply that $\omega^2 g(q^{\omega})^T g(q^{\omega}) = \mathcal{O}(1)$. Hence, $g(q^{\omega}(s)) = \mathcal{O}(1/\omega)$.

By a similar energy argument, since $[\dot{q}^{\omega}]^T M \dot{q}^{\omega}/2 \geq 0$ and $g(q^{\omega})^T g(q^{\omega}) \geq 0$, $V(q^{\omega})$ is bounded from above too, which implies that $q^{\omega}$ remains bounded.

Lemma A.2 Consider the solution to a conserved mechanical system

$$\begin{cases}
\ddot{x}^{\omega} = f_1(x^{\omega}, y^{\omega}), \\
\ddot{y}^{\omega} = f_2(x^{\omega}, y^{\omega}) - \omega^2 g(y^{\omega})^T \nabla g(y^{\omega}),
\end{cases}$$

(A.1)

where $x^{\omega}$ and $y^{\omega}$ are vectors, and $x^{\omega}(0) = x_0, \dot{x}^{\omega}(0) = \dot{x}_0, y^{\omega}(0) = y_0, \dot{y}^{\omega}(0) = \dot{y}_0$. Suppose that $f_1, f_2$ and $\nabla g$ are $C^1$ with bounded derivatives, $\dot{x}^{\omega}$ and $\dot{y}^{\omega}$ are bounded, $g(y_0) = 0$ and $(d/dt)g(y_0) = 0$, and $g(\cdot)$ has a non-degenerate Jacobian in a neighbourhood of $y_0$; then

$$\lambda(t) := -\lim_{T \to 0} \lim_{\omega \to \infty} \frac{1}{T} \int_t^{t+T} \omega^2 g(y^{\omega}(s)) \, ds,$$

(A.2)

exists and is finite. Denote by $x(t), y(t)$ the solution to

$$\begin{cases}
\ddot{x} = f_1(x, y), \\
\ddot{y} = f_2(x, y) + \lambda^T \nabla g(y), \\
g(y) = 0,
\end{cases}$$

(A.3)

with the same initial conditions $x_0, \dot{x}_0, y_0, \dot{y}_0$; then, as $\omega \to \infty$,

$$\begin{cases}
x^{\omega} \to x, \\
y^{\omega} F \to y, \\
g(y^{\omega}) \to 0.
\end{cases}$$

(A.4)

Proof. We employ the multiscale averaging framework described in Tao et al. (2010) to demonstrate the convergence. Here, $x^{\omega}$ is a slow variable and its evolution corresponds to the constrained dynamics; $y^{\omega}$ is a fast variable corresponding to a fluctuating deviation from the constraint manifold at a characteristic timescale of $\mathcal{O}(1/\omega)$, and it lies in the normal bundle of the constraint manifold.
First, consider the linear constraint case in which \( g(y) = Cy^T \) for some nonsingular \( C \) (the affine case can be similarly treated by shifting \( y \)). We see that \( y \) dynamics is governed by

\[
\ddot{y}^o = f_2(x^o, y^o) - \omega^2 y^o C^T C.
\] (A.5)

This is a forced harmonic oscillator, and its solution can be written as

\[
y^o(t) = \int_0^t f_2(x^o(s), y^o(s)) \sin(\omega \tilde{C} s) \tilde{C}^{-1/\omega} ds,
\] (A.6)

where \( \tilde{C} = \sqrt{C^T C} \) is the well-defined matrix square root, and matrix sin is defined either by Taylor expansion or diagonalization. Note that there is no propagation of the initial condition because \( y^o(0) = 0 \).

It can be shown from (A.6) (for instance, by Tao et al., 2011, Lemma 3.8; the idea is that an addition \( 1/\omega \) comes from the sin due to integration by parts) that \( y^o(t) \) is \( O(\omega^{-2}) \) at least up to \( t = o(1) \), and \( y^o \) is asymptotically periodic (because (A.5) is asymptotically linear), and hence locally ergodic on the energy shell (with Dirac ergodic measure).

Since \( y^o \) is locally ergodic on the energy shell, (A.2) well defines \( \lambda \), and Tao et al. (2010, Theorem 1.2) guarantees that the effective equation for (A.5) is

\[
\ddot{y} = f_2(x^o, y) + \lambda^T C,
\] (A.7)

in the sense that \( y^o \xrightarrow{\mathcal{F}} y \) and \( x^o \rightarrow x \). Note that the convergence on \( x \) is in the strong sense, i.e., \( \lim_{\omega \to \infty} x^o(t) \rightarrow x(t) \) for all bounded \( t > 0 \). This is because \( x \) is purely slow, for which case, \( \mathcal{F} \)-convergence implies strong convergence.

Now, consider a fully nonlinear \( g(\cdot) \) with a nondegenerate Jacobian. Lemma A.1 gives that \( g(y^o) = \mathcal{O}(1/\omega) \). Since \( y^o \) is by assumption bounded, inverting \( g \) leads to \( y^o - y_0 = \mathcal{O}(1/\omega) \). Consequently, the dynamics of \( y^o \) approaches that of a forced oscillator (with equilibrium at \( y_0 \)) at an \( \mathcal{O}(1/\omega) \) timescale, because \( g(\cdot) \) is approximated by its first-order Taylor expansion:

\[
\ddot{y}^o = f_2(x^o, y^o) - \omega^2 g(y^o)^T \nabla g(y^o)
= f_2(x^o, y^o) - \omega^2 \left( \nabla g(y_0)(y^o - y_0)^T + \mathcal{O}(\omega^{-2}) \right) ^T \left( \nabla g(y_0) + \text{Hess } g(y_0)(y^o - y_0)^T + \mathcal{O}(\omega^{-2}) \right)
= f_2(x^o, y^o) - \omega^2 (y^o - y_0) \nabla g(y_0)^T \nabla g(y_0) + \mathcal{O}(1),
\]

where nonlinearity \( f_2 + \mathcal{O}(1) \) again manifests as a slow force, which is dominated by the linear term that leads to asymptotically periodic oscillations. Hence, similar to the linear case, \( y^o \) is locally ergodic on the energy shell, the Lagrange multiplier \( \lambda \) is well-defined and the solution \( x^o, y^o \) \( \mathcal{F} \)-converges to the effective solution \( x, y \). \( \square \)

**Sketch of the proof of Theorem 1.3.** (Figure A1 illustrates the notation used in the proof to help understand the geometry.) Since \( g(q^o) \) is at most \( \mathcal{O}(1/\omega) \) (Lemma A.1), \( q^o \) is close to the constraint manifold.
$g^{-1}(0)$ in the sense that if we define, for all $t$,

$$q_0(t) := \min_{q \in g^{-1}(0)} \| q - q^{o}(t) \|$$

(A.8)

then $q^{o}(t) - q_0(t) = O(1/\omega)$. Indeed, given that $\nabla g$ has the maximum rank, $\nabla g(q_0(t))$ spans the normal section (i.e., the subspace perpendicular to the tangent subspace) of the constraint manifold, in which $q^{o}(t) - q_0(t)$ also lies. Moreover, $g$ restricted to each normal section is an isomorphism, and both the restricted map and its inverse have bounded norms due to the boundedness of $q^{o}$ (i.e., compactness of the solution space)—this is why $g(q^{o}) = O(1/\omega)$ implies $q^{o}(t) - q_0(t) = O(1/\omega)$.

The idea is that since $q^{o}$ is close enough, the constraint manifold can be locally viewed as a flat subspace, and F-convergence for this case has been proved in Lemma A.2. More precisely, there exists a linear isomorphism $A_{q_0(t)}$, such that

$$A_{q_0(t)} (q^{o}(t) - q_0(t)) = \begin{bmatrix} 0 \\ y \end{bmatrix},$$

(A.9)

where $y$ is a vector with codimension of $g^{-1}(0)$ and $0$ is a null vector.

For $q^{o}(s)$ with $s - t = O(1/\omega)$, we will have a full-dimensional representation:

$$A_{q_0(t)} (q^{o}(s) - q_0(t)) = \begin{bmatrix} x \\ y \end{bmatrix},$$

(A.10)

and $x$ and $y$ will, respectively, be the slow and fast variables, representing the constrained dynamics and fluctuations away from the constraint manifold (analogous to Lemma A.2). This is because

$$\frac{d^2}{ds^2} \begin{bmatrix} x \\ y \end{bmatrix} = A_{q_0(t)} \left( -\nabla V(q^{o}(s)) - \omega^2 g(q^{o}(s)) \nabla g(q^{o}(s)) \right)$$

$$= \begin{bmatrix} f_1(x, y) \\ f_2(x, y) \end{bmatrix} + \begin{bmatrix} \mathcal{O}(1) \\ -\omega^2 \tilde{g}(y) \nabla \tilde{g}(y) + \mathcal{O}(1) \end{bmatrix},$$

(A.11)
where \( f_1 \) and \( f_2 \) are defined as \( A_{q_0(t)}(-\nabla V(q^\omega(s))) \). The \( O(1) \) in the first row of the right-hand side of (A.11) is because \( A_{q_0(t)} \) rotates the normal section to the \( y \)-direction, i.e.,

\[
A_{q_0(t)} \nabla g(q^\omega(s)) = A_{q_0(t)}(\nabla g(q_0) + O(1/\omega)) = \begin{bmatrix} 0 \\ * \end{bmatrix} + O(1/\omega),
\]

(A.12)

where \( * \) is some nonzero expression, and certainly \( O(1/\omega) = O(1) \).

The \( O(1) \) in the second row of the right-hand side of (A.11) can also be intuitively obtained by using an analogous geometric argument, together with Taylor expansion.

Since (A.11) corresponds to the locally flat system (A.1), Lemma (A.2) proved the existence of an equivalent Lagrange multiplier as well as the F-convergence towards it. Moreover, (A.11) and the global dynamics near the curved constraint manifold (1.7) is linked via a coordinate transformation \( q_0 \mapsto A_{q_0}(q^\omega - q_0) \), which, naturally, is slowly varying as \( q_0 \) changes. Since averaging via F-convergence (Tao et al., 2010, Theorem 1.2) still works if the slow and fast variables are images of the original variable under a slowly varying diffeomorphism, the global dynamics (1.7) is F-convergent to a solution of (1.9). Note that \( g(q(t)) = 0 \) in (1.9) is automatically satisfied, because \( \lim_{\omega \to \infty} g(q^\omega(t)) = 0 \).

Finally, the solution to (1.9) is also the solution to (1.6). This is by the existence and uniqueness of the solution to differential algebraic equations with initial conditions.

(Only main lines of the proof are provided; details are similar to analysis in Tao et al., 2010, 2011.)

\[ \square \]

Remark A.3 The above proofs show that \( g(q^\omega(s)) \) is not only \( O(\omega^{-1}) \), but also \( O(\omega^{-2}) \).