Long Time Accurate Simulations of Mechanical Systems with Constraints.

Charles Cuell

CMPT 898

April 22, 2005

Contents

1. Abstract

In this paper we outline a technique for developing variational integrators for Lagrangian systems with constraints linear in the velocity. An example shows the excellent potential this technique has for producing long time accurate and geometrically faithful integrators.

The general example of constructing a constrained variational from any explicit method, such as Euler or Runge–Kutta is given here.

Routh's sphere is introduced as an example of a rolling rigid body with a non-holonomic constraint that we would like to develop an integrator for.

2. Classical Lagrangian Mechanics

We briefly describe the mathematical setting for Lagrangian mechanics and show how to obtain the Euler–Lagrange differential equations of motion. For a more detailed description of the classical formulation of Lagrangian mechanics, see, for example, Marion and Thornton [8].

Configuration space is the set of all possible configurations of a system. This can be as simple as the set $\mathbb{R}^3$ for a free particle or as complicated as a product of Lie groups for coupled rigid bodies. The configuration space of a system is traditionally denoted by Q. We restrict our attention to systems for which $Q = \mathbb{R}^n$ and write a point as $(q^1, q^2, \ldots, q^n)$.

Phase space is the combined set of all possible configurations and all possible velocities of a system. It is the tangent space of Q, denoted TQ. When $Q = \mathbb{R}^n$, $TQ = \mathbb{R}^{2n}$. The points of TQ will be written $(q^1, q^2, \ldots, q^n, \dot{q}^1, \dot{q}^2, \ldots, \dot{q}^n)$. For a physical motion of the system, a phase point specifies its actual configuration and its actual velocity.

Example: Two dimensional simple harmonic oscillator. This system consists of a mass in the plane, m, attached to two mass-less springs with spring constant k. The configuration space is $\mathbb{R}^2$, where we will write $q^1 = x$ and $q^2 = y$.

Phase space is $\mathbb{R}^4$, where each point is specified as $(x, y, \dot{x}, \dot{y})$. See figure (2). $\diamondsuit$

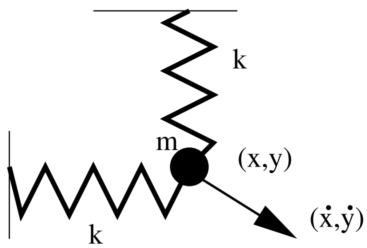

FIGURE 1. Two dimensional simple harmonic oscillator

Let the kinetic energy of a system be denoted by T and the potential energy by U. The *Lagrangian* for a system is defined to be $L = T - U$.

Example: Two dimensional simple harmonic oscillator continued. The kinetic energy of the mass is $T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2)$. The potential energy is $U = \frac{k}{2}(x^2 + y^2)$. This gives the Lagrangian as

$$L = T - U = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - \frac{k}{2}(x^2 + y^2). \tag{1}$$

$\diamondsuit$

Let $q(t)$ be a curve in $\mathbb{R}^n$ on the interval $[\,t_a, t_b\,]$ and $\dot{q}(t)$ its derivative. The *action* is given by

$$S[q(t)] = \int_{t_a}^{t_b} L(q(t), \dot{q}(t))dt. \tag{2}$$

Note that the action is a line integral over the curve $(q(t), \dot{q}(t))$ in $\mathbb{R}^{2n}$.

The calculus of variations provides a necessary condition for the curve $(q(t), \dot{q}(t))$ to be a physical trajectory. See for example Gelfand and Fomin [5] for the theory of variational calculus. The *variational principle* requires that the curve $q(t)$ be a critical point of the action function with respect to the variational derivative when the endpoints, $q(t_a)$ and $q(t_b)$ are fixed. That is, we search for trajectories such that

$$\delta S[q(t)]\delta q(t) = 0, \tag{3}$$

where δS is the variation of S and $\delta q(t)$ is the variation of $q(t)$. We typically suppress the t in the notation for $\delta q(t)$ and write only δq. In what follows a repeated index indicates a sum, so

$$\frac{\partial L}{\partial \dot{q}^i}\delta q^i \quad \text{means} \quad \sum_{i=0}^{n} \frac{\partial L}{\partial \dot{q}^i}\delta q^i.$$

The fixed endpoint condition requires $\delta q(t_a) = \delta q(t_b) = 0$. The variation of S is

$$\begin{aligned}
\delta S[q(t)]\delta q &= \int_{t_a}^{t_b} \delta L(q(t)\dot{q}(t))dt \\
&= \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot{q}^i}\frac{d}{dt}\delta q^i \right) dt \\
&= \left.\frac{\partial L}{\partial \dot{q}^i}\delta q^i\right|_{t_a}^{t_b} + \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} \right) \delta q^i \, dt.
\end{aligned} \tag{4}$$

The last equality follows by an integration by parts. By the fixed endpoint condition, the first term in equation (4) vanishes. Then $\delta S[q(t)]\delta = 0$ only if

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = 0 \quad i = 1 \ldots n. \tag{5}$$

Equations (5) are called the *Euler–Lagrange* equations. They are a system of second–order differential equation on configuration space $\mathbb{R}^n$ in the variables $(q^1, q^2, \ldots, q^n)$.

Example: Two dimensional simple harmonic oscillator continued. The Lagrangian for the two dimensional simple harmonic oscillator is given in equation (1). Configuration space is $\mathbb{R}^2$ with points denoted (x, y), so there are two second–order differential equations: one in x and one in y. The equation for x is:

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = 0,$$

$$-kx - \frac{d}{dt}(m\dot{x}) = 0,$$

$$kx + m\ddot{x} = 0,$$

where $\ddot{x} = \dfrac{d^2x}{dt^2}$. The y equation is derived similarly as

$$ky + m\ddot{y} = 0.$$

Writing $\omega = \sqrt{\frac{k}{m}}$, the solution for initial conditions $(x_0, y_0, \dot{x}_0, \dot{y}_0)$ is

$$x(t) = x_0 \cos(\omega t) + \dot{x}_0 \sin(\omega t),$$

$$y(t) = y_0 \cos(\omega t) + \dot{y}_0 \sin(\omega t).$$

$\Diamond$

2.1. **The Euler–Lagrange Equations as a First Order System.** To write the Euler–Lagrange equations as a set of first order equations on phase space, $\mathbb{R}^{2n}$ assume that the kinetic energy as a function on phase space $\mathbb{R}^{2n}$ is

$$T(q,\dot{q}) = \frac{1}{2}\dot{q}^T M(q)\dot{q},$$

where M is an $n \times n$ positive definite symmetric q dependent matrix. Written in coordinates, it looks like

$$T(q,\dot{q}) = \frac{1}{2}M_{ij}(q)\dot{q}^i\dot{q}^j.$$

Also, assume that the potential energy is a function of q only. Then the Euler–Lagrange equations are

$$\begin{aligned}
\frac{\partial}{\partial q^i}(T-U) - \frac{d}{dt}\frac{\partial}{\partial \dot{q}^i}(T-U) &= 0,\\
\frac{\partial T}{\partial q^i} - \frac{\partial U}{\partial q^i} - \frac{d}{dt}(\frac{\partial T}{\partial \dot{q}^i}) &= 0,\\
\frac{\partial T}{\partial q^i} - \frac{\partial U}{\partial q^i} - \frac{\partial^2 T}{\partial q^j \partial \dot{q}^i}\dot{q}^j - \frac{\partial^2 T}{\partial \dot{q}^j \partial \dot{q}^i}\ddot{q}^j &= 0,
\end{aligned}$$

where $i,j = 1\ldots n$ and $\ddot{q} = \dfrac{d^2 q}{dt^2}$. To simplify this discussion, assume that $\dfrac{\partial T}{\partial q^i} = 0$. Then, $\dfrac{\partial^2 T}{\partial \dot{q}^j \partial \dot{q}^i} = M_{ji}$ and the Euler–Lagrange equations are

$$-\frac{\partial U}{\partial q^i} - M_{ij}\ddot{q}^j = 0.$$

Write the inverse of M in components as M^{ij} and solve for $\ddot{q}$,

$$\ddot{q}^j = -M^{ij}\frac{\partial U}{\partial q^i}.$$

The Euler–Lagrange equations, when T does not depend on q, are then

$$\frac{dq^i}{dt} = \dot{q}^i,$$
$$\frac{d\dot{q}^i}{dt} = -M^{ij}\frac{\partial U}{\partial q^j}.$$

This first order system corresponds to the flow of the vector field X_E on $\mathbb{R}^{2n}$ where

$$X_E(q^1,\ldots,q^n,\dot{q}^1,\ldots,\dot{q}^n) = (\dot{q}^1,\ldots,\dot{q}^n,-M^{1i}\frac{\partial U}{\partial q^i},\ldots,-M^{ni}\frac{\partial U}{\partial q^i}).$$

The flow of X_E is denoted F_t^L. It is a smooth map of $\mathbb{R}^{2n}$ to $\mathbb{R}^{2n}$ that evolves a point along solutions for a time t. F_t^L can be written in the form

$$F_t^L(q,\dot{q}) = (F_t^q(q,\dot{q}), F_t^{\dot{q}}(q,\dot{q})),$$

where F_t^q evolves the q components and $F_t^{\dot{q}}$ evolves the $\dot{q}$ components. $F_t^q(q_0,\dot{q}_0)$ corresponds to solutions on Q of the Euler–Lagrange equations with initial conditions $(q_0,\dot{q}_0)$.

2.2. **Summary.** The process of Lagrangian mechanics involves the following steps:

(1) Find an appropriate configuration space, Q.

(2) Write the Lagrangian, $L = T - U$.

(3) Write the Euler–Lagrange equations (5).

(4) Solve the Euler–Lagrange equations.

The last step of the process is often problematic as the equations can be quite complicated. In such cases where a solution is difficult, or impossible, to find, we rely on numerical methods.

3. Constrained Lagrangian Mechanics

Given a mechanical system, other mechanical systems can obtained by constraining the phase points to a subset of phase space. For example, a pendulum system is obtained from a free particle system by constraining the mass to a circle of fixed radius. The variational principle, as given above, is only applicable to unconstrained systems. To include constraints in the theory, the variational principle must be augmented with D'Alembert's principle.

See Rosenberg [11] for a complete discussion of D'Alembert's principle. We follow Marsden and Ratiu [9] in our treatment of the related Lagrange–D'Alembert principle.

3.1. **Lagrange–D'Alembert Principle.** The variations, $\delta q(t)$, as introduced in the variational principle equation (4), are meant to represent the directions in which a configuration is allowed to change at time t. For the two dimensional simple harmonic oscillator, and all other similar systems, the variations are arbitrary. This freedom is what allows us to conclude that

$$\int_{t_a}^{t_b} \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}\right) \delta q^i \, dt = 0 \quad \Rightarrow \quad \frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = 0, \tag{6}$$

hence obtaining the Euler–Lagrange equations.

Let $\mathcal{D}$ be a d–dimensional linear subspace of $\mathbb{R}^n$, where $d < n$, and let $M = \{\,(q, \dot{q}) \,|\, \dot{q} \in \mathcal{D}\,\}$. Note that $\mathcal{D}$ is a linear subspace of $\mathbb{R}^n$ whereas M is subset of $\mathbb{R}^{2n}$ that is linear in the velocity variables. The sets $\mathcal{D}$ and M are equivalent, but serve different purposes. $\mathcal{D}$ restricts variations and velocities and M restricts phase points. Also note that the specification of $\mathcal{D}$ gives M. When δq is restricted to lie in $\mathcal{D}$ the implication in (6) need not hold. Adding constraints removes degrees of freedom in choosing the δq^i so that equation (3) may hold for many curves. The Lagrange–D'Alembert principle directly removes this arbitrariness by forcing the following:

$$\left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}\right)\delta q^i = 0, \tag{7}$$

$$\delta q \in \mathcal{D}, \tag{8}$$

$$(q(t), \dot{q}(t)) \in M. \tag{9}$$

Since the Euler–Lagrange equations (7) vanish on the linear subspaces of $\mathcal{D}$, equations (7) to (9) can be replaced by

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = \lambda_a \phi_i^a, \tag{10}$$

$$\phi_j^a(q)\dot{q}^j = 0, \tag{11}$$

where $a = 1 \ldots n - d$ and each ϕ^a is row vector orthogonal to $\mathcal{D}$. Each λ^a is a Lagrange undetermined multiplier. As first order equations for the Lagrangian $L = \frac{1}{2}\dot{q}^T M \dot{q} - V(q)$, equations (10) and (11) are

$$\frac{dq^i}{dt} = \dot{q}^i, \tag{12}$$

$$\frac{d\dot{q}^i}{dt} = -M^{ij}\frac{\partial V}{\partial q^j} + \lambda_a \phi_i^a, , \tag{13}$$

$$\phi_i^a(q)\dot{q}^i = 0. \tag{14}$$

Equations (12) and (13) are an index two set of differential–algebraic equations. See Ascher and Petzold [2] for more details. If the constraints are holonomic, then equation (14) can be integrated and replaced by $f^a(q) = 0$, where $\frac{df^a}{dt} =$

$\phi^a_i(q)\dot{q}^i$. This transforms the equations to an index three set of differential–algebraic equations.

Example: Two dimensional simple harmonic oscillator constrained to the unit circle. Let e_1 and e_2 be the standard unit vectors in $\mathbb{R}^2$. To constrain the mass to the unit circle $\mathcal{D} = \operatorname{span}\{\, -ye_1 + xe_2 \,\}$. Construct M as follows:

$$(\dot{x}, \dot{y}) \in \mathcal{D} \Leftrightarrow (\dot{x}, \dot{y}) = \alpha(-y, x) \ \text{ for some } \alpha.$$

Eliminating α, we get $M = \{\, (x, y, \dot{x}, \dot{y}) \,|\, x\dot{x} + y\dot{y} = 0 \,\}$. That is, M is the zero level set of the function $c(x, y, \dot{x}, \dot{y}) = x\dot{x} + y\dot{y}$.

The Lagrangian is still kinetic energy plus potential energy:

$$L(x, y, \dot{x}, \dot{y}) = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - \frac{k}{2}(x^2 + y^2).$$

The variation $\delta q = (\delta x, \delta y)$ has to be in $\mathcal{D}$, so we must have $x\delta x + y\delta y = 0$. Equations (7), (8), (9) are then, respectively,

$$(kx + m\ddot{x})\delta x + (ky + m\ddot{y})\delta y = 0, \tag{15}$$

$$x\delta x + y\delta y = 0, \tag{16}$$

$$x\dot{x} + y\dot{y} = 0. \tag{17}$$

Multiply equation (15) through by x and substitute equation (16) into it. We obtain,

$$(kx + m\ddot{x})(-y\delta y) + (ky + m\ddot{y})x\delta y = 0. \tag{18}$$

The δy variation is arbitrary, so that

$$-y(kx + m\ddot{x}) + (ky + m\ddot{y})x = 0. \tag{19}$$

Differentiate equation (17) with respect to t and use it to eliminate $\ddot{x}$ in equation (19) to get

$$\ddot{y} + (\dot{x}^2 + \dot{y}^2)y = 0, \tag{20}$$

where we used the fact that $x^2 + y^2 = 1$. Using the constraint equation (17) eliminate $\dot{x}$ in equation (20). Also, since $x^2 + y^2 = 1$, eliminate x^2 to get the final equation:

$$(1 - y^2)\ddot{y} + y\dot{y} = 0.$$

This is a second order equation in y. $\diamondsuit$

Example: Nonholonomic free particle. The nonholonomic free particle is a free particle in $\mathbb{R}^3$ subject to the constraint $\dot{z} = y\dot{x}$. This gives

$$M = \{\, (x, y, z, \dot{x}, \dot{y}, \dot{z}) \mid \dot{z} - y\dot{x} = 0 \,\},$$

$$\mathcal{D} = \operatorname{span}\{\, e_1 + ye_3,\ e_2 \,\}.$$

Following the same routine as in the two dimensional simple harmonic oscillator, the equations of motion of motion are:

$$\ddot{x} + \frac{y}{1 + y^2}\dot{x}\dot{y} = 0,$$

$$\ddot{y} = 0.$$

$\Diamond$

A fundamental difference exists between the examples of the constrained two dimensional simple harmonic oscillator and the nonholonomic free particle in that the constraint of the former can be integrated, whereas the latter cannot. This means that there is nothing more than can be done with the equations for the nonholonomic free particle. For the two dimensional simple harmonic oscillator, however, there is a change of coordinates that adapts the equations to the constraints and results in a lower dimensional Lagrangian system. The following example illustrates this.

Example: Two dimensional simple harmonic oscillator constrained to the unit circle continued. Change to polar coordinates (r, θ).

$$x = r\cos\theta,$$

$$y = r\sin\theta.$$

Then, differentiating, $x\dot{x} + y\dot{y} = \dot{r} = 0$. Since the mass is on the unit circle, $\dot{r} = 0$ integrates to $r = 1$ for the constraint in polar coordinates. This gives $\delta q = (\delta r, \delta\theta) = (0, \delta\theta)$, so that $\delta\theta$ is free and the equations of motion from the variational principle in polar coordinates are

$$\frac{\partial L}{\partial \theta} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} = 0,$$

where $L(r, \theta, \dot{r}, \dot{\theta}) = \frac{m}{2}\dot{\theta}^2 - \frac{k}{2}$. $\lozenge$

Mechanical systems with integrable constraints are called *holonomic*, while those with non integrable constraints are called *nonholonomic*. Note that it is not necessary that each constraint be integrable, but that the constraint set, M be the phase space for some configuration space. These constraint sets need not look like $\mathbb{R}^{2k}$, even if they are integrable, as the example with the two dimensional simple harmonic oscillator illustrates.

Holonomic systems can be reduced to mechanical systems on configuration spaces of the same dimension as $\mathcal{D}$, while nonholonomic systems must be analysed on the full phase space. In essence, this allows holonomic systems to be treated as unconstrained systems by adapting the coordinate system to the constraints.

4. Discrete Lagrangian Mechanics

In this section we show how to use curves to develop a discrete Lagrangian mechanical theory in analogy with the continuous theory.

Let $\psi_h : (-a, a) \times U \to Q$ be a differentiable map for some $a \in \mathbb{R}$ and U open in TQ. Let $\alpha^+ : [0, a) \to [0, \infty)$ and $\alpha^- : [0, a) \to (-\infty, 0]$ have the property $\alpha^+(t) - \alpha^-(t) = t$ for all t in $[-a, a]$. Require ψ_h to be such that

(1) $\psi_h(0, (q, \dot{q})) = q$,

(2) $\dfrac{\partial \psi_h}{\partial t}(0, (q, \dot{q})) = \dot{q}$,

(3) $\psi_h(t, (q, \dot{q})) = q + t\dot{q} + \mathcal{O}(t^2)$.

These conditions guarantee that the image of ψ_h is a differentiable curve through q with tangent vector $\dot{q}$ at $t = 0$. Define the maps

$$\partial_h^+(q, \dot{q}) = \psi_h(\alpha^+(h), (q, \dot{q})) \quad \partial_h^-(q, \dot{q}) = \psi_h(\alpha^-(h), (q, \dot{q})).$$

The number a must be small enough that the map $\partial_h^{\pm}(q,\dot{q}) = (\partial_h^+(q,\dot{q}), \partial_h^-(q,\dot{q}))$ be invertible for each $(q,\dot{q})$ in U when $h \in (0,a]$.

The map ψ_h is a curve in configuration space $\mathbb{R}^n$ for every $(q,\dot{q})$ in phase space $\mathbb{R}^{2n}$. The maps ∂_h^+ and ∂_h^- are then the endpoints of the curve over the interval $[\alpha^-(h), \alpha^+(h)]$.

Let L be the Lagrangian for a mechanical system and define the discrete Lagrangian by

$$L_d(q,\dot{q}) = \frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)} L(\psi_h(q(t),\dot{q}(t)), \frac{\partial \psi_h}{\partial t}(t,(q,\dot{q})))\, dt + \mathcal{O}(h^2).$$

The discrete Lagrangian is, therefore, an approximation to the average of the continuous Lagrangian on the curve given by ψ_h over the interval $[\alpha^-(h), \alpha^+(h)]$.

Let $q_d = \{q_k\}_{k=0}^N$ and $\dot{q}_d = \{\dot{q}_k\}_{k=0}^N$ be $N+1$ element sequences in $\mathbb{R}^n$. Require that $\dot{q}_d$ be such that $\partial_h^+(q_k,\dot{q}_k) = q_{k+1}$ and $\partial_h^-(q_k,\dot{q}_k) = q_k$. This forces the endpoints of the curve segments $\psi_h(t,(q_k,\dot{q}_k))$ and $\psi_h(t,(q_{k+1},\dot{q}_{k+1}))$ to join up. Define the *discrete action* on the sequence q_d to be

$$S_h(q_d) = \sum_{k=0}^{N} L_d(q_k,\dot{q}_k)h.$$

This is an approximation of the continuous action in equation (2) by approximating the integral. Denote the derivative operator by D. Let $\delta q_d = \{\delta q_k\}_{k=0}^N$ be a sequence of vectors such that the vector δq_k has base point at q_k. We define another sequence of vectors $\delta \dot{q}_d = \{\delta \dot{q}_k\}_{k=0}^N$ that satisfy the conditions

$$D\partial_h^+(q_k,\dot{q}_k)(\delta q_k, \delta \dot{q}_k) = \delta q_{k+1}, \quad D\partial_h^-(q_k,\dot{q}_k)(\delta q_k, \delta \dot{q}_k) = \delta q_k. \tag{21}$$

A sequence of vectors $\{\, \delta q_k \,,\, \delta \dot{q}_k \,\}$ that satisfy equation (21) are called a *discrete variation.* The *discrete variational principle* is

$$DS_d(q_d)\delta q_d = 0, \tag{22}$$

$$q_0, q_N \text{ fixed}, \tag{23}$$

$$\partial_h^+(q_k, \dot{q}_k) = \partial_h^-(q_{k+1}, \dot{q}_{k+1}). \tag{24}$$

The resulting equations, for each $k = 1, \ldots N-1$ are

$$D\partial_h^-(q_k, \dot{q}_k)(\delta q_k, \delta \dot{q}_k) = 0, \tag{25}$$

$$D\partial_h^+(q_{k+1}, \dot{q}_{k+1})(\delta q_{k+1}, \delta \dot{q}_{k+1}) = 0, \tag{26}$$

$$D\partial_h^+(q_k, \dot{q}_k)(\delta q_k, \delta \dot{q}_k) = D\partial_h^-(q_{k+1}, \dot{q}_{k+1})(\delta q_{k+1}, \delta \dot{q}_{k+1}), \tag{27}$$

$$DL_d(q_k, \dot{q}_k)(\delta q_k, \delta \dot{q}_k) + DL_d(q_{k+1}, \dot{q}_{k+1})(\delta q_{k+1}, \delta \dot{q}_{k+1}) = 0, \tag{28}$$

$$\partial_h^+(q_k, \dot{q}_k) = \partial_h^-(q_{k+1}, \dot{q}_{k+1}). \tag{29}$$

Equations (25) and (26) are the fixed endpoint conditions equation (23). Equation (27) guarantees that the sequence $\{\, \delta q_k \,,\, \delta \dot{q}_k \,\}$ is a discrete variation. Equation (29) is equation (22) written out for the kth term.

To obtain evolution equations, the system in equations (25) to (29) must be solved in two stages. The first stage is to find n linearly independent variations from equations (25) to (27). These equations are a set of $3n$ linear equations in the $4n$ unknowns $(\delta q_k, \delta \dot{q}_k, \delta q_{k+1}, \delta \dot{q}_{k+1})$ that hold for $(q_k, \dot{q}_k, q_{k+1}, \dot{q}_{k+1})$. When h is small enough to guarantee that $\partial_h^{\pm}$ is invertible, an n parameter set of solutions can be guaranteed for this system.

To obtain the discrete evolution equations, choose a set of n linearly independent solutions. These go into the $(\delta q_k, \delta \dot{q}_k)$ and $(\delta q_{k+1}, \delta \dot{q}_{k+1})$ of equation (28) one at a time to produce n equations. With equation (29) there are $2n$ equations in the $2n$ unknowns $(q_{k+1}, \dot{q}_{k+1})$.

The following example shows how to obtain the symplectic Euler method using this technique.

Example: Symplectic Euler. Let M be a constant mass matrix and $V(q)$ a potential function. Then $L(q, \dot{q}) = \frac{1}{2}\dot{q}^T M q - V(q)$. Let $\psi_h(t, (q, \dot{q})) = q + t\dot{q}$, $\alpha^+(t) = t$ and $\alpha^-(t) = 0$. Use a left endpoint approximation to

$$\frac{1}{h}\int_0^h L(q + t\dot{q}, \dot{q})\, dt$$

to obtain the discrete Lagrangian

$$L_d(q, \dot{q}) = L(q, \dot{q}).$$

Then $\partial_h^+(q, \dot{q}) = q + h\dot{q}$ and $\partial_h^-(q, \dot{q}) = q$. Compute the derivatives of these maps to get

$$D\partial_h^+(q, \dot{q}) = \begin{bmatrix} I & hI \end{bmatrix}, \quad D\partial_h^-(q, \dot{q}) = \begin{bmatrix} I & 0 \end{bmatrix}$$

where I is the $n \times n$ identity matrix and 0 is the $n \times n$ zero matrix. This gives

$$D\partial_h^-(q_k, \dot{q}_k)(\delta q_k, \delta \dot{q}_k) = \delta q_k,$$

$$D\partial_h^+(q_{k+1}, \dot{q}_{k+1})(\delta q_{k+1}, \delta \dot{q}_{k+1}) = \delta q_{k+1} + h\delta \dot{q}_{k+1}.$$

Then equations (25), (26), and (27) give

$$\delta q_k = 0 \text{ and } \delta \dot{q}_k \text{ arbitrary,}$$

$$\delta q_{k+1} = -h\delta \dot{q}_{k+1},$$

$$h\delta \dot{q}_k = \delta q_{k+1}.$$

Parametrise the solution set by $\delta \dot{q}_k$ to get

$$\delta q_k = 0,\ \delta \dot{q}_k \text{ arbitrary, } \delta q_{k+1} = h\delta \dot{q}_k,\ \delta \dot{q}_{k+1} = -\delta \dot{q}_k.$$

Then, equation (28) is

$$\frac{\partial L_d}{\partial \dot{q}^i}(q_k, \dot{q}_k)\delta \dot{q}_k^i - \frac{\partial L_d}{\partial q^i}(q_{k+1}, \dot{q}_{k+1})h\delta \dot{q}_k^i - \frac{\partial L_d}{\partial \dot{q}^i}(q_{k+1}, \dot{q}_{k+1})\delta \dot{q}_k^i = 0.$$

For each i, let

$$\delta \dot{q}_k^i = 1, \quad \delta \dot{q}_i^j = 0 \text{ for} j = 1 \ldots n, j \neq i. \tag{30}$$

Equation (30) selects n linearly independent variations, and gives the set of equations

$$M_{ij}\dot{q}_k^j - h\frac{\partial V}{\partial q^i}(q_{k+1}) - M_{ij}\dot{q}_{k+1}^j = 0.$$

Rearranged, and with equation (29)

$$\dot{q}_{k+1}^i = \dot{q}_k^i - hM^{ij}\frac{\partial V}{\partial q^j}(q_{k+1}), \tag{31}$$

$$q_{k+1} = q_k + h\dot{q}_k, \tag{32}$$

for $i = 1 \ldots n$. These equations are formally implicit in the first set and explicit in the second set and are called the *symplectic Euler* method. See Hairer *et. al.* [6] for more on the symplectic Euler method. Of course, one can substitute equation 32 into equation 31 to obtain a fully explicit set of equations. $\Diamond$

Example: Two dimensional simple harmonic oscillator – symplectic Euler integration. Here we have

$$M = \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix}, \quad M^{-1} = \begin{bmatrix} \frac{1}{m} & 0 \\ 0 & \frac{1}{m} \end{bmatrix}, \quad V(x,y) = \frac{k}{2}(x^2 + y^2).$$

Then, equations (31) and (32) are

$$\dot{x}_{k+1} = \dot{x}_k - \frac{k}{m} h x_{k+1},$$
$$\dot{y}_{k+1} = \dot{y}_k - \frac{k}{m} h y_{k+1},$$
$$x_{k+1} = x_k + h\dot{x}_k,$$
$$y_{k+1} = y_k + h\dot{y}_k.$$

Or, equivalently,

$$\dot{x}_{k+1} = \dot{x}_k - \frac{k}{m} h (x_k + h\dot{x}_k),$$
$$\dot{y}_{k+1} = \dot{y}_k - \frac{k}{m} h (y_k + h\dot{y}_k).$$

$\Diamond$

4.1. **Two Dimensional Simple Harmonic Oscillator.** In this section, we compare the symplectic Euler method with the classic fourth order Runge–Kutta and *ode45* package from MatLab for the two dimensional simple harmonic oscillator integrated over 20000 oscillations. Over this long time, the symplectic Euler provides considerably better results than the other methods, even though it is only first order. The step size for the Runge–Kutta and the symplectic Euler methods is 0.25.

First, recall the equations of motion,

$$kx + m\ddot{x} = 0, \quad ky + m\ddot{y} = 0.$$

Energy and angular momentum are, respectively,

$$E = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) + \frac{k}{2}(x^2 + y^2), \quad p = \dot{y}x - \dot{x}y.$$

Both of these quantities are conserved, since

$$\frac{dE}{dt} = 0, \quad \frac{dp}{dt} = 0.$$

See figure (4.1) for the plots of energy and momentum. It is clear from this picture that the symplectic Euler method preserves these natural invariants while the others do not. The initial conditions are $(x(0), y(0), \dot{x}(0), \dot{y}(0)) = (1, 0, 0, 1)$, which produces the solution curve $(x(t), y(t)) = (\cos t, \sin t)$, which is a circle of radius 1 when x and y are plotted against each other. Figure (4.1) plots the errors

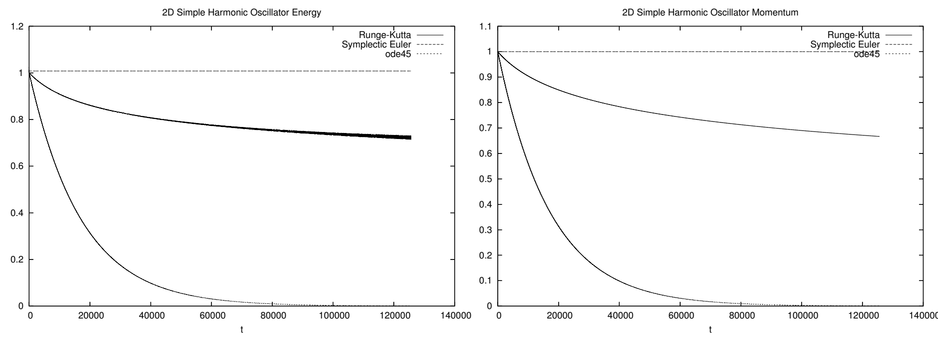

FIGURE 2

in the various methods for $x^2 + y^2 - 1$. Again, the symplectic Euler method gives a

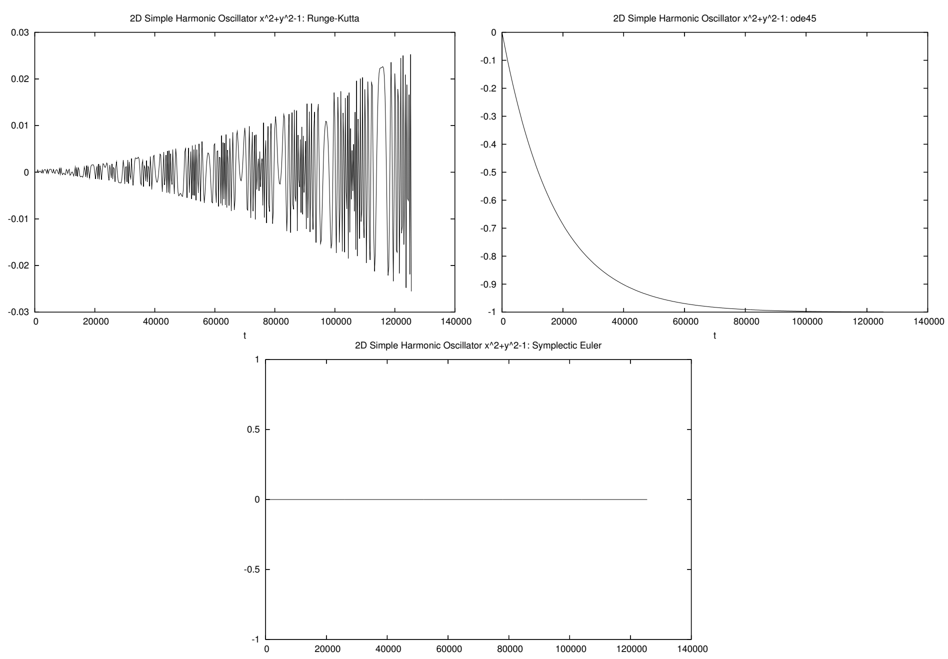

FIGURE 3

superior plot. It should be noted that none of these methods are explicitly designed to preserve energy, momentum or the unit circle.

5. Discrete Constrained Lagrangian Mechanics

Let $\mathcal{D}$ be a constraint, as described in the Lagrange–D'Alembert principle from section 3.1. To implement constraints in discrete mechanics,introduce the differentiable mapping $\phi_h : (-a, a) \times U \to Q$, for an open subset $U \subset TQ$ and require that ϕ_h satisfy the following conditions:

(1) $\phi_h(0, (q, \dot{q}) = q$,

(2) $\frac{\partial \phi_h}{\partial t}(0, (q, \dot{q})) = \dot{q}$,

(3) $\frac{\partial \phi_h}{\partial t}(t, (q, \dot{q})) \in \mathcal{D}$ for $t \in [\alpha^-(h), \alpha^+(h)]$.

For a given $(q, \dot{q})$, the image of ϕ_h is a curve for which every velocity vector satisfies the constraint. Define $\phi_h^+(q, \dot{q}) = \phi_h(\alpha^+(h), (q, \dot{q}))$ and $\phi_h^-(q, \dot{q}) = \phi_h(\alpha^-(h), (q, \dot{q}))$.

5.1. **Discrete Lagrange–D'Alembert Principle.**

The discrete variational principle needs to be augmented in order to constrain it to $\mathcal{D}$. The *discrete Lagrange–D'Alembert* principle is

$$DS_d(q_d)\delta q_d = 0, \tag{33}$$

$$q_0, q_N \text{ fixed}, \tag{34}$$

$$\partial^+(h)(q_k, \dot{q}_k) = \partial^-(q_{k+1}, \dot{q}_{k+1}), \tag{35}$$

$$\phi_h^+(\bar{q}_k, \dot{\bar{q}}_k) = \partial_h^+(q_k, \dot{q}_k), \quad \phi_h^-(\bar{q}_k, \dot{\bar{q}}_k) = \partial_h^-(q_k, \dot{q}_k) \ \text{ some } (\bar{q}_k, \dot{\bar{q}}_k), \tag{36}$$

(37)

$$\phi_h^+(\bar{q}_{k+1}, \dot{\bar{q}}_{k+1}) = \partial_h^+(q_{k+1}, \dot{q}_{k+1}), \quad \phi_h^-(\bar{q}_{k+1}, \dot{\bar{q}}_{k+1}) = \partial_h^-(q_{k+1}, \dot{q}_{k+1}) \ \text{ some } (\bar{q}_{k+1}, \dot{\bar{q}}_{k+1}),$$

$$D\partial_h^+(q_k, \dot{q}_k)(\delta q_k, \delta \dot{q}_k) \in \mathcal{D}. \tag{38}$$

Equations (33), (34) and (35) are the usual ones for the discrete variational principle. Equations (36) and (37) force the unconstrained curves ψ_h to have endpoints that can be joined by curves that satisfy the constraint. Note that ψ_h need not satisfy the constraint, but all points in the evolution will be reachable by curves that do satisfy the constraint. Also, the points $\bar{q}$ and $\dot{\bar{q}}$ are not necessarily points in the evolution, but auxiliary points that must be found in order to select the correct curve. The final equation (38) forces the variation to be in $\mathcal{D}$ for each k.

The discrete evolution equations are

$$D\partial_h^-(q_k, \dot{q}_k)(\delta q_k, \delta \dot{q}_k) = 0, \tag{39}$$

$$D\partial_h^+(q_{k+1}, \dot{q}_{k+1})(\delta q_{k+1}, \delta \dot{q}_{k+1}) = 0, \tag{40}$$

$$D\partial_h^+(q_k, \dot{q}_k)(\delta q_k, \delta \dot{q}_k) = D\partial_h^-(q_{k+1}, \dot{q}_{k+1})(\delta q_{k+1}, \delta \dot{q}_{k+1}), \tag{41}$$

$$DL_d(q_k, \dot{q}_k)(\delta q_k, \delta \dot{q}_k) + DL_d(q_{k+1}, \dot{q}_{k+1})(\delta q_{k+1}, \delta \dot{q}_{k+1}) = 0, \tag{42}$$

$$\partial_h^+(q_k, \dot{q}_k) = \partial_h^-(q_{k+1}, \dot{q}_{k+1}), \tag{43}$$

$$\phi_h^+(\bar{q}, \dot{\bar{q}}) = \partial_h^+(q_{k+1}, \dot{q}_{k+1}), \tag{44}$$

$$\phi_h^-(\bar{q}, \dot{\bar{q}}) = \partial_h^-(q_{k+1}, \dot{q}_{k+1}), \tag{45}$$

$$D\partial_h^+(q_k, \dot{q}_k)(\delta q_k, \delta \dot{q}_k) \in \mathcal{D}. \tag{46}$$

As in the unconstrained case, an n parameter set of solutions of the linear equations (39), (40) and (41) can be found. The constraint in equation (46) also must be satisfied, which reduces the dimension of the solution set from n to d, where $n-d$ is the number of constraints. Equations (44) and (45) give $2n$ more equations and add $n+d$ variables $(\bar{q}, \dot{\bar{q}})$. There are only $n+d$ variables because there are only

d degrees of freedom in choosing the $\dot{\bar{q}}$ that guarantee $(q, \dot{\bar{q}})$ satisfy the constraints. In total there are $3n+d$ equations in the $3n+d$ variables $(q_{k+1}, \dot{q}_{k+1}, \bar{q}, \dot{\bar{q}})$.

We emphasise that no assumptions were made on the integrability of the constraints. If the set $\mathcal{D}$ is a holonomic constraint, then the curves ϕ_h will respect the integrability and remain on the same constraint set as the initial conditions. The definition of a discrete holonomic system and theorems regarding the discrete variational principle for holonomic systems are the subject of current research.

The following example illustrates the procedure for generating the discrete constrained equations.

Example: Two dimensional simple harmonic oscillator constrained to the unit circle. The constraint is $x\dot{x}+y\dot{y}=0$, so that $\mathcal{D} = \mathrm{span}\{\ -y\ e_1 + x\ e_2\ \}$. Let

$$\psi_h(t,(q,\dot{q})) = q + t\dot{q},$$

$$\alpha^+(h) = \frac{h}{2},$$

$$\alpha^-(h) = -\frac{h}{2},$$

$$\phi_h(t,(\bar{q},\dot{\bar{q}})) = (\bar{x}\cos(\gamma t) - \bar{y}\sin(\gamma t), \bar{x}\sin(\gamma t) + \bar{y}\cos(\gamma t)),$$

where $(\dot{\bar{x}}, \dot{\bar{y}})$ is decomposed as $\dfrac{\gamma}{\bar{x}^2+\bar{y}^2}(-\bar{y},\bar{x}) + \dfrac{\eta}{\bar{x}^2+\bar{y}^2}(\bar{x},\bar{y})$. The discrete Lagrangian is obtained by a midpoint approximation to

$$\frac{1}{h}\int_{-\frac{h}{2}}^{\frac{h}{2}} L(q+t\dot{q},\dot{q})\, dt$$

to give

$$L_d(q,\dot{q}) = \frac{m}{2}(\dot{x}^2+\dot{y}^2) - \frac{k}{2}(x^2+y^2).$$

Write $\delta v_k = (\delta x_k, \delta y_k, \delta \dot{x}_k, \delta \dot{y}_k)$. For the unconstrained system, $(\delta v_k, \delta v_{k+1})$ satisfies equations (39), (40) and (41) if

$$\delta v_k = (\frac{h}{2}\delta \dot{x}_k, \frac{h}{2}\delta \dot{y}_k, \delta \dot{x}_k, \delta \dot{y}_k),$$
$$\delta v_{k+1} = (\frac{h}{2}\delta \dot{x}_k, \frac{h}{2}\delta \dot{y}_k, -\delta \dot{x}_k, -\delta \dot{y}_k).$$

To satisfy the discrete Lagrange–D'Alembert principle, we also require

$$D\partial_h^+(q_k\dot{q}_k)\delta v_k \in \mathcal{D},$$

which gives

$$\delta v_k = (\frac{h}{2}(y_k + \frac{h}{2}\dot{y}_k), -\frac{h}{2}(x_k+\frac{h}{2}\dot{x}_k), y_k + \frac{h}{2}\dot{y}_k, -(x_k+\frac{h}{2}\dot{x}_k))\delta \dot{x}_k,$$
$$\delta v_{k+1} = (\frac{h}{2}(y_k + \frac{h}{2}\dot{y}_k), -\frac{h}{2}(x_k+\frac{h}{2}\dot{x}_k), -(y_k + \frac{h}{2}\dot{y}_k), (x_k+\frac{h}{2}\dot{x}_k))\delta \dot{x}_k.$$

There is only one parameter, $\delta \dot{x}_k$, in the variations, hence only one discrete Euler–Lagrange equation:

$$dL_d(q_k, \dot{q}_k)\delta v_k + dL_d(q_{k+1}, \dot{q}_{k+1})\delta v_{k+1} = 0.$$

Setting $\delta \dot{x}_k = 1$,

$$\dot{x}_{n-1}(y_{n-1}+\frac{h}{2}\dot{y}_{n-1}) - \dot{y}_{n-1}(x_{n-1}+\frac{h}{2}+\dot{x}_{n-1}) - \dot{x}_n(y_{n-1}+\frac{h}{2}\dot{y}_{n-1}) + \dot{y}_n(x_{n-1}+\frac{h}{2}+\dot{x}_{n-1}) -$$

$$\frac{h}{2}x_{n-1}(y_{n-1}+\frac{h}{2}\dot{y}_{n-1})+\frac{h}{2}y_{n-1}(x_{n-1}+\frac{h}{2}+\dot{x}_{n-1})-$$
$$\frac{h}{2}x_n(y_{n-1}+\frac{h}{2}\dot{y}_{n-1})+\frac{h}{2}(x_{n-1}+\frac{h}{2}+\dot{x}_{n-1})=0.$$

We also have

$$x_n-\frac{h}{2}=x_{n-1}+\frac{h}{2}\dot{x}_{n-1},$$
$$y_n-\frac{h}{2}=y_{n-1}+\frac{h}{2}\dot{y}_{n-1},$$

The constraint equations are

$$x_n+\frac{h}{2}\dot{x}_n=\bar{x}\cos\gamma-\bar{y}\sin\gamma,$$
$$y_n+\frac{h}{2}\dot{y}_n=\bar{x}\sin\gamma+\bar{y}\cos\gamma,$$
$$x_n-\frac{h}{2}\dot{x}_n=\bar{x}\cos\gamma+\bar{y}\sin\gamma,$$
$$y_n-\frac{h}{2}\dot{y}_n=-\bar{x}\sin\gamma+\bar{y}\cos\gamma,$$

where $(\dot{\bar{x}},\dot{\bar{y}})=\frac{\gamma}{-\dot{\bar{y}}^2+\dot{\bar{x}}^2}+\frac{\eta}{\dot{\bar{x}}^2+\dot{\bar{y}}^2}$. Figure (5.1) shows the value of the constraint $x\dot{x}+y\dot{y}$ over 20000 oscillations of the system for the classic fourth order Runge–Kutta and the discrete system developed in this example. It is clear that the method developed here preserves the constraint better than the Runge–Kutta.

Note that the constraint in this example was given as $x\dot{x}+y\dot{y}=0$ rather than in its integrated form $x^2+y^2-r^2=0$ illustrating the fact that this method respects holonomic constraints even when they are not explicitly written in their holonomic

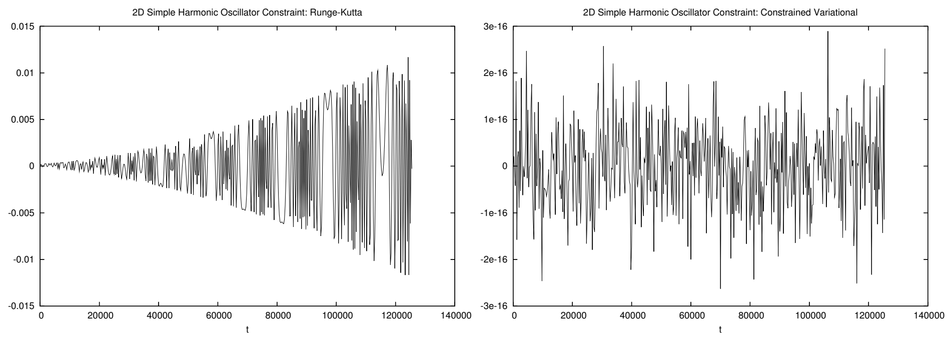

FIGURE 4

form. $\diamondsuit$

6. Discrete Lagrangian Methods from Explicit Methods

Explicit constant time step methods, such as forward Euler or the classic fourth order Runge–Kutta can be used as the basis for constructing discrete Lagrangian integrators. If the system is holonomic, then the resulting integrator is symplectic. See Patrick and Cuell [10] for the details.

An explicit method is a mapping R_h from $\mathbb{R}^{2n}$ to $\mathbb{R}^{2n}$ corresponding to the integration step $(q_k, \dot{q}_k) \mapsto (q_{k+1}, \dot{q}_{k+1})$. The curves are $\psi_h(t, (q, \dot{q})) = \tau_Q \circ R_t(q, \dot{q})$, where $\tau_Q(q, \dot{q}) = q$. This gives the curve as the sequence of configurations in $\mathbb{R}^n$ obtained by increasing the integrator time step from 0 to h. This implies that $\alpha^-(t) = 0$ and $\alpha^+(t) = t$ and therefore $\partial_h^-(q, \dot{q}) = q$ and $\partial_h^+(q, \dot{q}) = R_h(q, \dot{q})$.

The derivative of ∂_h^- is

$$D\partial_h^-(q, \dot{q})(\delta q, \delta \dot{q}) = \delta q.$$

Therefore, the variations $(\delta q_k, \delta\dot{q}_k)$ and $(\delta q_{k+1}, \delta\dot{q}_{k+1})$ must satisfy (see equations (39) and (40)):

$$D\partial_h^-(q_k, \dot{q}_k)(\delta q_k, \delta\dot{q}_k) = \delta q_k = 0, \tag{47}$$

$$D\partial_h^+(q_{k+1}, \dot{q}_{k+1})(\delta q_{k+1}, \delta\dot{q}_{k+1}) = DR_h(q_{k+1}, \dot{q}_{k+1})(\delta q_{k+1}, \delta\dot{q}_{k+1}) = 0. \tag{48}$$

Using equations (47) and (48) in equation (41) gives

$$DR_h(q_k, \dot{q}_k)(0, \delta\dot{q}_k) = \delta q_{k+1}. \tag{49}$$

Substituting equation (49) into equation (48) produces the following n linear equations in the $2n$ unknowns $(\delta\dot{q}_k, \delta\dot{q}_{k+1})$:

$$DR_h(q_{k+1}, \dot{q}_{k+1})(DR_h(q_k, \dot{q}_k)(0, \delta\dot{q}_k), \delta\dot{q}_{k+1}) = 0. \tag{50}$$

In components, equation (50) is

$$\frac{\partial R_h^i}{\partial q^j}(q_{k+1}, \dot{q}_{k+1})\frac{\partial R_h^j}{\partial \dot{q}^\ell}(q_k, \dot{q}_k)\delta\dot{q}_k^\ell + \frac{\partial R_h^i}{\partial \dot{q}^\ell}(q_{k+1}, \dot{q}_{k+1})\delta\dot{q}_{k+1}^\ell = 0, \tag{51}$$

for $i = 1 \ldots n$. Equations (50) or (51) will produce an n parameter set of solutions.

If there are no constraints, then the discrete Euler–Lagrange equations (42) are:

$$\frac{\partial L_d}{\partial \dot{q}^\ell}(q_k, \dot{q}_k)\delta\dot{q}_k^\ell + \frac{\partial L_d}{\partial q^i}(q_{k+1}, \dot{q}_{k+1})\frac{\partial R_h^i}{\partial \dot{q}^\ell}(q_k, \dot{q}_k)\delta\dot{q}_k^\ell + \frac{\partial L_d}{\partial \dot{q}^\ell}\delta\dot{q}_{k+1}^\ell = 0, \tag{52}$$

for each solution $(\delta\dot{q}_k, \delta\dot{q}_{k+1})$ of equations (50) or (51).

Without constraints, we need only find n linearly independent solutions of equations (50) or (51) to generate n independent discrete Euler–Lagrange equations (52).

Together with equations (43)

$$q_{k+1} = R_h(q_k, \dot{q}_k) \tag{53}$$

there are n equations in the n unknowns $(q_{k+1}, \dot{q}_{k+1})$.

Let $\mathcal{D}$ be a d dimensional constraint set, as in section 5.1. The discrete Lagrange–D'Alembert principle requires $DR_h(q_k, \dot{q}_k)(\delta q_k, \delta\dot{q}_k) \in \mathcal{D}$ (equation (46)). It will be more convenient to represent the constraints by their annihilators

$$\mathcal{D}^\circ = \{\, \phi \in \mathbb{R}^{n*} \mid \phi(X) = 0, X \in \mathcal{D}\,\}.$$

$\mathcal{D}^\circ$ are the vectors that are orthogonal to $\mathcal{D}$. If $\mathcal{D}$ has dimension d then $\mathcal{D}^\circ$ has dimension $n - d$. Equation (46) is then

$$\phi^a(R_h(q_k, \dot{q}_k))DR_h(q_k, \dot{q}_k)(\delta q_k, \delta\dot{q}_k) = 0, \tag{54}$$

for $a = 1 \ldots n - d$. Using equations (47) and (49), (54) and (53)

$$\phi^a(q_{k+1})\delta q_{k+1} = 0, \tag{55}$$

which is better written as

$$\phi^a_\ell(q_{k+1})\frac{\partial R^\ell_h}{\partial \dot{q}^j}(q_k, \dot{q}_k)\delta\dot{q}^j_k = 0, \tag{56}$$

for $a = 1 \ldots n - d$.

Equations (56) and (51) are $2n-d$ linear equations in the $2n$ variables $(\delta\dot{q}_k, \delta\dot{q}_{k+1})$, leaving only a d parameter set of solutions. The constraint equations (44) and (44) are

$$\phi^+(\bar{q}, \dot{\bar{q}}) = R_h(q_{k+1}, \dot{q}_{k+1}), \tag{57}$$

$$\phi^-(\bar{q}, \dot{\bar{q}}) = q_{k+1}, \tag{58}$$

where $\dot{\bar{q}} \in \mathcal{D}$. As in section 5.1, the number of equations and variables goes to $3n + d$.

The discrete Euler–Lagrange equations require a set of linearly independent solutions to equations (51). Without constraints, one can safely pick

$$\delta\dot{q}_k^i = 1, \quad \delta\dot{q}_k^j = 0, \text{ for} j \neq i, \tag{59}$$

for $i = 1 \dots n$. Then, because $\frac{\partial R_h}{\partial \dot{q}}(q, \dot{q})$ is invertible by requirement (section 5), $\delta\dot{q}_{k+1}$ can be solved for in equation (51).

With constraints, one has to be careful in picking a parametrisation, since the constraint equations (56) force relationships between the variables $(\delta\dot{q}_k, \delta\dot{q}_{k+1})$. Let $\{ X_i \}$ be a basis for $\mathcal{D}$. Then equations (50) and (55) can be completed to a set of $2n$ independent equations by adding

$$X_i \cdot \delta\dot{q}_k = 1, \quad X_j \cdot \delta\dot{q}_k = 0, \text{ for } j = 1 \dots d, j \neq i, \tag{60}$$

for $i = 1 \dots d$.

7. Routh's Sphere

7.1. **Equations of Motion.** The configuration space of a ball is the Lie group $\mathcal{E}(3) = \mathbb{R}^3 \times \mathcal{SO}(3)$, the space of rigid motions and rotations of $\mathbb{R}^3$. Fix a reference

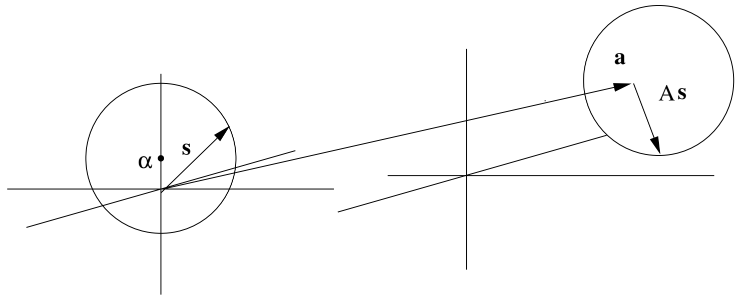

FIGURE 5. Configuration of a rolling sphere

sphere in $\mathbb{R}^3$ with its centre of mass at the origin. See figure (5). A position of the ball is given by applying an element (a, A) of $\mathcal{E}(3)$ to the reference sphere. This amounts to translating the reference sphere by a and rotating it by A. For every configuration of the ball, there is a unique point, s on the reference sphere such that $a + As$ is the point of contact of the ball with the surface.

The phase space of the rolling ball system is $T\mathcal{E}(3) = T\mathbb{R}^3 \times T\mathcal{SO}(3)$. Following Cushman [4] and Hermans [7], write the equations of motion on the left trivialisation $\mathcal{E}(3) \times \mathfrak{e}(3) = \mathbb{R}^3 \times \mathcal{SO}(3) \times \mathbb{R}^3 \times \mathfrak{so}(3)$. And further, identify the Lie algebra $\mathfrak{so}(3)$ with $\mathbb{R}^3$. In short, an element of phase space is (a, A, b, ω), where a is the position of the centre of mass, A is the orientation represented by rotating the reference sphere by A and $b = A^{-1}\dot{a}$. $\omega = A^{-1}\dot{A}$ is the angular velocity of the rolling sphere with respect to the fixed axis of the reference sphere.

The Lagrangian on $\mathcal{E}(3) \times \mathfrak{e}(3)$ is

$$L(a, A, b, \omega) = \frac{1}{2} I\omega \cdot \omega + \frac{1}{2} mb \cdot b - mga \cdot e_3,$$

where $I = \mathrm{diag}(I_1, I_1, I_3)$ is the moment of inertia tensor of the sphere, m is the mass of the sphere and g is the acceleration due to gravity. The e_3 axis is lined up with the axis of symmetry in the reference sphere.

The rolling constraint is enforced by requiring the instantaneous axis of rotation to be through the contact point and parallel to ω. With respect to the reference sphere, this is

$$b = s \times \omega.$$

The force of constraint on $\mathbb{R}^3 \times \mathfrak{so}(3)$ is $F = (\lambda, s \times \lambda)$. Adding this constraint force to the unconstrained Euler equations (see Arnold [1]) and using the Lagrange multiplier $\lambda = (\lambda_1, \lambda_2, \lambda_3)$, the equations of motion are

$$\begin{aligned}
\dot{a} &= Ab, \\
\dot{A} &= A\omega, \\
m\dot{b} &= mb \times \omega - mgA^{-1}e_3 + \lambda, && (61) \\
I\dot{\omega} &= I\omega \times \omega + s \times \lambda. && (62)
\end{aligned}$$

Eliminating λ and using the constraint, equations (61) and (62) are replaced by

$$\dot{b} = \frac{d}{dt}(s \times \omega), \tag{63}$$

$$I\dot{\omega} + m(s \times \dot{\omega}) \times s + m(\dot{s} \times \omega) \times s = I\omega \times \omega + m(s \cdot \omega)\omega - mgA^{-1}e_3 \times s, \tag{64}$$

which are implicitly defined ordinary differential equations.

Representing an element of $\mathcal{SO}(3)$ by a 3×3 orthogonal matrix requires six constraints to enforce the orthonormality of the columns. It is better to represent A by a unit quaternion since it is an element of $\mathbb{R}^4$ with only the one constraint enforcing unit length. See Bates and Cushman [3] for a more detailed discussion on this representation.

If A is a rotation about an axis (q_1, q_2, q_3) through the angle θ, the equivalent quaternion is

$$q = \left(\cos(\frac{\theta}{2}), \frac{\sin(\frac{\theta}{2})}{\|(q_1, q_2, q_3)\|}(q_1, q_2, q_3)\right)$$

Let $x = (x_1, x_2, x_3)$ be in $\mathbb{R}^3$ and write x in $\mathbb{R}^4$ as $x = (0, x_1, x_2, x_3)$. Write $q = (q_0, q_1, q_2, q_3)$ Using quaternion multiplication, define R_q by

$$R_q x = q x q^{-1}.$$

The result of this operation is to rotate x by the angle $\theta = 2\cos^{-1}(q_0)$ about the axis (q_1, q_2, q_3).

If $\omega = (\omega_1, \omega_2, \omega_3)$ is the angular velocity, then the corresponding quaternion in $\mathbb{R}^4$ is $\gamma = 2(0, \omega_1, \omega_2, \omega_3)$. An arbitrary vector tangent to the unit quaternions at the point q is $\dot{q} = (\dot{q}_0, \dot{q}_1, \dot{q}_2, \dot{q}_3) = q\gamma$ for some $\gamma = (0, \gamma_1, \gamma_2, \gamma_3)$ with base point $(1, 0, 0, 0)$.

Using quaternions, the configuration space for the sphere is $\mathbb{R}^3 \times \mathcal{S}^3$ where $\mathcal{S}^3$ is the three–sphere in $\mathbb{R}^4$. With quaternion multiplication $\mathcal{S}^3$ is a Lie group with identity element $(1, 0, 0, 0)$. Phase space is $T\mathbb{R}^3 \times T\mathcal{S}^3 = \{\, (a, q, \dot{a}, \dot{q}) \,\}$, where the coordinates have been arranged so that the configuration coordinates are the first $2n$ components and the velocity coordinates are the last $2n$ components. The Lagrangian is

$$L(a, q, \dot{a}, \dot{q}) = 2I(q^{-1}\dot{q}, q^{-1}\dot{q}) + \frac{m}{2}\dot{a} \cdot \dot{q} - mga \cdot e_3, \tag{65}$$

with the constraint

$$q_0\dot{q}_0 + q_1\dot{q}_1 + q_2\dot{q}_2 + q_3\dot{q}_3 = 0. \tag{66}$$

The rolling constraint is

$$\dot{a} = -\frac{d}{dt}(R_q)s. \tag{67}$$

To find a formula for s, let a configuration (a, q) be in $\mathbb{R}^3 \times \mathbb{R}^4$. Let $n(s)$ be the outward unit normal to the reference sphere at the point s. Restricting to flat surfaces only, s must satisfy $An(s) = -e_3$. For a sphere of unit radius and centre of mass a distance α from the geometric centre (see figure 5), s as a function of q is

$$s = (s_1, s_2, s_3) = \frac{1}{\|q\|^2}\left(2(q_0 q_2 - q_1 q_3)\,,\, -2(q_0 q_1 + q_2 q_3)\,,\, q_1^2 + q_2^2 - q_0^2 - q_3^2 + \alpha \right).$$

The constraint equation (67) is computed by expanding the time derivative.

$$\dot{a} = -\dot{q}sq^{-1} + qs(\dot{q^{-1}}) = \dot{q}sq^{-1} - qsq^{-1}\dot{q}q^{-1}.$$

Rearranging, the constraint can be written as

$$q^{-1}\dot{q}s - sq^{-1}\dot{q} + q^{-1}\dot{a}q = 0.$$

The three components of the constraint are

$$\begin{aligned} &\frac{2}{\|q\|^2}\dot{q}_0(s_2 q_3 - s_3 q_2) - \frac{2}{\|q\|^2}\dot{q}_1(s_2 q_2 + s_3 q_3) + \frac{2}{\|q\|^2}\dot{q}_2(s_3 q_0 + s_2 q_1) + \frac{2}{\|q\|^2}\dot{q}_3(s_3 q_1 - s_2 q_0) + \\ &\qquad \frac{1}{\|q\|^4}\dot{a}_1(q_1^2 + q_0^2 - q_3^2 - q_2^2) + \frac{2}{\|q\|^4}\dot{a}_2(q_1 q_2 + q_0 q_3) + \frac{2}{\|q\|^4}\dot{a}_3(q_1 q_3 - q_0 q_2) = 0, \end{aligned} \tag{68}$$

$$\begin{aligned} &\frac{2}{\|q\|^2}\dot{q}_0(s_3 q_1 - s_1 q_3) + \frac{2}{\|q\|^2}\dot{q}_1(s_1 q_2 - s_3 q_0) - \frac{2}{\|q\|^2}\dot{q}_2(s_1 q_1 + s_3 q_3) + \frac{2}{\|q\|^2}\dot{q}_3(s_1 q_0 + s_3 q_2) + \\ &\qquad \frac{2}{\|q\|^4}\dot{a}_1(q_1 q_2 - q_0 q_3) + \frac{1}{\|q\|^4}\dot{a}_2(q_2^2 + q_0^2 - q_1^2 - q_3^2) + \frac{2}{\|q\|^4}\dot{a}_3(q_2 q_3 + q_0 q_1) = 0, \end{aligned} \tag{69}$$

$$
\begin{aligned}
(70)\quad & \frac{2}{\|q\|^2}\dot q_0(s_1q_2 - s_2q_1) + \frac{2}{\|q\|^2}\dot q_1(s_2q_0 + s_1q_3) + \frac{2}{\|q\|^2}\dot q_2(s_2q_3 - s_1q_0) - \frac{2}{\|q\|^2}\dot q_3(s_2q_2 + s_1q_1) + \\
& \frac{2}{\|q\|^4}\dot a_1(q_1q_3 + q_0q_2) + \frac{2}{\|q\|^4}\dot a_2(q_2q_3 - q_0q_1) + \frac{1}{\|q\|^4}\dot a_3(q_3^2 + q_0^2 - q_2^2 - q_1^2) = 0.
\end{aligned}
$$

7.2. **Numerical Simulation.** For a first numerical simulation, take the curves $\psi_h(a, q, \dot a, \dot q) = (a + t\dot a, q + t\dot q$, $\alpha^-(t) = 0$, $\alpha^+(t) = t$, so that the integrator will be derived from the explicit Euler method, as in section 6. Take the discrete Lagrangian to be the left endpoint approximation of

$$\frac{1}{h}\int_0^h L(a + t\dot a, q + t\dot q, \dot a, \dot q)\, dt,$$

so that $L_d(a, q, \dot a, \dot q) = L(a, q, \dot a, \dot q)$, where L is given in equation (65). Recall from section 4 that this generates the symplectic Euler method.

The constraint, $\mathcal{D}$, is given by the vectors $(\dot a, \dot q)$ that satisfy the constraint equations (66), (68), (69) and (70).

Write the exponential map from $\mathfrak{so}(4)$ to $\mathcal{SO}(4)$ by $\exp_{\mathcal{SO}(4)}$. This will distinguish it from the exponential map exp from $\mathfrak{so}(3)$ to $\mathcal{SO}(3)$ which is also needed.

Let $(0, \gamma_1, \gamma_2, \gamma_3)$ be a vector tangent to $\mathcal{S}^3$ at $(1, 0, 0, 0)$. Then let $\gamma \in \mathfrak{so}(3)$ be

$$
\gamma = \begin{bmatrix} 0 & -\gamma_1 & -\gamma_2 & -\gamma_3 \\ \gamma_1 & 0 & \gamma_3 & -\gamma_2 \\ \gamma_2 & -\gamma_3 & 0 & \gamma_1 \\ \gamma_3 & \gamma_2 & -\gamma_1 & 0 \end{bmatrix}. \tag{71}
$$

Also, let $\omega = 2(\gamma_1, \gamma_2, \gamma_3) \in \mathfrak{so}(3)$. Then, $\exp_{\mathcal{SO}(4)}(t\gamma)$ generates a rotation of the configuration which corresponds to a rotation $\exp(t\omega)$ of the sphere. The discretized constraints are:

$$\phi_h(a,q) = (a + R_q s - \mathbf{R} + \exp(hA\omega)(\mathbf{R} - R_q s) + h\mathbf{R} \times A\omega, \exp_{\mathcal{SO}(4)}(h\gamma)q), \tag{72}$$

where $\mathbf{R} = (0,0,-1)$. Let $r = \|\gamma\|$. The formulae for the exponential maps are,

$$\exp_{\mathcal{SO}(4)}(t\gamma) = \begin{bmatrix} \cos(tr) & -\frac{\gamma_1}{r}\sin(tr) & -\frac{\gamma_2}{r}\sin(tr) & -\frac{\gamma_3}{r}\sin(tr) \\ \frac{\gamma_1}{r}\sin(tr) & \cos(tr) & \frac{\gamma_3}{r}\sin(tr) & -\frac{\gamma_2}{r}\sin(tr) \\ \frac{\gamma_2}{r}\sin(tr) & -\frac{\gamma_3}{r}\sin(tr) & \cos(tr) & \frac{\gamma_1}{r}\sin(tr) \\ \frac{\gamma_3}{r}\sin(tr) & \frac{\gamma_2}{r}\sin(tr) & -\frac{\gamma_1}{r}\sin(tr) & \cos(tr) \end{bmatrix}, \tag{73}$$

$$\exp(tA\omega) = I + \frac{\sin(2tr)}{2tr}tA\omega + \frac{1-\cos(2tr)}{4t^2r^2}(tA\omega)^2. \tag{74}$$

The idea behind using equations 72, 73 and 74 is that γ generates a rotation of $q \in \mathcal{S}^3$. There is a corresponding rotation generated by $A\omega$ of the centre of mass about the geometric centre of the sphere. There must also be a rigid translation in the direction $\mathbf{R} \times A\omega$ by an amount equal to the the arc length traced out by the contact point on the sphere.

Note that $A\omega$ corresponds to the angular velocity in the space frame, whereas ω is the angular velocity with respect to a coordinate system fixed in the reference body.

7.3. **Numerical Results.** Here we present the results of some simulations.

In the first case, $\alpha = 0$ and $I_3 = I_2 = I_1 = 1$ corresponding to a uniform sphere. The initial conditions are an angular velocity in the e_2 direction to correspond to rolling along the e_1 axis.

Figure 6 shows a linear drift in the energy of very small slope. This is not typical behaviour for a variational integrator and may be a result of our constraint discretization. Further experiments need to be conducted on this.

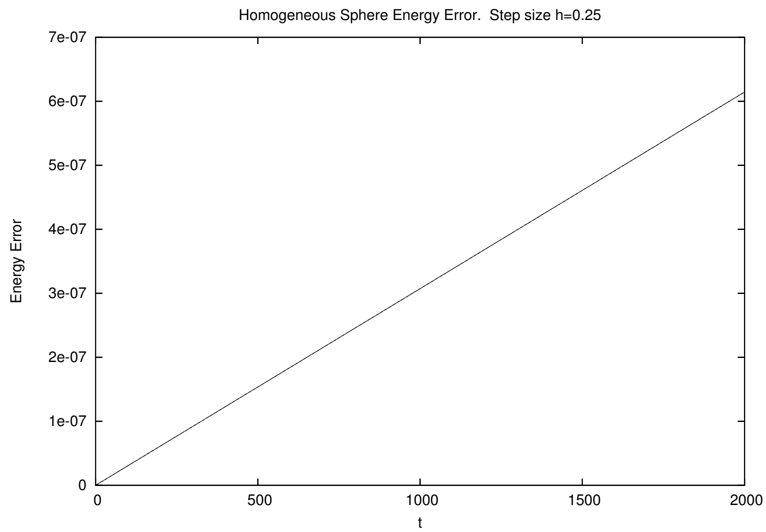

FIGURE 6

Figure 7 shows that the angular velocity in the e_2 direction also has a small linear drift. Note that the initial conditions are given as $(\gamma_1, \gamma_2, \gamma_3) = (0, .5, 0)$. This produces an angular velocity of $\omega = (0, 1, 0)$ in the continuous system, and slightly less than that in the simulated system due to the discretized constraints.

The plot in figure 8 is of q_0 and shows that it remains nicely bounded, as should be expected.

Now let $\alpha = 0.2$ and $I_3 = 0.8$. The initial conditions for these plots are $(a_1, a_2, a_3) = (-0.199, 0, 0)$ and $\omega = (0, 0, 0)$. This produces a rocking motion along the e_1 axis and a corresponding oscillation in ω_2.

The plot in figure 9 shows the bounded energy behaviour we expect. Note that only every tenth point is plotted.

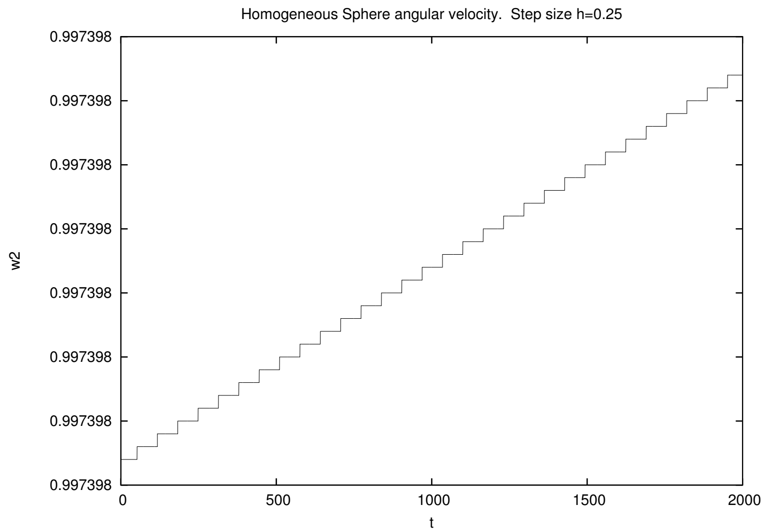

FIGURE 7

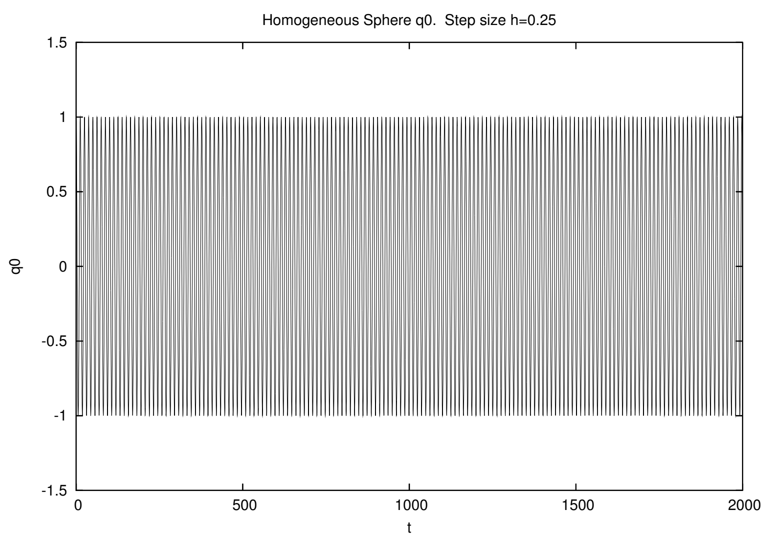

FIGURE 8

Figure 10 shows the a_3 component of the centre of mass. Again, only every tenth data point is shown.

Finally, figure 11 shows the oscillations of the ω_2 component of angular velocity.

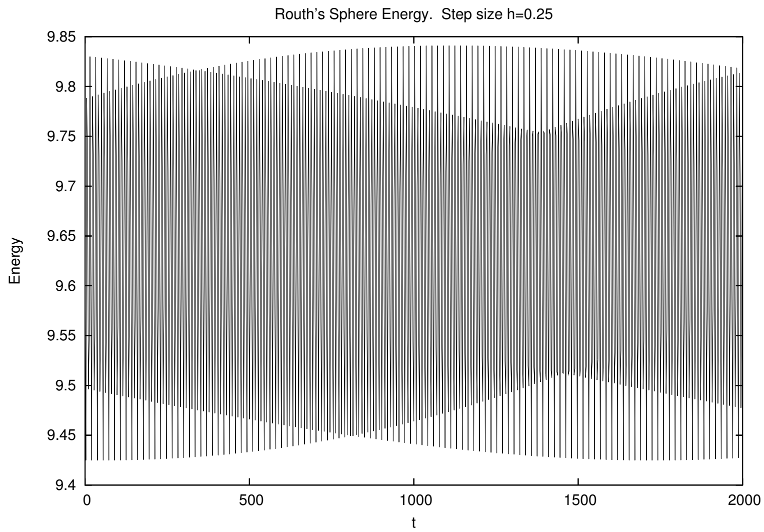

FIGURE 9

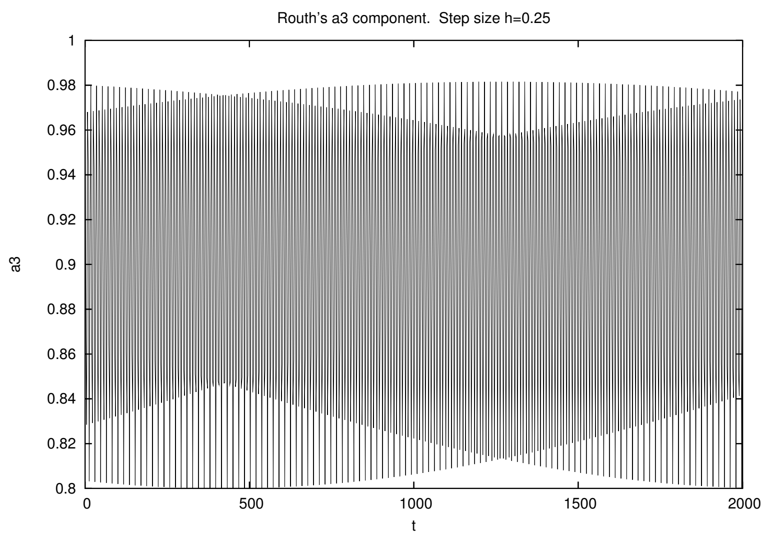

FIGURE 10

8. Conclusions and Further Work

In this project, we introduced a new technique that has been developed for generating variational integrators for constrained systems. We showed, through examples, that there is good potential for this method to produce accurate long time integrators.

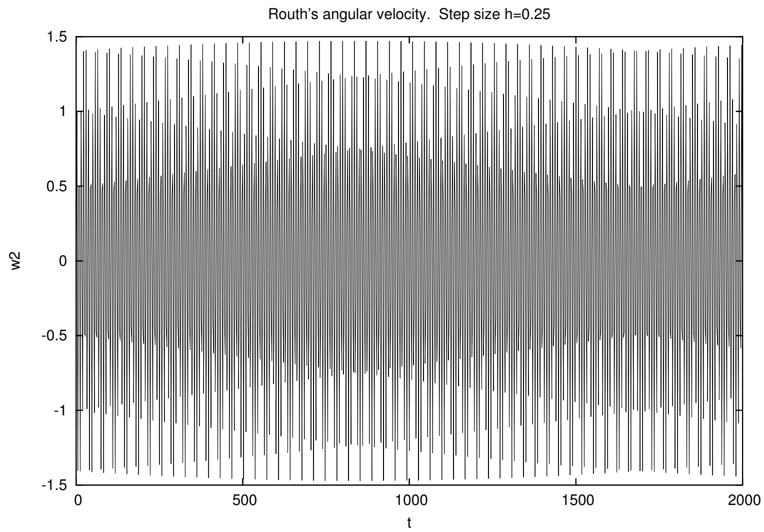

FIGURE 11

There is still considerable work to do on this project.

Discretization. Producing the discretized constraints has proven to be the greatest difficulty, since it involves finding the mapping ϕ_h, which is typically the flow of a vector field. Techniques need to be developed to incorporate a very accurate differential equation solver into the scheme to compute ϕ_h at every time step.

Rolling Rigid Bodies. Rolling rigid bodies are proving to be a good test case, as they incorporate technical difficulties (i.e. nontrivial configuration space and implicitly defined equations of motion) while still providing examples for which analytic solutions and qualitative behaviour are known.

Routh's Sphere. It is known that there are motions of Routh's sphere that result in the centre of mass moving asymptotically to a position directly above the geometric centre. The integrator should be tested on these solutions.

High Order Integrators. The ability to generate high order variational integrators from Lagrangian integrators may be one of the most beneficial aspects of this technique. Examples need to be written and tested against other high order methods.

Numerical Analysis. A study of the order and stability properties of various methods developed by this technique needs to be done.

Geometry. A holonomically constrained integrator can be written as a lower dimensional system by adapting the curves ψ_h to the constraint. This assertion still needs to be proven and its utility tested.

Software. Good, general purpose software needs to be written. The current code is adequate for small simple systems. Compound rigid bodies, such as the mobile robot with three wheels, may have many repeated components (i.e. three ball wheels, nine rollers) which suggests that a more strongly object oriented approach should be considered.

References

[1] V. Arnold. *Mathematical Methods of Classical Mechanics.* Springer-Verlag, 2 edition, 1989.

[2] Uri Ascher and Linda Petzold. *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations.* The Society for Industial and Applied Mathematics, 1998.

[3] L. Bates and R. Cushman. *Global Aspects of Classical Integrable Systems.* Birkhäuser Verlag, 1997.

[4] Richard Cushman. Routh's sphere. *Reports on Mathematical Physics*, 42:47–70, 1998.

[5] I. Gelfand and S. Fomin. *Calculus of Variations.* Prentice-Hall, 1963.

[6] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration.* Springer–Verlag, 2002.

[7] Hermans Joost. *Rolling Rigid Bodies with and without Symmetries.* PhD thesis, University of Utrecht, 1995.

[8] J. Marion and S. Thornton. *Classical Dynamics of Particles and Systems.* Saunders College Publishing, 1995.

[9] Jerrold Marsden and Tudor Ratiu. *Introduction to Mechanics and Symmetry*. Springer-Verlag, 2 edition, 1999.

[10] G. Patrick and C. Cuell. In preperation. 2005.

[11] R. Rosenberg. *Analytical Dynamics of Discrete Systems*. Plenum Press, 1977.