Alex Malins 14/03/2007

TIME STEP ERROR IN MOLECULAR DYNAMICS

1. Abstract

In this paper solutions to dynamical systems obeying Hamiltonian mechanics are obtained using algorithms from numerical integration. Symplectic integrators are preferred and applied to test problems such as a simple pendulum and a spring pendulum. Ascertaining the most computationally efficient route to solution has traditionally been a trade off between a employing higher order integrators and lowering the time step at a set order. It is proposed that a new set of time reversible integrators, which solve a 'modified' Hamiltonian instead of the exact, may perform better than standard symplectic integrators of an equivalent order. Our investigation of the performance of the new schemes against established integrators on the test problems finds no evidence to support this claim. The study moves on to analyse the effect of discrete jumps in energy of the numerical solution, which exist for certain trajectories of the spring pendulum, on the performance of the integrators. Finally we study numerical solutions to a Lennard-Jones fluid, where energy jumps are prevalent and the system provides a much sterner test for the integrators than the two pendulum systems.

2. Introduction

This paper studies numerical integration schemes used to solve the ordinary differential equations (ODEs) arising from a number of physical systems. It is common for a formulation of a dynamical system to result in a set of ODEs which cannot be solved using analytical methods. Under these circumstances the best route to a solution trajectory is to use a numerical integrator to solve the equations for a future time, given an initial state of the system. Integrators can approximate the exact solution to any specified degree of accuracy, with the only limit being computational time required to execute the algorithm - hence the integrator efficiency is critical!

This paper builds on the research in *Practical Construction of Modified Hamiltonians* by Skeel and Hardy (2001), where the focus is on the economics of numerical schemes in solving dynamical systems arising from Hamiltonian mechanics. The study is limited to conservative systems, where the total energy of the system remains constant with respect to time. Symplectic integrators are designed to solve conservative systems, with the solution Hamiltonian being bounded in time, unlike time irreversible schemes where the solution energy is allowed to drift.

We implement a series of established integrators on some physical test problems and develop measures for the performance. This requires careful analysis of the relative errors of the numerical solutions, with the theoretical rate of convergence to the exact solution trajectory being judged against the observed rate of convergence.

The performance analysis first requires an understanding of the motion exhibited by the system. Graphs of the motion of the system are examined, and in certain cases Poincaré sections are plotted, to establish underlying nature of the motion. The data for these graphs is obtained via brute force – using a small time step and an algorithm with a high theoretical degree of accuracy.

The expertise gained is then transferred to an analysis of a new type of integrator, proposed in *An Augmented Verlet Algorithm* by R.C. Ball (preprint 2007). These integrators solve a 'modified' Hamiltonian, instead of the exact Hamiltonian of the physical model. A simple procedure is presented here to generate the 'modified' Hamiltonian and then the performance of the scheme is directly contrasted against the traditional schemes. Of pertinent interest is establishing the most computationally efficient scheme, with least burden on resources, over the test problems. The paper then moves on to discuss the limitations of each scheme and suggests future directions research in this area could take.

3. Background Theory

The study focuses solely on systems obeying Hamiltonian mechanics. The Hamiltonian of a system can be regarded as the total energy of the system and is a constant with respect to time for conservative systems. The Hamiltonian $H(q_j, p_j)$ is a function of the generalised coordinates of the system $\{q_j \mid j = 1,..., N\}$ and their respective canonical momenta $\{p_j \mid j = 1,..., N\}$. In non-magnetic systems this may be written as the sum of the total kinetic energy T and total potential energy V of the system in question,

$$H = T + V \tag{1}.$$

Hamilton's equations can then be used to find the differential equations of motion which determine the motion of the system,

$$\frac{d}{dt} q_i(t) = \frac{\partial H(q_j, p_j)}{\partial p_i}; \frac{d}{dt} p_i(t) = -\frac{\partial H(q_j, p_j)}{\partial q_i} \tag{2}.$$

The study examines three Hamiltonian systems: the simple and spring pendulums, and a Lennard-Jones Fluid. The mathematical models for these systems, including any implied assumptions, are presented below.

3.1. The Simple Pendulum

Polar (r, θ) coordinates form the natural coordinate system for the simple pendulum, which moves on a 2-dimensional plane in space.

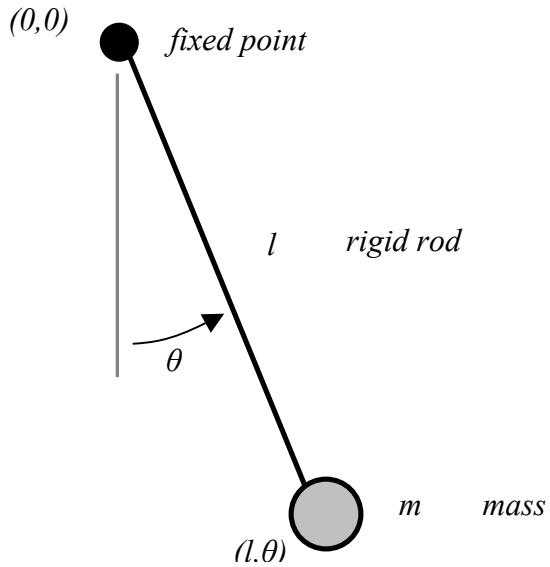

Fig.1. Diagram of the simple pendulum system.

Assume the only forces acting on the mass are due to its weight and tension in the rigid rod, with the weight of the rod being insignificant. The rod is free to pivot at the fixed point. Using (1) the Hamiltonian is written:

$$H(\theta, p_\theta) = \frac{p_\theta^2}{2ml^2} - mgl\cos(\theta) \tag{3}.$$

Note how the Hamiltonian is independent of time. An application of Hamilton's equations (2) gives,

$$\frac{d}{dt}\theta(t) = \frac{p_\theta}{ml^2}; \frac{d}{dt}p_i(t) = -mgl\sin(\theta) \tag{4}.$$

These are the equations of motion for the system (4), with the non-linearity of the second meaning obtaining a solution for all times via analytical techniques is not possible.

3.2. The Spring Pendulum

Cartesian coordinates (x, y) are used for the spring pendulum (Fig. 2), with motion restricted to a two dimensional plane. Fig. 1 depicts the setup involved.

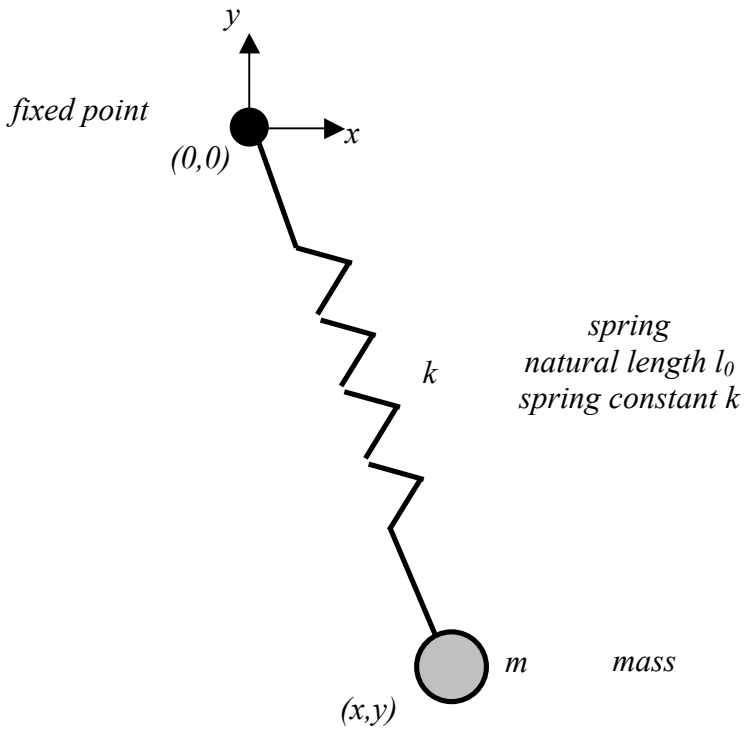

Fig. 2. Diagram of the spring pendulum system.

The spring has a very small mass compared to the mass which is supports, and is free to pivot at the fixed point. Assume the only forces which act upon the mass is its weight and a restoring/tension force due to the spring. The Hamiltonian and Hamilton's equations are:

$$H = \frac{p_x^2 + p_y^2}{2m} + mgy + \frac{1}{2}k(\sqrt{x^2 + y^2} - l_0)^2 \quad (5),$$

$$\frac{d}{dt}x(t) = \frac{p_x}{m}; \frac{d}{dt}p_x(t) = -kx(1 - \frac{l_0}{\sqrt{x^2 + y^2}})$$

$$\frac{d}{dt}y(t) = \frac{p_y}{m}; \frac{d}{dt}p_y(t) = -mg - ky(1 - \frac{l_0}{\sqrt{x^2 + y^2}}) \quad (6).$$

There is only one dimensionless parameter in this model of the spring pendulum, *K,*

$$K = \frac{l_0 k}{mg} \quad (7).$$

Attention is paid to the special case where the natural frequency of the transverse oscillations, where the system is equivalent to a simple pendulum, is twice that of the natural frequency of the vertical oscillations of an equivalent mass spring system. For this to occur,

$$K = 3 \quad (8).$$

3.3. The Lennard-Jones Fluid

The Lennard-Jones potential between two particles separated by a distance *r* is given by,

$$V(r) = 4\varepsilon\left[\left(\frac{\sigma}{r^{12}}\right) - \left(\frac{\sigma}{r^{6}}\right)\right] \quad (9).$$

where σ is the radius of zero potential and ε is the depth of the potential well. Refer to Figure 3.

An ensemble of *N* point particles in a three dimensional space interact subject to the Lennard-Jones potential. At long ranges each particle pair experiences an attractive force and at a short range the force is repulsive. In this study, the particles are constrained by a harmonic potential well centred on the origin. There are no gravitational forces acting on the particles.

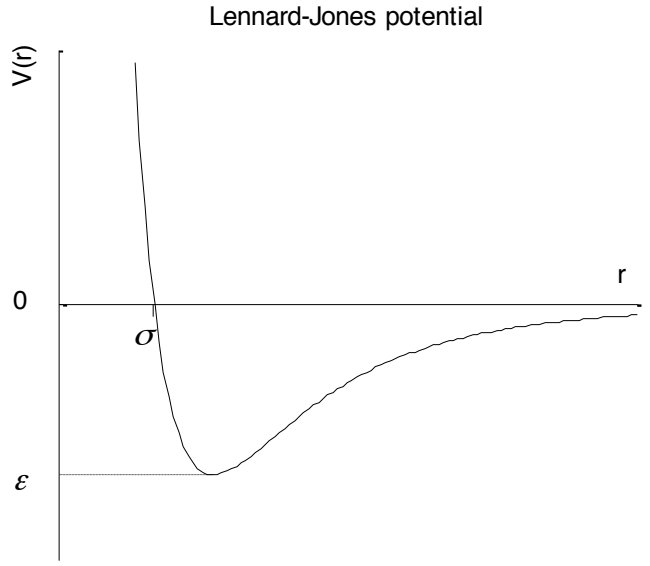

Fig 3. Graph of Lennard-Jones potential versus separation distance of particles.

For an ensemble of *N* particles, with particle *i* at position (x_i, y_i, z_i) and mass m_i, the Hamiltonian of the system is,

$$H = \tfrac{1}{2}\sum_{i=1}^{N} m_i({p_{x_i}}^2 + {p_{y_i}}^2 + {p_{z_i}}^2) + \tfrac{1}{2}k\sum_{i=1}^{N}({x_i}^2 + {y_i}^2 + {z_i}^2) + \tfrac{1}{2}4\varepsilon\sum_{i=1}^{N}\sum_{j=1}^{N}[(\tfrac{\sigma}{r_{ij}^{12}}) - (\tfrac{\sigma}{r_{ij}^{6}})] \quad (10),$$

where $r_{ij} = [(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2]^{0.5}$ is the distance between particles *i* and *j*. Application of Hamilton's equations for particle *i* give,

$$\begin{aligned}
\frac{d}{dt}x_i(t) &= \frac{p_{x_i}}{m_i}; \frac{d}{dt}p_{x_i}(t) = -kx_i + 24\varepsilon\sigma^6\sum_{j=1}^{N} 6(x_i - x_j)r_{ij}^{-8}(2\sigma^6 r_{ij}^{-6} - 1) \\
\frac{d}{dt}y_i(t) &= \frac{p_{y_i}}{m_i}; \frac{d}{dt}p_{y_i}(t) = -ky_i + 24\varepsilon\sigma^6\sum_{j=1}^{N} 6(y_i - y_j)r_{ij}^{-8}(2\sigma^6 r_{ij}^{-6} - 1) \\
\frac{d}{dt}z_i(t) &= \frac{p_{z_i}}{m_i}; \frac{d}{dt}p_{z_i}(t) = -kz_i + 24\varepsilon\sigma^6\sum_{j=1}^{N} 6(z_i - z_j)r_{ij}^{-8}(2\sigma^6 r_{ij}^{-6} - 1)
\end{aligned} \quad (11).$$

4. The Numerical Schemes

Numerical schemes for solving the differential equations, resulting from Hamiltonian mechanics, involve marching the system forwards through time in discrete stages. The schemes are called integrators and the process of converting continuous differential equations to discrete difference equations is known as discretization. For basic integrators the discrete stages move the system forwards a fixed amount of time called the time step, calculating new estimates for the positions and momenta of the system from previously known values. All the schemes studied involve underlying updates of the system coordinates using difference equations (12), concatenations of which are employed in more complex integrators to form an overall stage.

$$q_i(t+\alpha) = q_i(t) + \alpha\frac{p_i(\tau)}{m} \quad (12a),$$

$$p_i(t+\beta) = p_i(t) + \frac{\beta}{m}F(q_i(\tau)) \quad (12b).$$

The values α, β, τ are constants unique to each integrator. $F(q_i(\tau))$ is a function, independent of the momenta, equivalent to a force.

For simplicity an overall time step Δt is defined for each step of an integrator (13). In general the number of the $F(q_i(\tau))$ 'force' calculations involved per time step is the dominant computational exertion when running the numerical simulations.

$$\begin{aligned}
q_i(t) &\to q_i(t+\Delta t) \\
p_i(t) &\to p_i(t+\Delta t)
\end{aligned} \quad (13).$$

The subject of this paper to analyse how accurately the exact dynamics of the physical test systems are predicted by the integrators, as a trade off with the computational requirements needed for them to run. Analytically the rate of convergence of the estimates to the exact trajectories can be calculated by comparing the integration method to a Taylor expansion of the true dynamics. This shows that the rate of convergence is proportional

to $(\Delta t)^{\gamma}$, where γ is a positive integer known as the order of accuracy of the integrator. However this does not depict to full story of the errors introduced in the discretization of the continuous dynamics.

It is known that integrators which are symplectic perform more strongly than non-symplectic integrators in solving Hamiltonian. This is because symplectic integrators conserve a Hamiltonian slightly perturbed from the exact Hamiltonian through time, whereas non-symplectic integrators allow the energy of the system to drift.

The paper moves on to study the performance of a recent set of symplectic integrators, employing the philosophy of solving a 'modified' Hamiltonian. The integrators are now defined.

4.1. Euler's Method

Euler's method is the simplest integrator where both the positions and momenta are updated simultaneous for each time step. The algorithm is first order accurate and time irreversible, meaning there is no constraint on the energy of the predictions. There is one force calculation per time step.

$$\begin{aligned} q_i(t+\Delta t) &= q_i(t) + \tfrac{\Delta t}{m} p_i(t) \\ p_i(t+\Delta t) &= p_i(t) + \Delta t \tfrac{dp_i}{dt}(t) \end{aligned} \qquad (14).$$

4.2. Symplectic Algorithms

Symplectic algorithms are time reversible and theoretically their solution energies are bounded. In the algorithm the positions and momenta are updated alternately, using the most recent estimate of the corresponding coordinate in the calculation. It is natural to represent these methods using the following notation:

A	$\frac{1}{2}$
B	1
A	$\frac{1}{2}$

Table 1. Normalised 2nd order Verlet coefficients. Note how sum of both the *A* and *B* coefficients is 1.

where the time reversibility of the algorithm is clear.

In Table 1, *A* represents an update of the either positions or the momenta (there is no specification as to which) using respective formula in (12); the second column contains the

fractional value $\frac{\alpha}{\Delta t}$ or $\frac{\beta}{\Delta t}$ (the proportion of a time step update the coordinates experience) and τ is the time of the most recent estimate of the corresponding coordinate. Conversely for update *B*.

Table 1 depicts the standard, second order accurate, Verlet algorithm. If the *B* step is an update of the momenta, there is one force calculation per time step. It may appear that if the *A* steps were updates of momenta, there would be two force calculations per time step. However it is entirely equivalent to programme the scheme as,

A	1
B	1

where the two *A* updates in Table 1 have been merged. Thus the algorithm reduces to one force calculation per time step. Care must be taken as the positions and momenta are half a time step Δt out of phase, which has implications for the calculation of the Hamiltonian and for first time step away from the starting conditions. These problems are easily surmounted by interpolating one of the two coordinates to the times that at which the other is known.
The principle of unifying the first and last steps can be applied to all the symplectic algorithms.

It is worth noting that although there is no rule as to which of the coordinate updates *A* and *B* represent, the simulations are distinct and usually produce different numerical solutions – this implies for a given simulation one solution will be more accurate than the other! This is looked at in the results section of this paper.

Three types of symplectic integrators with fourth order accuracy are studied. The first is a three force calculation method (Forest and Ruth, 1989) with coefficients in Table 2a.

A	$\frac{1}{6}(2^{1/3} + 2^{-1/3} + 2)$
B	$\frac{1}{3}(2^{1/3} + 2^{-1/3} + 2)$
A	$-\frac{1}{6}(2^{1/3} + 2^{-1/3} - 1)$
B	$-\frac{2}{3}(2^{1/3} + 2^{-1/3} + \frac{1}{2})$
A	$-\frac{1}{6}(2^{1/3} + 2^{-1/3} - 1)$
B	$\frac{1}{3}(2^{1/3} + 2^{-1/3} + 2)$
A	$\frac{1}{6}(2^{1/3} + 2^{-1/3} + 2)$

Table 2a. A three force calculation, fourth order, symplectic integrator (Forest & Ruth, 1990).

Tables 2b and 2c contain the coefficients for two more fourth order accurate symplectic integrators; Table 2b contains a Mclachlan, 1995, scheme involving five force calculations per time step and 2c a six force calculation scheme (Blanes and Moan, 2000).

A	0.14
B	0.28
A	0.452733214233835
B	0.625466428467670
A	-0.092733214233835
B	-0.810932856935340
A	-0.092733214233835
B	0.625466428467670
A	0.452733214233835
B	0.28
A	0.14

Table 2b. A five force calculation, fourth order, symplectic integrator (Mclachlan, 1995).

A	0.079203696431196
B	0.209515106613362
A	0.353172906049774
B	-0.143851773179818
A	-0.042065080357720
B	0.434336666566456
A	0.219376955753500
B	0.434336666566456
A	-0.042065080357720
B	-0.143851773179818
A	0.353172906049774
B	0.209515106613362
A	0.079203696431196

Table 2c. A six force calculation, fourth order, symplectic integrator (Blanes & Moan, 2002).

4.3 The 'Modified' Hamiltonian Method

The philosophy in the standard symplectic integrators above is to cancel a specified number terms the Taylor expansion of the true trajectory, but this leaves a series of higher order 'truncation' errors unaccounted for. The concept behind the 'modified' Hamiltonian integrators involves cancelling terms in the Taylor series, once again to a specified degree, then counter terming higher order truncation errors by adding a correction to the Hamiltonian of the system. The integrators solve the dynamics of this '*modified*' Hamiltonian, not the exact Hamiltonian, resulting in numerical solutions closer to that of the true dynamics.

Two integrators are tested with a basis equivalent to that of a second order, two force calculation, symplectic integrator. However. The two schemes are defined in Table 3. Note how updates *A* and *B* are replaced by *Q* and *P*, which represent updates in the positions and momenta respectively, i.e. there is now a constraint imposed by the algorithm as to which of the position or momenta updates occur at stages *A* and *B*. This is a consequence of adding a term to the Hamiltonian of the system.

Q	$\frac{1}{2} - \frac{1}{2\sqrt{3}}$
P	$\frac{1}{2}$
Q	$\frac{1}{\sqrt{3}}$
P	$\frac{1}{2}$
Q	$\frac{1}{2} - \frac{1}{2\sqrt{3}}$

Table 3a. A fourth order, position update first, modified Hamiltonian symplectic integrator.

$\lambda = \frac{3+\sqrt{3}}{24(9+5\sqrt{3})}$

P	$\frac{1}{6}$
Q	$\frac{1}{2}$
P	$\frac{2}{3}$
Q	$\frac{1}{2}$
P	$\frac{1}{6}$

Table 3b. A fourth order, momentum update first, modified Hamiltonian symplectic integrator.

$\lambda = \frac{1}{72}$

Tables 3a and 3b. replace updates A and B by Q and P, as the coordinate updated during each stage is now explicit. Q represents an update in the position coordinates (12a), and P an update in momentum coordinates. Numerical schemes by R.C. Ball (preprint 2007).

The modified Hamiltonian $\widetilde{H}$ solved is given by:

$$\widetilde{H} = H + \lambda(\Delta t)^2 \mathbf{f}.\mathrm{M}^{-1}.\mathbf{f} \tag{15}$$

H is the exact Hamiltonian of the system, M the inertia tensor of the system, λ a constant of the algorithm, and $\mathbf{f} = -\nabla V$ the force. $\widetilde{H}$ is then used as the basis for the integration schemes in Table 3, in a completely equivalent manner to the Verlet algorithms and the exact Hamiltonian H.

The additional term in the perturbed Hamiltonian generates extra terms on the calculation the second of Hamilton's equations for each test system. The resulting equations for the simple pendulum (16) and spring pendulum (17) are given:

$$\frac{d}{dt} p_i(t) = -mgl\sin(\theta) + \lambda(\Delta t)^2 m^2 g^2 2\sin(\theta)\cos(\theta) \tag{16}$$

$$\frac{d}{dt} p_x(t) = -kx(1 - \frac{l_0}{\sqrt{x^2 + y^2}}) + \lambda(\Delta t)^2 2k[\tfrac{k}{m} x(1 - \frac{l_0}{\sqrt{x^2 + y^2}}) + \frac{kgxy}{(x^2 + y^2)^{3/2}}] \tag{17a}$$

$$\frac{d}{dt} p_y(t) = -mg - ky(1 - \frac{l_0}{\sqrt{x^2 + y^2}}) + \lambda(\Delta t)^2 2k[(\tfrac{k}{m} y + 2g)(1 - \frac{l_0}{\sqrt{x^2 + y^2}}) + \frac{kgy^2}{(x^2 + y^2)^{3/2}}] \tag{17b}$$

The scope of this paper does not extend to a solution of the Lennard-Jones fluid using the modified Hamiltonian integrators.

These increased complexities of the differential equations require more 'force'-type calculations per time step. Objective measurement of the extra computational exertion is difficult, but this means the 'modified' integrators require more than the two 'force' calculations per time step than first appearances suggest.

5. Computational Details

Numerical data is computed from original programs written jointly by the authors of this paper. The code was written in C on a Microsoft Windows platform and compiled using the GNU Compiler Collection. The programs output numerical data to files in basic text format. This data is then read into MATLAB, through which all graphs are drawn.

Although the code generalised to allow any physical variables, in this paper we set physical constants to one, unless specified otherwise, so the results are independent of any physical unit. In particular for the simple pendulum conditions (18) are used, and (19) (20) for the spring pendulum and Lennard-Jones fluid respectively.

$$m = g = l = 1 \tag{18},$$

$$m = g = l_0 = 1 \tag{19},$$

$$\varepsilon = 0.25, \sigma = 0.1, k = 1 \text{ and } m_i = 1 \; \forall 1 \le i \le N \tag{20}.$$

6. Results and Discussion – Simple Pendulum

The analysis of the performance of the numerical integrators begins with examining the difference between symplectic and non-symplectic integrators. Figure 4 shows how the total energy of the system (i.e. the Hamiltonian) is constrained in the symplectic Verlet solution, but diverges for the time irreversible Euler solution.

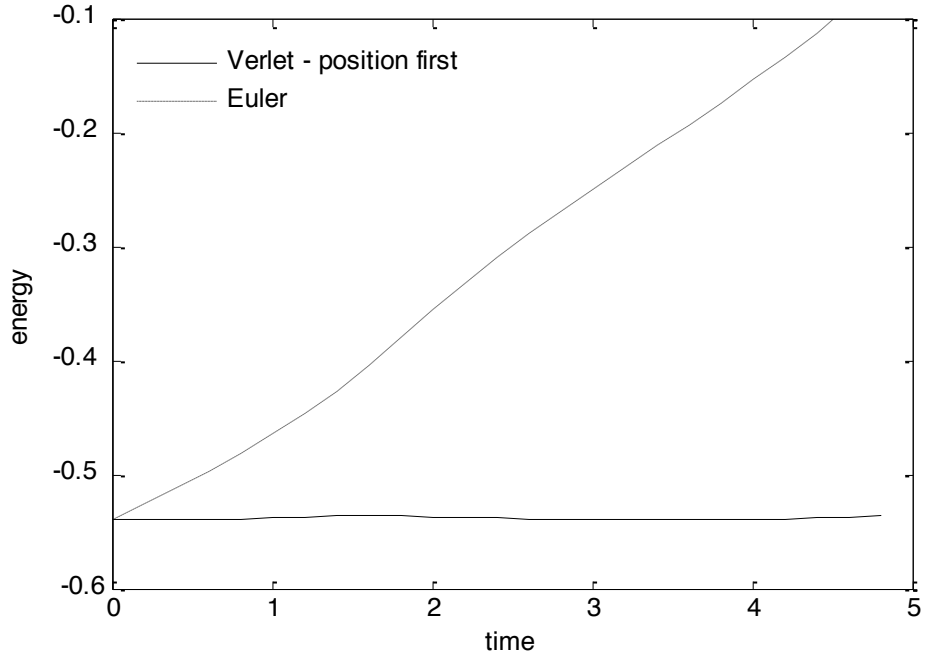

Fig 4. Energy versus time for simple pendulum. Conditions: $\Delta t = 0.2, \theta_0 = 1, p_{\theta_0} = 0, t \in [0,5]$ and (18) with an initial update in position for the Verlet method.

Closer examination of the energy time plots for the Verlet algorithm indicates the numerical energy oscillates within bounds of energy. This behaviour is mirrored in the energy time plots of the three higher order symplectic schemes and the 'modified' Hamiltonian schemes. To quantify this fluctuation, the root mean square variation δE_{RMS} (21) in the solution energies from the approximate Hamiltonian (the mean energy of the numerical solution $\overline{E}$) is calculated for a number in time steps.Let n be the number of data points, then,

$$\delta E_{RMS} = \sqrt{\frac{\sum_{i=1}^{n} (E_i - \overline{E})^2}{n}} \qquad (21),$$

$$\text{where } \overline{E} = \frac{1}{n}\sum_{i=1}^{n} E_i .$$

This measure of energy fluctuation we call the energy error and the results show this declines as the simulation time step is reduced. The theory predicts a $\delta E_{RMS} \propto (\Delta t)^{\gamma}$ proportionality between the energy error and the time step, where γ is the order of the integrator. To test this power law a graph of $\log(\delta E_{RMS})$ versus $\log(\Delta t)$ is plotted and it is checked that the data falls on a straight line. If so a linear line of regression is fitted to the data, with the gradient of the line providing an estimate for the order γ.

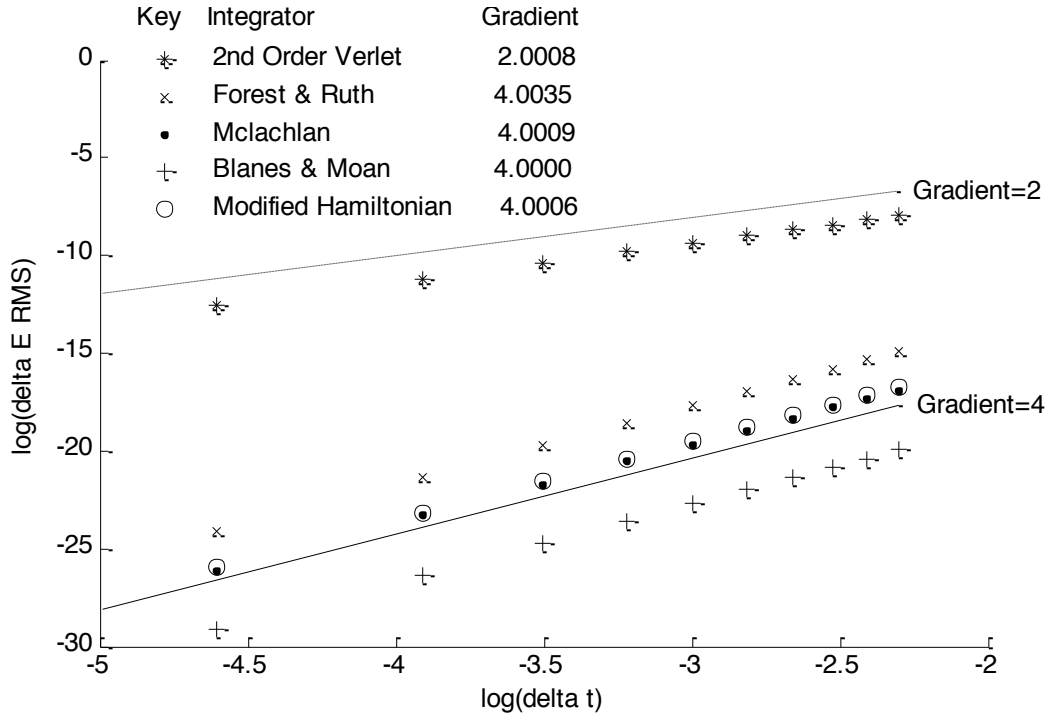

Fig 5. $\log(\delta E_{RMS})$ versus $\log(\Delta t)$, Conditions: $\Delta t = 0.01:0.01:0.1, \theta_0 = 1, p_{\theta_0} = 0, t \in [0,1000]$ and (18) with an initial update in position for all methods. Dashed line overlay has gradient 2, black line overlay gradient of 4. In the legend, the gradients are for the lines of regression fitted to the corresponding data set.

In Figure 5, the notation $\Delta t = 0.01 : 0.01 : 0.1$ means in the simulation Δt begins at 0.01, finishes at 0.1, and takes values of natural multiples of 0.01 in between. The graph shows clear support the theoretical proportionality relationships when taking into account the spread of the data and the gradients of the regression lines. The corresponding graph where the schemes have an initial update of the momentum of the pendulum also supports the theoretical power law linking the energy error and the time step.

As an aside to measure the rate of divergence of the energy of the Euler solutions from the exact Hamiltonian, liner regression lines are fitted to the straight sections of the energy time graphs (from $t = 0$) for a sample of time steps. The gradient of these lines decrease as the time step reduces. A logarithmic plot of the gradients versus time step results in a linear series of gradient one, which suggests the rate of divergence of the energy is directly proportional to the time step. The data supports this theoretical first order performance of Euler's method.

Another measure of the dependence of integrator accuracy on time step is to examine the rate of convergence of the individual coordinates. In the absence of an exact solution, a best estimate for a coordinate value at a given time is obtained from the smallest time step ($\min\{\Delta t\}$) solution available. Then a relative error of higher time step solutions with respect to this approximation of the solution is calculated. In Figure 6 the relative error in the position coordinate theta, $\hat{\theta} = \left|\theta(90) - \theta_{\min\{\Delta t\}}(90)\right|$ at simulation time $t = 90$, is shown as a function of the time step in a logarithmic graph.

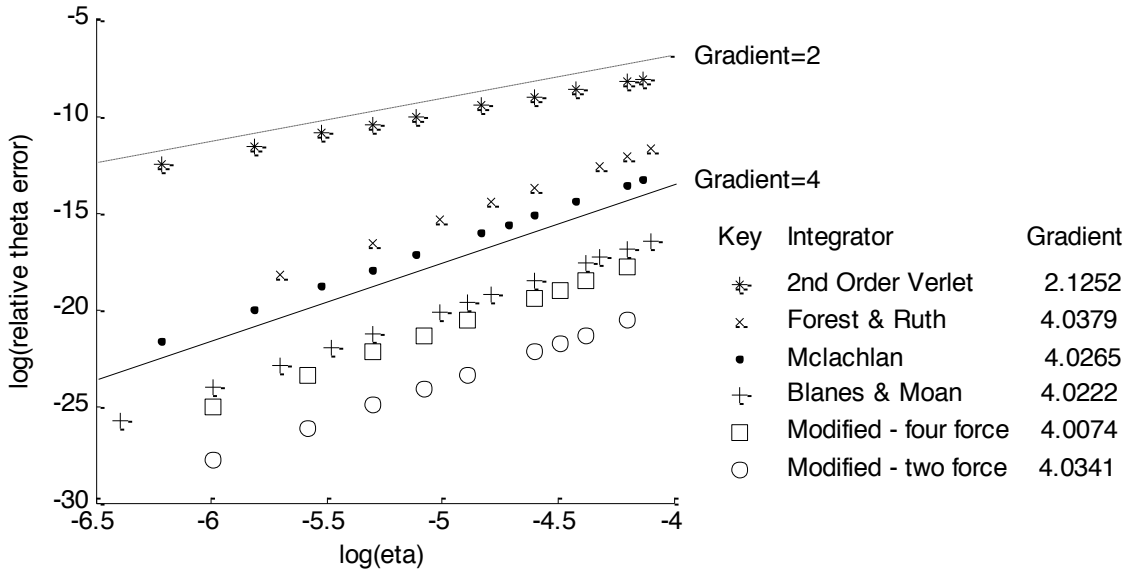

Fig 6. $\log(\hat{\theta}(90))$ versus $\log(\eta)$, Conditions: $\eta \in [0.0015, 0.017], \theta_0 = 1, p_{\theta_0} = 0, t = 90$ and (18) with an initial update in momentum for all methods. Dashed line overlay has gradient 2, black line overlay has gradient of 4.

The measure of the relative coordinate errors is less robust than the measure of the energy errors. As the time step tends to zero the energy error must also become zero, and the solution energy becomes exact. The measure of the relative coordinate errors is less robust however. When dealing with a coordinate error using a time step of a similar size to $\min\{\Delta t\}$, the magnitude of the coordinate error may be of the same order as the absolute difference between the solution coordinate and the exact coordinate value. This could result in the coordinate error data for time steps slightly greater than $\min\{\Delta t\}$ being unreliable.

Figure 6 introduces an effective time step measure η (eta) which accounts for the number of force calculations n_f used by the integrator per overall time step Δt. We assume that each force calculation is the only computational exertion during the running of an integration algorithm, meaning the updates the position coordinates and simulation requires zero processing power. This assumption implies that the computational workload to execute a time step of each integrator is directly proportional to the number of force calculations in that time step. For a given η this allows direct comparison of the magnitude of the errors in the approximate solution between the different integrators where,

$$\eta = \frac{\Delta t}{n_f} \tag{22}.$$

Objectively it is difficult to define the number of force calculations per time step for the modified Hamiltonian schemes because of the extra calculations required in the momenta update steps. The quantity the extra work involved depends firstly on the mathematical model of the system and secondly on the efficiency of the code (the extra terms resulting from the Hamiltonian perturbation in the momenta steps are in some way similar to the standard force calculations). This means there is no universal value of n_f for the modified integrators. To determine the exact value of n_f would involve somehow measuring the computational workload during the simulation, however such depth is considered superfluous in this study as one must not lose sight or appreciation of the initial assumption that all calculations, bar the force, are negligible.

On the most basic level the modified integrators require two standard force calculations per time step. Examination of the extra calculation involved in the pendulum systems leads us to a reasonable estimate of a total of four force calculations per overall time step. However conservative bounds for n_f would be between two and seven force calculations

per time step, and these are borne in mind when comparing the modified schemes to the standard fourth order integrators.

Returning to Figure 6 it is clear that the second Verlet algorithm is uncompetitive against the fourth order schemes with coordinate errors of at least three orders of magnitude greater at fixed η. This result is confirmed by the corresponding relative momentum error, $\hat{p}_\theta$, graph of Figure 6 and the η equivalent graph of Figure 5 (i.e. $\log(\delta E_{RMS})$ versus $\log(\eta)$).

For the simple pendulum there exists an exact solution for small amplitude oscillations, where the pendulum displays simple harmonic motion with a constant period of motion given by (23). From the numerical simulations at small amplitudes, a period is interpolated from the solution data by finding the time between two successive changes in sign of the momentum. When these zero values of momentum fall in between two data points, the zero time is found by linear interpolation. The error in this interpolation is small when the period of motion is very much greater than the time step Δt. The absolute difference between the interpolated period and the exact, from (23), can be used as a measure of the error in the numerical solution.

$$\mathrm{T} = 2\pi\sqrt{\frac{l}{g}} \tag{23}$$

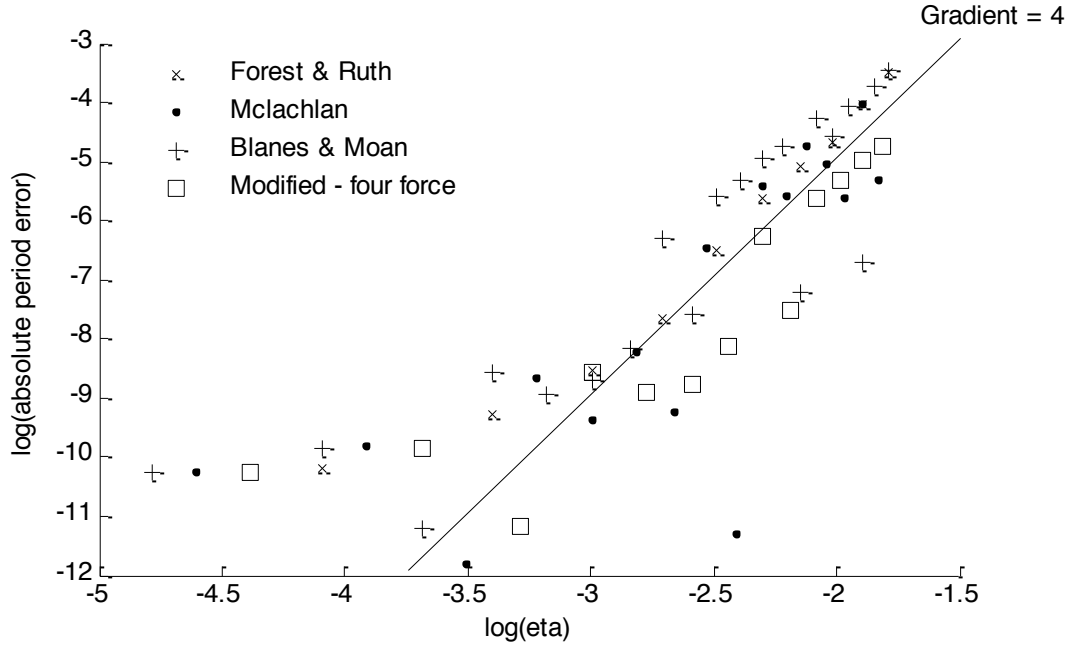

Fig 7. log(absolute period error) versus $\log(\eta)$, Conditions: $\eta \in [0.08, 0.17]$, $\theta_0 = 0.01$, $p_{\theta_0} = 0$ and (18) with an initial update in momentum for all methods. Exact period $\mathrm{T} = 2\pi$. Black line overlay has gradient of 4.

Figure 7 shows the absolute period error measure to be less stable than previous error measures. For larger values of η the linear interpolation of the momentum to find the period introduces a noise to the data. It is worth noting that at the zero value of the momentum, the acceleration of the bob is at greatest so the momentum value is changing at its fastest rate, which increases the possible error in the linear interpolation. Interpolating the zero value of the angle theta does not avoid this problem as the velocity of the bob is greatest at this point during the motion. However the section of the graph at larger values of η does suggest a fourth order power law relating the period error to the time step, as shown by the overlaid black line of gradient four. At smaller values of η the convergence breaks down. This is due to the small angle approximation used in the derivation of the exact solution. As a result at the smaller time steps the integrators converge to a period of motion which is exceeds the accuracy of the $T = 2\pi$ period calculated from (23).

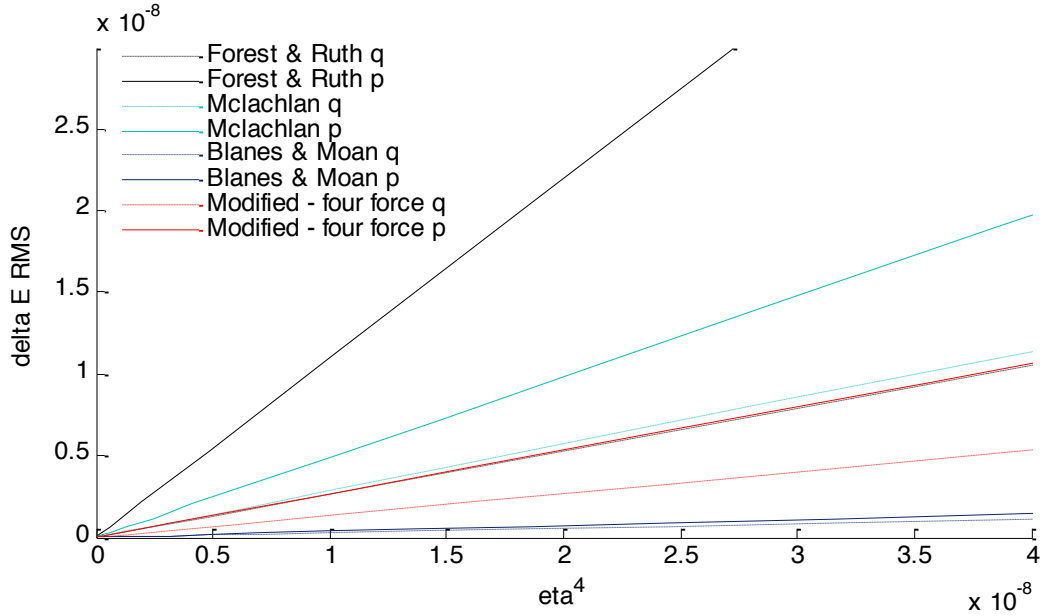

Fig 8. δE_{RMS} versus η^4, Conditions: $\eta \in [0.0008, 0.02], \theta_0 = 1, p_{\theta_0} = 0, t \in [0,1000]$ and (18). Four force calculation modified Hamiltonian integrators. Discrete data represented as continuous lines by joining up data points with straight lines as to bring out approximate gradients of the data – the data set for each line contains ten values equally spaced across the axis of η^4. Initial update in position (q) integrators are represented by dashed lines, initial update in momentum integrators (p) by solid lines.

Now we are in a position to compare the fourth order integrators in terms of performance. Theory says the error measures are directly proportional $(\Delta t)^4$, along with η^4 by adjustment of the constant of proportionality, and numerical evidence for the fourth order dependence is clear. Interest transfers to the constant of proportionality in the power law

relationship, which depends explicitly on the system and its physical conditions. The smaller this constant is, the smaller the magnitude of the errors in the numerical solution and hence the more accurate the integrators will be. Graphs of the error measures versus η^4 will have a gradient of this constant.

Figure 8 is a plot of the energy error versus η^4 and the first clear result is that the Blanes & Moan integrator is the most efficient of the standard fourth order schemes, under the specified conditions, as its graph has the shallowest gradient. It is difficult to draw a distinction between the Forest & Ruth and the Mclachlan integrators as the result depends on the initial coordinate updated in the scheme. For the modified Hamiltonian methods the lie of the results with respect to the standard fourth order integrators depends on the value of n_f used. This is discussed more in Figure 9, but for $n_f = 4$ the results do not indicate that the modified Hamiltonian integrators are in any way superior.

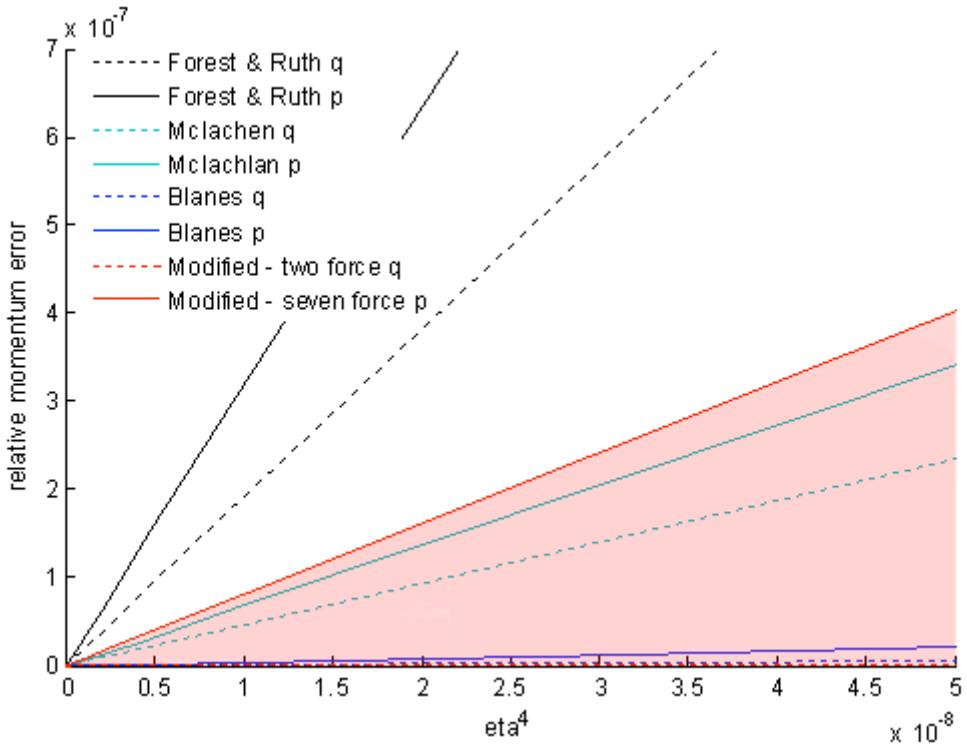

Fig 9. $\hat{p}_\theta(92)$ (relative momentum error) versus η^4, Conditions: $\eta \in [0.0008, 0.02], \theta_0 = 1, p_{\theta_0} = 0, t = 92$ and (18). The red shaded region depicts the spread of modified Hamiltonian integrator results depending on the value of n_f chosen from the region $[2,7]$. Discrete data represented continuously as in Figure 8.

Figure 9 addresses directly the discrepancy regarding the value of n_f used for the modified Hamiltonian schemes. The red shaded region shows the spread of possible graphs for the modified Hamiltonian integrators depending on value of n_f used from the suggested

bounds of $[2,7]$. This could mean the modified integrators are the best performing, or the second worst behind the Forest & Ruth scheme. This solidifies the conclusion drawn from Figure 8 – there is no evidence to suggest the new modified integrators show an improvement in computational performance terms over the standard fourth order integrators. Figure 9 also shows a clear distinction between the Forest & Ruth and Mclachlan integrators, with the latter performing more strongly.

6.1 Results and Discussion – Spring Pendulum

The motion exhibited by the spring pendulum is more complex than that of the simple pendulum, with three non-trivial types of motion displayed: periodic, quasiperiodic and chaotic. The system involves a total of four position and momentum coordinates, which allow Poincaré sections to be used to identify the underlying nature of the motion. Interest is taken in the special case, $K = 3$, where the natural frequency of transverse motion is twice the natural frequency of vertical oscillations and at certain energies chaotic motion is displayed.

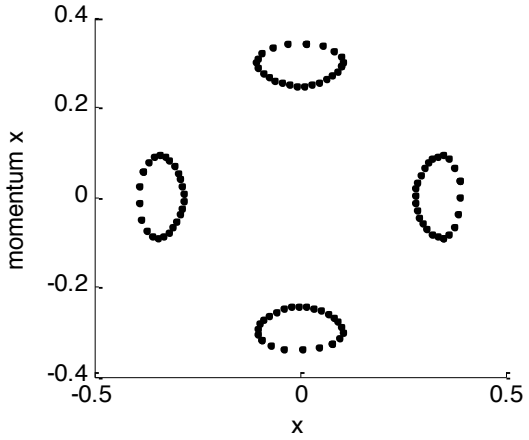

Fig 10a. Poincaré section showing quasiperiodic motion. Conditions: $x_0 = 0.16113, y_0 = -1.21307$, $p_{x_0} = 0.46079, p_{y_0} = 0.25229$ with $H = -1$.

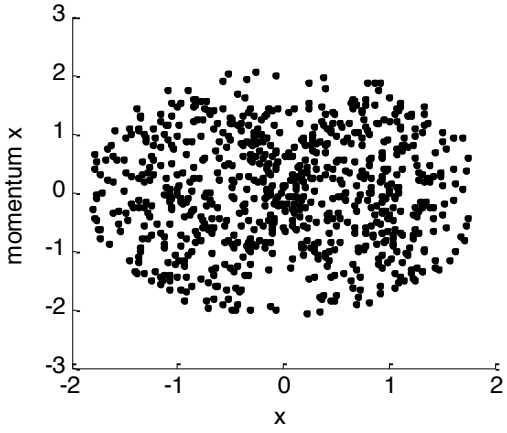

Fig 10b. Poincaré section showing chaotic motion. Conditions: $x_0 = 0.11673, y_0 = -1.52770$, $p_{x_0} = 1.05101, p_{y_0} = 1.76103$ with $H = 1$.

Uniform conditions for Figures 10a and 10b: $\Delta t = 0.01, t \in [0,200], K = 3$ and (19). Data generated using 2nd Order Verlet integrator with an initial update in position.

Analysis of the Poincaré sections indicates that the type of motion exhibited depends on the Hamiltonian of the system. For the absolute minimum system energy, $H_{\min} = -mg(\frac{mg}{k} + l_0)$, the motion is trivial as the bob is fixed in a stable equilibrium position, represented by a single point on a Poincaré section. Increasing the system energy results in

two types of motion being displayed, periodic and quasiperiodic. Periodic motion is shown by two or more fixed points on a Poincaré section and quasiperiodic by bounded curves in the phase space (Figure 10a). Further increase of the system energy can lead to chaotic motion, along with periodic and quasiperiodic, which is represented by a bounded region of points on the Poincaré section (Figure10b). At the high energy limit, the spring pendulum reverts to a periodic orbit, as the energy stored in the spring dominates the motion of the bob.

At a given energy the type of motion displayed depends on the system trajectory, and hence the starting conditions of the system. The type of motion displayed can be identified by drawing a Poincaré section from the numerical data. The knowledge borne from this analysis means the system can be initialized in a trajectory of displaying any type of motion, meaning integrator performance can be tested on each.

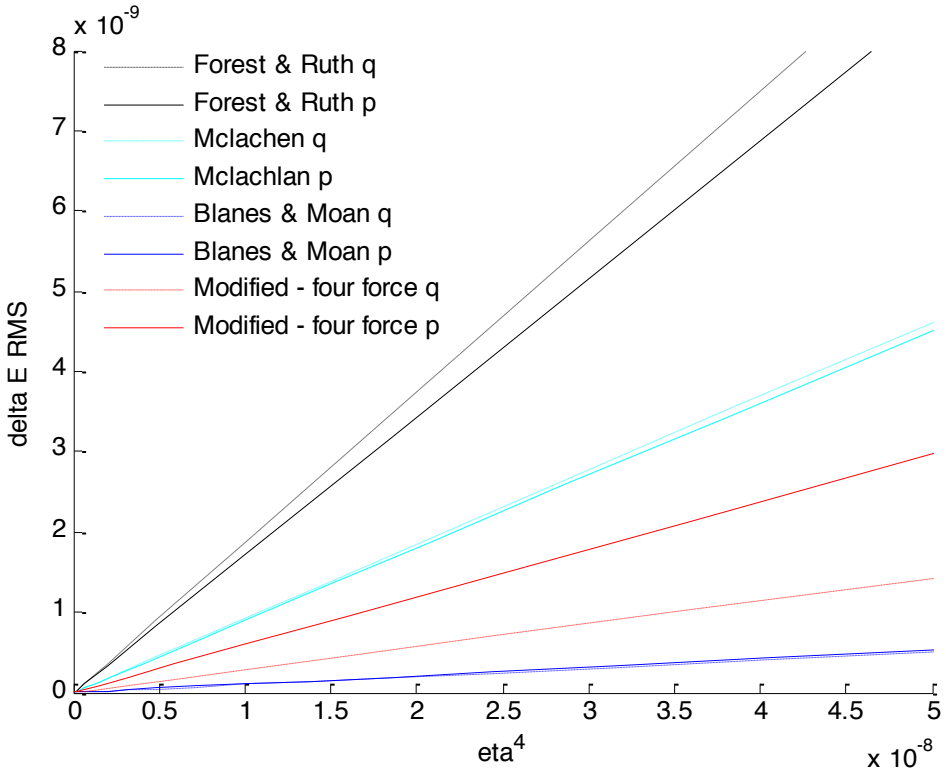

Fig 11. δE_{RMS} versus η^4, Conditions: $\eta \in [0.0008, 0.02]$, $x_0 = 0.2$, $y_0 = -1.4$, $p_{x_0} = p_{y_0} = 0$, $t \in [0,105]$, $K = 3$ and (19). Spring pendulum displaying quasiperiodic motion. Discrete data represented as continuously with ten equally spaced data points per line.

The story of the integrator performance on the simple pendulum transfers well to the results seen for the spring pendulum in periodic, quasiperiodic and chaotic motion under most conditions. The energy time graphs of the numerical solutions for these types of motion oscillate and are bounded (Figure 12a). Direct comparison of integrator performance using the

energy is shown in Figure 11, with the modified Hamiltonian integrators once again failing to outperform the standard fourth order integrators. It is interesting to note that the standard fourth order integrators have 'paired up', with position first and momentum first orientations showing similar performance, whereas the position first modified integrator clearly outperforms the momentum first. This is a consequence of the two modified schemes being mutually exclusive. The corresponding coordinate error graphs support Figure 11.

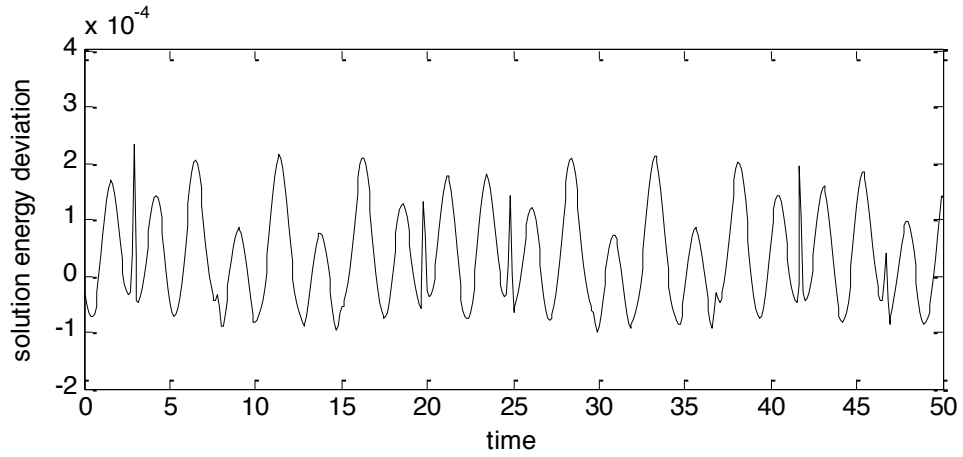

Fig 12a. Solution energy deviation from exact Hamiltonian, $H = 3.1300$, versus time. Conditions: $\Delta t = 0.1, x_0 = -0.9722$, $y_0 = -1.7120$, $p_{x_0} = 2.6193$, $p_{y_0} = 0.0883, t \in [0,50], K = 3$ and (19). System in quasiperiodic motion.

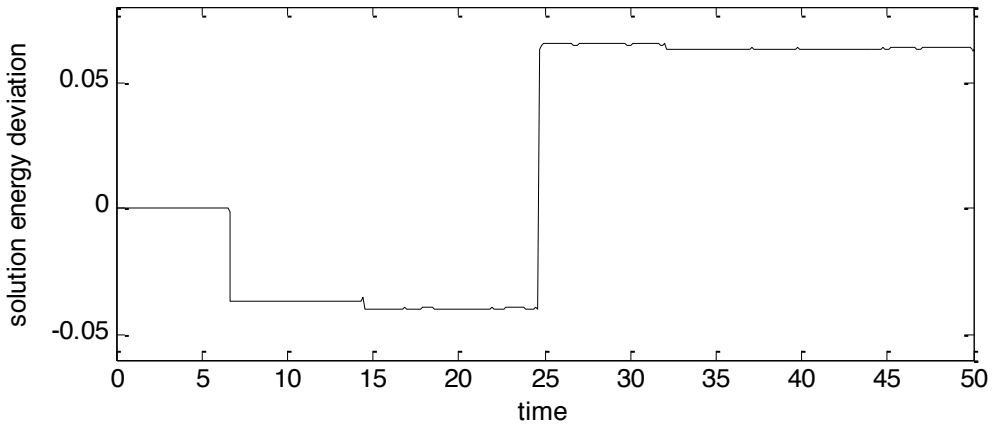

Fig 12b. Solution energy deviation from exact Hamiltonian, $H = 3.1300$, versus time. Conditions: $\Delta t = 0.1, x_0 = 0.1$, $y_0 = 1.5$, $p_{x_0} = -0.5$, $p_{y_0} = 1.5, t \in [0,50], K = 3$ and (19). System in chaotic motion. Notice the discrete jumps in energy of the solution, occurring when the trajectory is close to the origin.

We now discuss the trajectories of the motion which pass close to the origin, $(0,0)$, of the system. The magnitude of the errors in the numerical solution for these trajectories is appreciably larger than solutions to trajectories always away from the origin, even with the same energy, physical system variables (not initial conditions!) and integration scheme. This

is depicted by Figures 12; in 12a a system of exact energy (and hence initial energy) of 3.1300 undergoes a quasiperiodic trajectory. Figure 12b also shows a trajectory, this time chaotic, of energy 3.1300, with a solution is obtained using the same integration scheme and parameters as in 12a. The graphs show of the deviation of the solution energy from the exact energy (3.1300). There are clear discontinuities in the energy deviation of Figure 12b, resulting in an energy fluctuation one hundred times larger in magnitude than the trajectory in Figure 12a, resulting in a significantly less accurate solution. On an appropriate scale, the energy fluctuations between the discontinuities in Figure 12b are of the same order as those in 12a. Correlation of the times of the discontinuities to the solution data shows they occur when the trajectory approaches the origin. Conversely, for all simulations researched, whenever the trajectory approaches the origin, a discrete jump in energy is witnessed.

The reason why these jumps occur is clear on re-examination of Hamilton's equations for the spring pendulum (6). For the second equations, which describe the rate of change of the momenta, a singularity occurs at the origin (i.e. when $x = 0$ and $y = 0$). This creates problems for the numerical integrators dealing with trajectories close the origin due to the steepness of the $(x^2 + y^2)^{-0.5}$ term. This term accentuates the errors introduced in the discretization of continuous differential equations, reducing the accuracy of all the integrators.

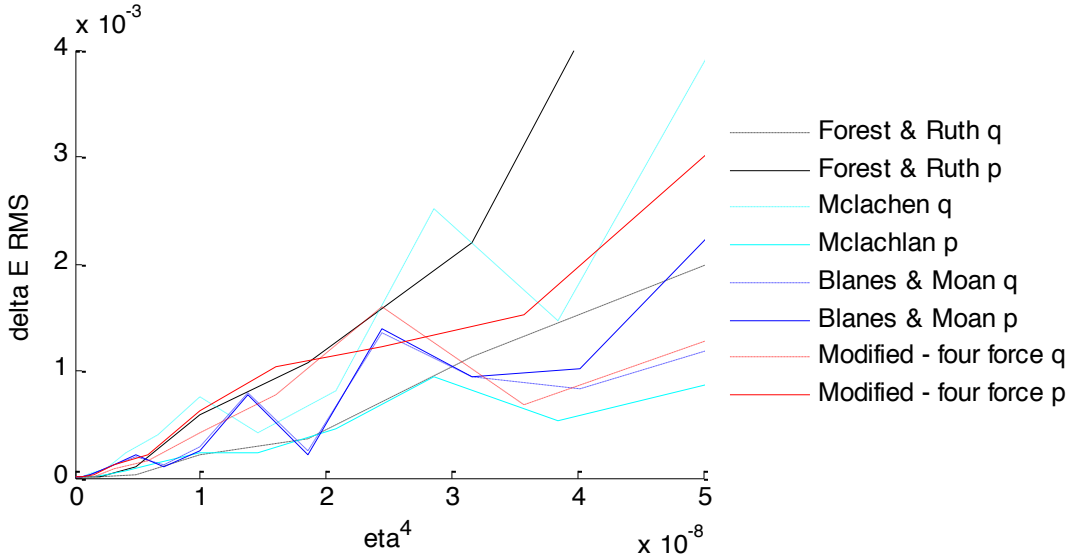

Fig 13. δE_{RMS} versus η^4, Conditions: $\eta \in [0.0008, 0.02]$, $x_0 = 0.1$, $y_0 = 1.5$, $p_{x_0} = -0.5$, $p_{y_0} = 1.5$, $t \in [0,100]$, $K = 3$ and (19). System displaying chaotic motion with discrete energy jumps in the solutions.

Comparing Figure 13 with Figure 11 it is clear the energy jumps are causing a detrimental effect on integrator performance. In Figure 13 the energy errors are five orders of magnitude greater for a given value of η. However the use of the energy error to measure the

effect of the jumps is questionable, as the solution energy is now unbounded and hence measure is heavily dependent on the trajectory (for example how many times the origin is approached in the given time period) and the time step.

To quantify the effect of the jumps on a numerical solution to such a trajectory, we use a statistical approach. Jump locations are established by detecting where the solution trajectory is arbitrarily close to the origin and then the size of the jump in energy is calculated from the differences in mean energies between the jumps. We then develop a measure for the 'average' jump size over the simulation and see how this is related to the time step.

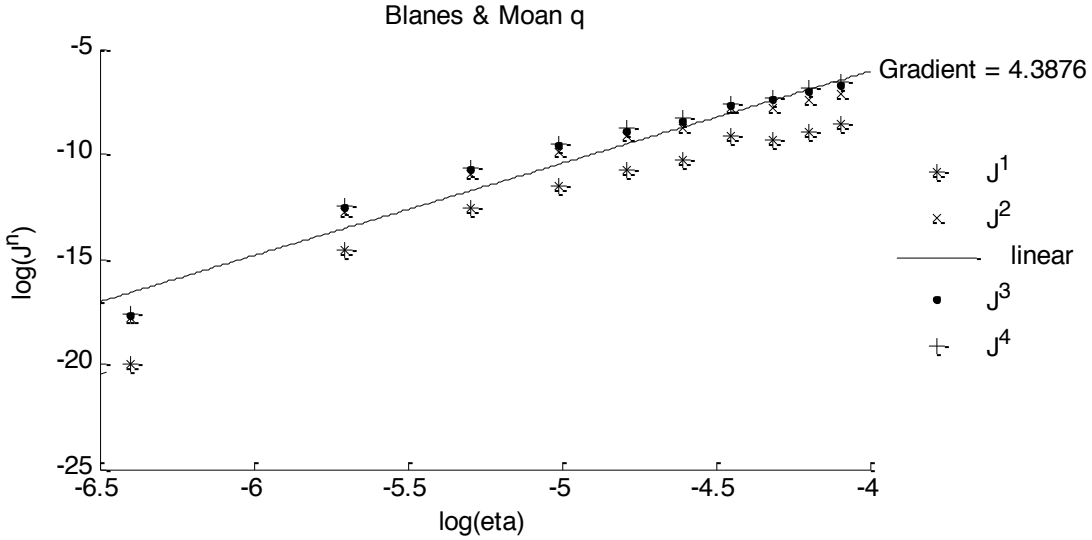

Fig 14a. Graph of log(jump measure) versus $\log(\eta)$ for Blanes & Moan integrator with initial update in position.

Conditions: $\eta \in [0.00015, 0.0017], x_0 = 0.1, y_0 = 1.5, p_{x_0} = -0.5, p_{y_0} = 1.5, t \in [0, 10000], K = 3$ and (19).

Number of jumps in simulation $= 299$. Black overlay of linear best fit for J^2.

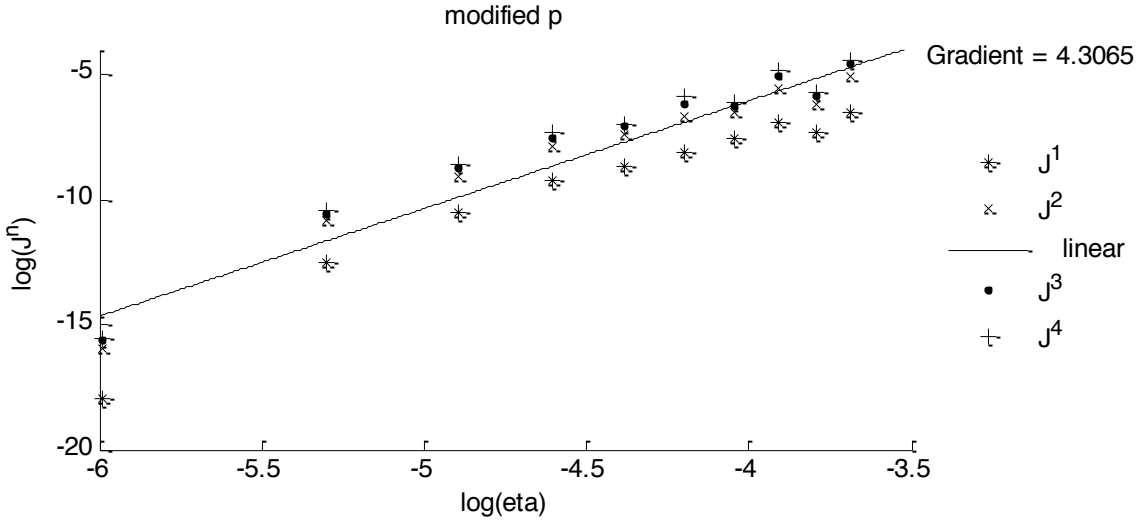

Fig 14b. Graph of log(jump measure) versus $\log(\eta)$ for modified integrator with initial update in momentum.

Conditions: $\eta \in [0.00015, 0.0017], x_0 = 0.1, y_0 = 1.5, p_{x_0} = -0.5, p_{y_0} = 1.5, t \in [0, 10000], K = 3$ and (19).

Number of jumps in simulation $= 314$. Black overlay of linear best fit for J^2.

A mean of the jump sizes could be distorted by a number of 'small' jumps with relatively little impact on the dynamics, when in fact the larger jumps cause the greatest errors in the solution due to the drift in the solution energy from the exact Hamiltonian. To account for this phenomenon 'weighted' averages of the jumps are calculated,

$$J^n = \frac{\sum_k {j_k}^n}{\sum_k {j_k}^{n-1}} \tag{24},$$

where j_k are individual jump sizes and k indexes all the jumps in the simulation. As n increases the average is skewed towards accounting for just the larger jumps.

Figures 14 show, as expected, that the jump sizes reduce as the time step decreases. As n increases the jump size measures J^n converge towards the limit of the largest jump size in the simulation. Jump measure J^2 is taken as the best measure of the average jump size with a tangible contribution to the energy errors. The overlaid black lines on the graphs are linear lines of best fit for J^2. Depending on the initial conditions of the system and the range of time steps examined the gradient of these lines vary, which indicates that a power law relationship between the jump size measure J^2 and the time step measure (Δt or η) is incorrect. However this is to be expected as the theoretical convergence measures do not apply for non-analytic functions. There is also no evidence for any integrator handling the computational problems caused by the singularity better than any others, which agrees with the breakdown of the error measure graphs (Figure 13).

A method for examining the extent to energy jumps are degradation in the accuracy of the numerical solution is proposed. An energy fluctuation $(\delta E_{RMS})_k$ is calculated from the sections of the trajectory in between the energy jumps, i.e. the sections of the motion in between points where the trajectory approaches the origin. A time weighted average over these sections is calculated forming an effective energy fluctuation value for the solution,

$$\delta\widetilde{E}_{RMS} = \frac{\sum_k t_k (\delta E_{RMS})_k}{\sum_k t_k} \tag{25},$$

where t_k is the length of time of section k. Clearly this effective energy fluctuation measure is always less than the overall energy error ($\delta\widetilde{E}_{RMS} \leq \delta E_{RMS}$). The smaller the difference between these two values, the lessening effect of the energy jumps on the solution accuracy. Future

study could examine how the magnitude of this difference translates to the stability and accuracy of the solution.

6.2 Results and Discussion – Lennard-Jones Fluid

A Lennard-Jones fluid of *N* particles has a total of *4N* positions and momenta. During the initialization of the simulation particles are randomly placed in a cube centred on the origin.

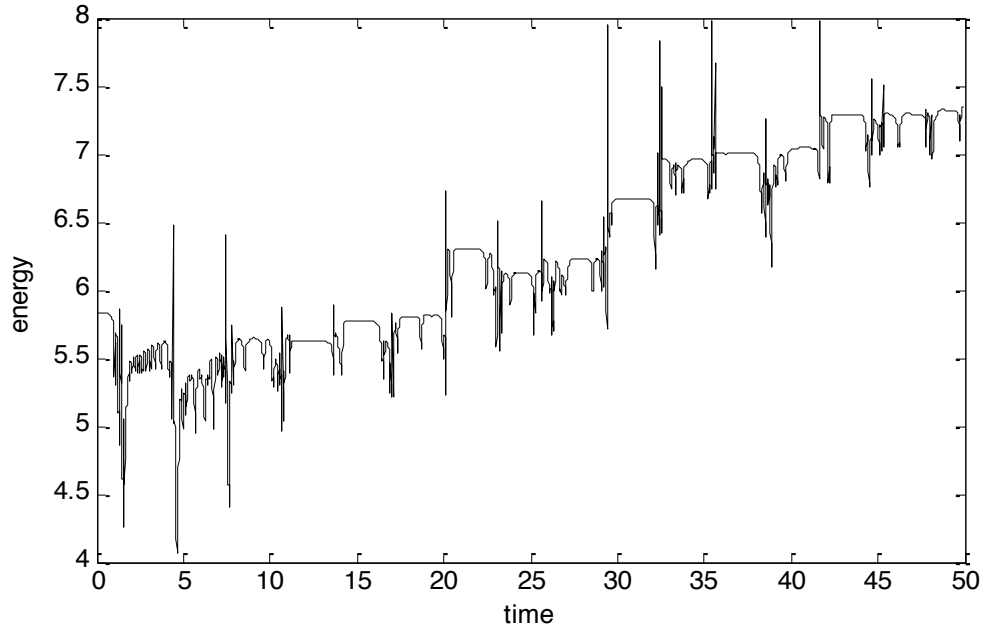

Fig 15. Graph of energy versus time for ten particle Lennard-Jones fluid. Particles initialised in random positions within a cube of side length 2 centred on the origin, with zero initial momentum. Simulation integrator Blanes & Moan position first. Conditions: $\Delta t = 0.01, t \in [0,50]$ and (20).

Figure 15 displays the solution energy for a Lennard-Jones simulation. It shows similar traits to the Figure 12b, with discrete jumps in the solution energy and relatively small oscillations in the energy in between the jumps. For two particles within a short range there is a strong repulsive force, resulting from the singularity in the Lennard-Jones potential at zero separation, which acts to separate them. Like the singularity in Hamilton's equations for the spring pendulum, this accentuates the errors introduced in discretization, causing problems for all the integrators.

The theoretical rates of convergence for the error measures are not mirrored in any of the numerical results, except in trivial cases. The accuracy of the solutions over long time periods is questionable, as the solution energy often drifts far from the exact Hamiltonian, in a situation similar to that of the time-irreversible Euler integrator on the simple pendulum.

7. Conclusions

The first result we note is that symplectic integrators produce more accurate solutions to time irreversible integrators on the simple pendulum system. On the same system the fourth order symplectic integrators are more economical in computational terms than the second order Verlet scheme. There is no evidence to support the claim that the modified Hamiltonian integrators are more efficient than any of the standard fourth order symplectic integrators, other than the most basic of which by Forest & Ruth.

The spring pendulum is described by non-analytic ordinary differential equations. These are continuous everywhere but the origin and all the results for the simple pendulum in the above paragraph transfer to numerical solutions to trajectories everywhere away from the origin singularity, regardless of the type of motion displayed by that trajectory. To stress again, there is no case for the modified Hamiltonian integrators having the highest computational efficiency.

We note the performance of all the integrators reduces for trajectories which approach the origin singularity, and because of this the theoretical performance measures lose their value in the analysis of the integrators. The numerical solutions resulting from such trajectories display discrete jumps in energy, which cast doubt over validity of the solution over long time periods of simulation. Numerical solutions for the Lennard-Jones fluid also display discrete energy jumps in the solution for all but the most trivial cases. There is no case for any of the integrators dealing with the jumps best, which indicates that using smaller time steps for the most basic symplectic integrators may prove more efficient than employing more computationally expensive schemes. Future research could test this claim and study further the performance of the numerical integrators on non-analytic ordinary to differential equations in order establish the most computationally efficient approach to solving them.

8. References

Allen, M.P., Tildesley, D.J., 1987. *Computer Simulation of Liquids*. Oxford: Clarendon Press.

Ball, R.C., 2007. An Augmented Verlet Algorithm. Preprint.

Blanes, S., Moan, P.C., 2000. Practical Symplectic Partitioned Runge-Kutta and Runge-Kutta-Nyström Methods. *Journal of Compuational and Applied Mathematics,* 142 (2), 313-330.

Forest, E., Ruth, R.D., 1990. Fourth-Order Symplectic Integration. *Physica,* 43 (1), 105-117.

Leimkuhler, B., Reich, S., 2004. *Simulating Hamiltonian Dynamics.* Cambridge: Cambridge University Press.

Mclachlan, R.I, 1995. On the Numerical Integration of Ordinary Differential Equations by Symmetric Composition Methods. *Scientific Computing,* 16 (1), 151-168.

Skeel, R.D., Hardy, D.J., 2001. Practical Construction of Modified Hamiltonians. *Scientific Computing*, 23 (4), 1172-1188.

Skeel, R.D., Zhang, G., Schlick, T., 1997. A family of symplectic integrators: stability, accuracy, and molecular dynamics applications. *Scientific Computing,* 18 (1), 203-222.