

Long time scale molecular dynamics using least action.

B. A. Gladwin* T. Huber†

(Received 8 August 2003)

Abstract

We present here an efficient method for evaluating molecular trajectories over long time scales. The method is based on optimisation of the path action defined by classical mechanics. We test the technique on non-trivial examples drawn from the literature and discuss the effectiveness of this approach in the study of molecular processes. Many of the present techniques for calculating molecular trajectories are limited computationally. Standard forward integration of Newton's equations of motion yields accurate results for a range of systems whose transition times are many orders of magnitude less than most biologically interesting processes. If one wants to extend these

*Department of Mathematics, University of Queensland, Brisbane, AUSTRALIA.
<mailto:gladwin@maths.uq.edu.au>

†Department of Mathematics, University of Queensland, Brisbane, AUSTRALIA
See <http://anziamj.austms.org.au/V45/CTAC2003/Glad> for this article, © Austral. Mathematical Soc. 2004. Published 26 June 2004, amended July 8, 2004. ISSN 1446-8735

calculations to biologically relevant time scales, it is necessary to develop methodologies which avoid this limitation. The process outlined in this paper has been tested on simple systems using harmonic and Lennard–Jones potential energy functions. The algorithm yields stable trajectories and is adjustable to suite available computational resources. In theory, this algorithm is applicable to any molecular system where the initial and final states are known. This could include investigation of chemical reactions, ligand/receptor binding and work cycles of molecular machinery.

Contents

1 Introduction	C535
2 Theory	C537
3 Results and discussion	C541
3.1 1D harmonic potential, two particle system	C541
3.2 Lennard–Jones potentials	C541
3.3 Experiment 2: three particle system	C544
3.4 Experiment 3: seven particle system	C547
4 Conclusions	C547
References	C549

1 Introduction

Mathematical modelling of protein structures aims to understand the link between structure and function in the biomolecular processes behind cellular mechanics. Key ideas and methodologies of previous studies have focused on

‘real time’ molecular transitions using classical dynamics. By simulating the molecules environment, the dynamics can be approximated by numerically solving Newton’s equation of motion. These processes provide insight into the behaviour of simple structures but are limited computationally to time scales that are often not biologically relevant.

One goal of the present research is to avoid the computational constraints that simulation of real time molecular transitions presents. Techniques which have contributed to solving this limitation include Targeted Molecular Dynamics (TMD) [2, 3], statistical study of rare events [4, 5, 6] and a group of techniques which we refer to as *Action Derived Molecular Dynamics* (ADMD) [7, 9, 8]. All of these exploit the knowledge of the molecule’s starting and ending conformations to target the biologically relevant pathways and guide the path search.

ADMD uses the classical Action of a trajectory as a means of limiting the path search space. The principle of least Action is attributed to Hamilton and states that for any real path taken by a system of particles the variation in the Lagrangian (Kinetic minus potential Energy) is zero.

$$\mathcal{S}' = \delta \int_{t_{\text{initial}}}^{t_{\text{final}}} \mathcal{L}(\mathbf{x}(t), \dot{\mathbf{x}}(t)) dt. \quad (1)$$

Mathematically, equation (1) means that the Action will take an extremum for any real path.

If $\mathbf{x}(t)$ is known, the kinetic energy of each particle can be determined from the first derivative with respect to time of the path and the particles mass. The potential energy for a particle is more complex as it is dependent on interactions with all other particles in the system. For this paper we look at only two particle interaction types. There is no loss of generality by limiting our discussion in this way providing all relevant interactions can be expressed in terms of the positions of each particle and physical constants such as the mass.

The target of the present work is to calculate long time scale paths in biomolecular systems. Defining the paths approximately in terms of a controllable set of parameters, a reformulation of the Action principle is applied which uses the error between the approximate path and the physical path in place of the Action. By adjusting the path, we find the path which minimises this error. This approach successfully locates paths satisfying the physical constraints on the systems. Further work will determine the robustness of this algorithm for situations of higher complexity.

2 Theory

Olender and Elber [1] reformulate the least action principle in terms of the error between an approximate path and the physical path. In this section key elements of this approach are highlighted and the details of our algorithm discussed.

Any path defined by the boundary conditions $[\mathbf{x}(t_{\text{initial}}), \mathbf{x}(t_{\text{final}})]$ can be approximated by the set of coordinates $\mathbf{x}_{\text{approx}}(t)$. The probability of a transition between a pair of these coordinates $(\mathbf{x}_{\text{approx}}(t_i), \mathbf{x}_{\text{approx}}(t_j))$ in time $\Delta t = t_j - t_i$ is the conditional probability

$$\Pr(\mathbf{x}_{\text{approx}}(t_j) \mid \mathbf{x}_{\text{approx}}(t_i); \Delta t) .$$

The transition which maximises this probability is the nearest transition to the true trajectory. By applying this reasoning to each transition along the trajectory the most likely path is located. If we assume that the error arising from the discrete step is Gaussian distributed around the true trajectory, then the error correlation between two steps

$$\langle \epsilon(t_i) \epsilon(t_j) \rangle \approx \frac{\sigma^2 \delta_{ij}}{\Delta t} , \quad (2)$$

and we define the conditional probability of a particular transition as

$$\begin{aligned} & \Pr(\mathbf{x}(t_i) \mid \mathbf{x}(t_i + \Delta t); \Delta t) \\ &= \left(\frac{1}{2\pi\langle\epsilon^2\rangle} \right)^{d/2} \exp \left[-\frac{(\mathbf{x}(t_i + \Delta t) - \mathbf{x}_{\text{approx}}(t_i + \Delta t))^2}{2\langle\epsilon^2\rangle} \right]. \end{aligned} \quad (3)$$

This in turn allows us to formulate an expression for the conditional probability of an entire trajectory as the product of all of the transition probabilities of that trajectory.

$$\prod_i \Pr(\mathbf{x}(t_i) \mid \mathbf{x}(t_i + \Delta t); \Delta t) \approx \exp \left[-\frac{1}{2\sigma^2} \int (\epsilon(t))^2 dt \right]. \quad (4)$$

Equation (4) represents a single exponential function in which the probability that a certain trajectory is correct expressed in terms of the error from the true path. The path with zero error at each point will have the highest probability. Onsager and Machlup state that the probability that a given trajectory $\mathbf{x}(t)$ will satisfy Langevin's equation of motion is

$$\Pr[\mathbf{x}(t)] \propto \exp \left[\frac{-\mathcal{S}}{k_B T} \right], \quad (5)$$

where the k_B is Boltzmann's constant and T is the temperature of the system and when this expression is compared to equation 4 one directly arrives at the expression for the Onsager–Machlup Action

$$\mathcal{S} = \int (\epsilon(t))^2 dt. \quad (6)$$

It is now necessary to formulate a function for the error in the path. As in Olender and Elber's paper we choose to formulate this error based on Newton's equations of motion. For the correct trajectory, we assume that Newton's equation motion holds at every time point along the path:

$$\mathbf{F}_i - m_i \ddot{\mathbf{x}}_i = 0.$$

In an approximate path the right hand side of this equation is equal to the error in the path $\epsilon(t)$ which is non-zero and the Action becomes

$$\mathcal{S} = \int (\mathbf{F}(t) - \mathbf{m}\ddot{\mathbf{x}}(t))^2 dt. \quad (7)$$

The advantage of this formulation over the one proposed by Hamilton is immediately apparent. A physical path will have zero Action rather than being an extremum and the path search becomes a well defined global minimisation problem. It remains now to specify an approximate path in terms of a parameter set and apply standard minimisation algorithms.

In this analysis we choose to expand the path as a linear trend plus a Fourier sine series:

$$\mathbf{x}(t) = \mathbf{x}_0 + \mathbf{a}t + \sum_i b_i \sin\left(\frac{i\pi t}{\tau}\right), \quad (8)$$

where \mathbf{x}_0 is the initial positions of the particles, \mathbf{a} represents an initial linear approximation to the path, τ is the total time of the event and \mathbf{b} are the Fourier expansion coefficients for each particle. This type of expansion allows effective filtering of high frequency oscillations by removing terms in the expansion, and therefore lends itself to the use of hierarchical path searching regimes. Successive minimisation runs with increasing numbers of expansions allows efficient minimisation by first optimising the gross movements of the molecule using lower modes and making finer adjustments using higher modes.

Since the motion of the a particle is governed by its interactions with surrounding particles, a potential energy function which captures these interactions needs to be specified. As mentioned previously, two standard energy functions will be investigated. The first of these is the harmonic (bond) potential function

$$\mathcal{V}(\mathbf{r}_i, \mathbf{r}_j) = \frac{1}{2}k_b (r_{ij} - r_{\text{equilib}})^2, \quad (9)$$

where k_b is the elastic coefficient of the bond and r_{equilib} is the equilibrium length of the bond.

The second potential function investigated in this paper is the Lennard–Jones interaction function to simulate non-bonded atom interactions.

$$\mathcal{V}(\mathbf{r}_i, \mathbf{r}_j) = \frac{B_{ij}}{r_{ij}^{12}} - \frac{A_{ij}}{r_{ij}^6}, \quad (10)$$

where $r_{ij} = |\mathbf{r}_j - \mathbf{r}_i|$ and interactions exist across all pairs of atoms. Given these interaction functions we numerically approximate the Action and its derivative with respect to the expansion coefficients \mathbf{b} along a path

$$\begin{aligned} \mathcal{S} &= \sum_t (\mathbf{F}(t) - \mathbf{m}\ddot{\mathbf{x}}(t))^2, \\ \frac{\partial \mathcal{S}}{\partial \mathbf{b}} &= 2 \sum_t (\mathbf{F}(t) - \mathbf{m}\ddot{\mathbf{x}}(t)) \left(\frac{\partial \mathbf{F}(t)}{\partial \mathbf{b}} - \mathbf{m} \frac{\partial \ddot{\mathbf{x}}(t)}{\partial \mathbf{b}} \right), \end{aligned}$$

where

$$\mathbf{F}(t) = -\frac{\partial \mathcal{V}(t)}{\partial \mathbf{x}},$$

and the $\ddot{\mathbf{x}}(t)$ represents the second time derivative of the path given by the expansion in equation (8). By calculating these derivatives analytically it is possible to employ standard minimisation algorithms, such as conjugate gradient minimisation, in order to locate the set of coefficients $\{\mathbf{b}\}$ for which the Action is zero. The steps of the algorithm are summarised below:

1. Define a initial guess of the path in terms of an adjustable set of parameters $\{\mathbf{b}\}$.
2. Specify all the interactions of interest in terms of a potential energy function $\mathcal{V}(t)$.
3. Evaluate the action and its derivatives at each time step from the expressions (7) and (11) and apply conjugate gradient minimisation in order to locate the coefficient which minimises the action.

3 Results and discussion

3.1 1D harmonic potential, two particle system

The first system under investigation is a simple two bodied harmonic oscillator. This simplified experiment uses the potential function specified by equation (9) where the equilibrium separation and the spring constant are both set to 1. The mass for both particles is 10 and the initial separation between the particles is zero. In this one space, one time dimensional problem the particles have been allow to crossover. The initial choice of 20 coefficients for the path expansion were chosen to be random.

To increase the complexity of the problem, the initial path estimate was chosen from a two dimensional path search space. Figure 1 shows a plot of path given by the initial guess in dashed line and the final solution for the system shown by the solid lines. Initially, there is a larger error between the physically realisable path and the ‘estimate’. This figure shows a sample of the trajectory to highlight the periodic final path. It is clear that the algorithm is able to find a path in one dimension which is periodic and satisfies Newton’s equation of motion. Figure 2 shows progress of the algorithm in terms of the number of minimisation steps. The path rapidly converges to a symmetric path with an overall Action of zero. In the next two experiments we investigate non-linear, more complex situations.

3.2 Lennard–Jones potentials

For the following two experiments, the interaction potential is the non-bonded Lennard–Jones potential described by equation (10). The particles start in their force equilibrium positions. This is important for achieving zero Action pathways in simulations where there is initial rotation.

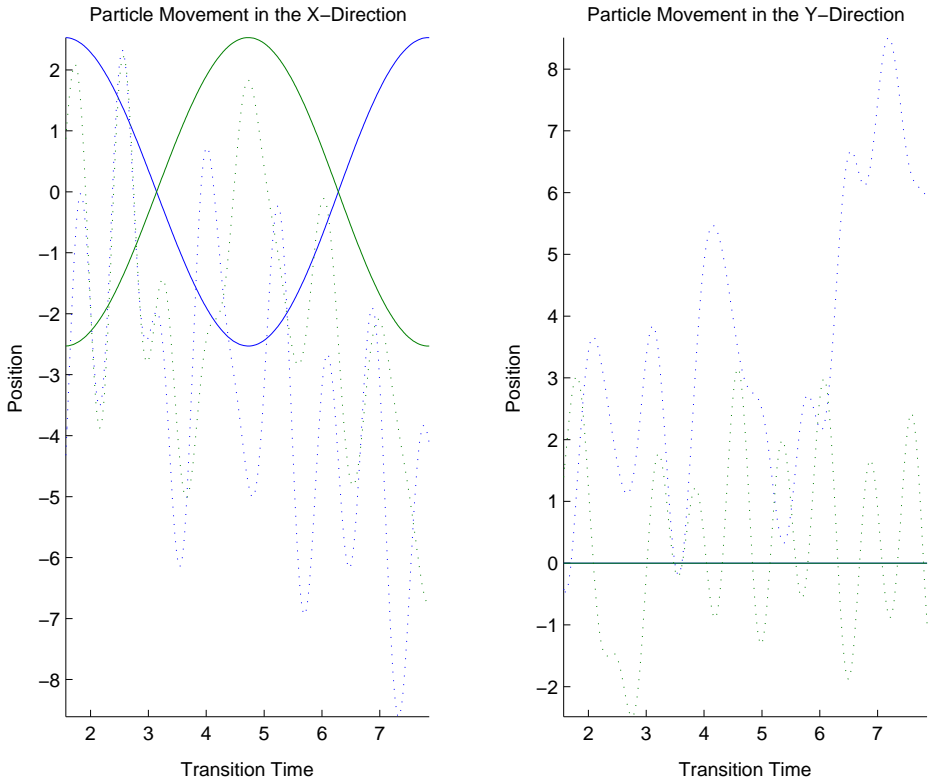


FIGURE 1: Particle path.

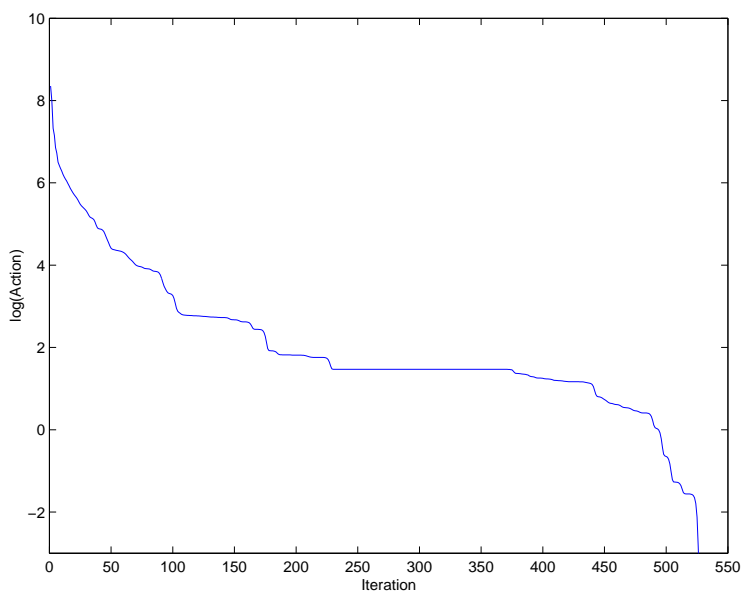


FIGURE 2: Action against minimisation steps.

Experimental Setup.

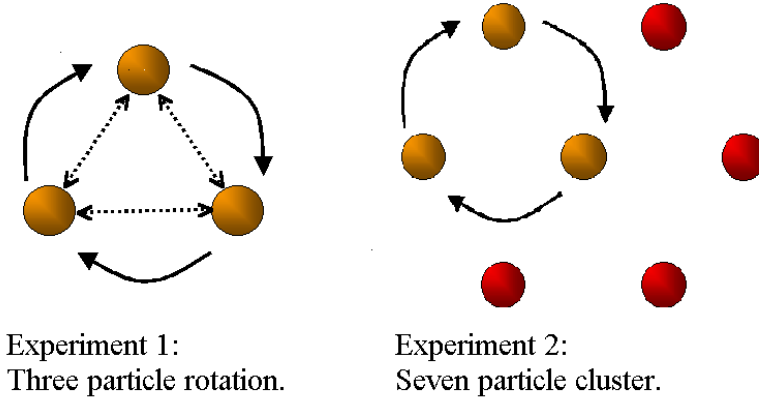


FIGURE 3: Rotation experiments using Lennard–Jones Potential.

In the both experiments, three particles rotate in a clockwise direction. The initial and final separations for all pairs of particles in each of the systems are the same. This is to ensure that the initial and final forces are the same. Figure 3 shows the experimental setup for the experiments.

3.3 Experiment 2: three particle system

Figure 3 shows the setup for experiment two. All particles are again chosen to have equal mass and the transition time τ is chosen arbitrarily. Lennard–Jones coefficients are chosen to maintain the zero initial force condition. Figure 4 shows the final path converged to for the three particle cluster. Twenty coefficients were chosen randomly and the result of the minimisation

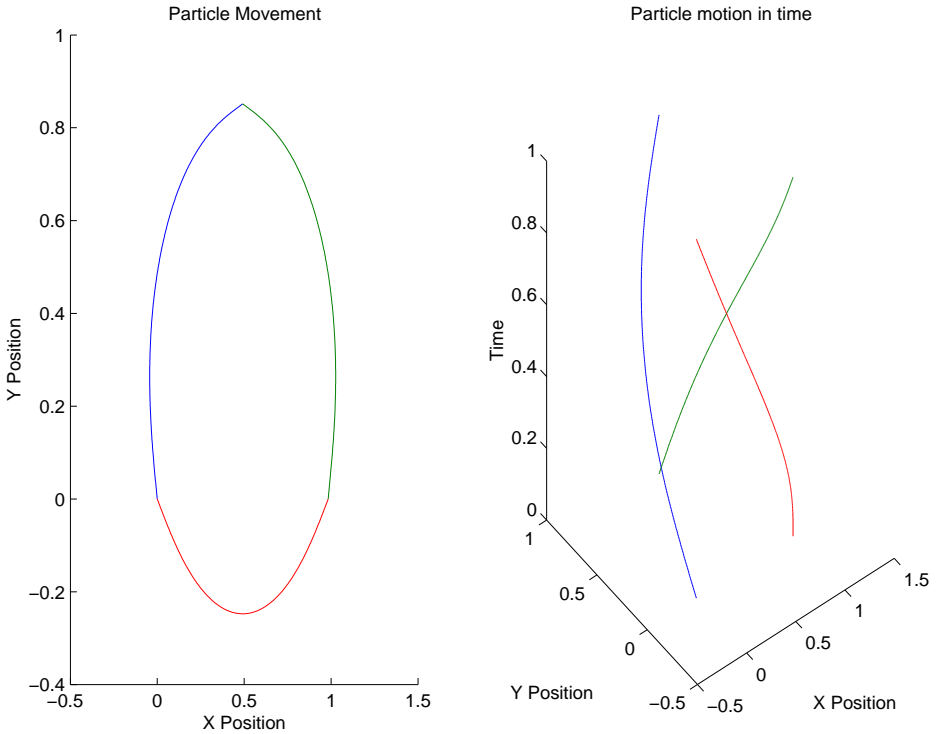


FIGURE 4: Three particle path.

is shown. The difference between the circular path and the path in Figure 4 is due to the limitations in the approximation used in the expansion. To investigate this in more detail, Figure 5 compares the force on a single particle with the difference between that particles path, and a circular path. Because of the enforced rotation, it was necessary to balance the angular momentum of the particles with the Lennard–Jones force. The rotation in this experiment requires an initial force. This initial force shows up in the action calculation yielding a non-zero final action. For the first simple example, it is possible to adjust the starting position of the particles in order to achieve a zero final action. The resulting path will be circular and the particles will

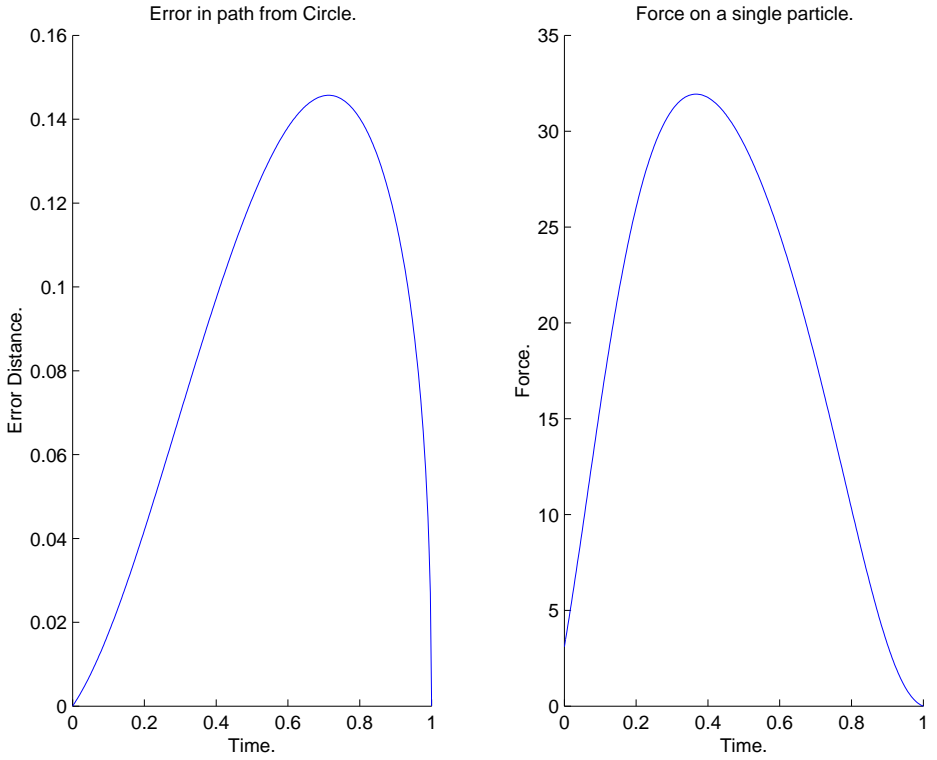


FIGURE 5: Comparison of the force on an individual particle and the error in the approximation.

stay at their equilibrium positions for the whole transition. The purpose of placing the particles at equilibrium positions was to ensure that forces at the boundaries were balanced. This has been done for comparison purposes so that the simple expansion we have used would be able to closely approximate the path.

This demonstrates the need for a more accurate expansion by showing that the inability of the expansion to accurately represent the true path is the main limitation in achieving a dynamical solution. A more detailed choice of expansion is needed to improve accuracy of the path and allow broader application of this method.

3.4 Experiment 3: seven particle system

In this experiment, the three rotating particles from experiment two are embedded in a symmetrical seven bodied cluster. Figure 3 show the seven particle cluster used in this experiment and the interaction function used is given by equation (10). As a first approximation to the path we choose to use the result achieved from the previous experiment. Figure 6 shows the converged solution. Convergence to the final solution was much slower than in the previous examples; however, by applying hierarchical optimisation, minimisation can be done in a short time. In this case, the final path did not have zero Action. There are two possibilities for this: either the result was a local minima in the Action surface; or the expansion was inappropriate for describing this type of motion.

4 Conclusions

It is clear from this initial analysis that the key limitation encountered in this approach arises from the expansion choice and not the algorithm. Be-

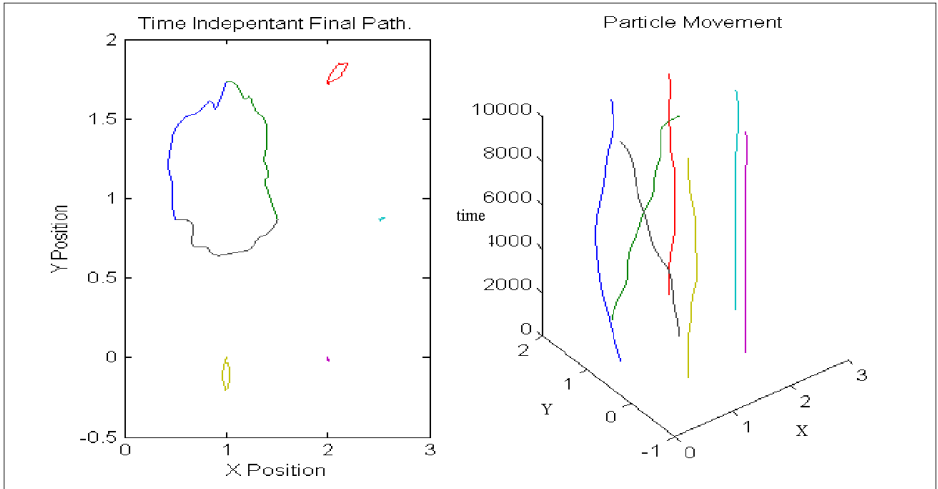


FIGURE 6: Seven particle final path.

cause the second derivative of the expansion shown in equation (8) is initially zero regardless of the physical system, it is inappropriate when applied to situations in which forces are acting on the system at the boundaries. In simple cases it was possible to adjust conditions to achieve the desired zero-Action path. Recent work of Passerone and Parrinello [7] uses an iterative process to attain dynamical trajectories from known boundary conditions. Their approach avoids the use of second derivatives by successive minimisation steps which progressively optimise the energy and minimise the Action. We choose to use the analytic second derivative to enable the implementation of fast optimisation and avoid an iterative scheme.

One of the more apparent advantage to this algorithm is that it is greatly adjustable to suit the computational limitations. In theory, the coarseness of the time sampling used in the simulation does not effect the existence of a stable trajectory. This allows the location in time of key events to discovered at a coarse grained level and then investigated at greater detail. In practice this is not the case. Insufficient sample points would potentially avoid key

events in the transition, resulting in a path which does not capture important real dynamics accurately. This results from the algorithm searching for paths whose Action are zero.

A computational advantage to this formulation is its suitability for parallel processing as the Action at different time points in the path can be computed independent of each other. Processors can be assigned individual time slices Action calculations which then take place simultaneously. Since the method uses a boundary value approach the transition will always finish in the final state. This is not a guarantee in a finite time for forward integration of the equations of motion.

These advantages make this a useful, efficient technique for solving problems where the boundary information is known.

Acknowledgments: The first author gratefully acknowledges the Australian Research Council Centre of Excellence for Mathematics and Statistics of Complex Systems and the University of Queensland School of Physical Sciences for supporting this work.

References

- [1] R. Olender and R. Elber, Calculation of classical trajectories with a very large time step: Formalism and numerical examples. *Journal of Chemical Physics* **105** (1996), 9299–9315. [C537](#)
- [2] P. Ferrara, J. Apostolakis and A. Caffisch, Targeted molecular dynamics simulations of protein unfolding. *Journal of Physical Chemistry B* **104** (2000), 4511–4518. [C536](#)

- [3] P. Ferrara, J. Apostolakis and A. Caffisch, Computer simulations of protein folding by targeted molecular dynamics. *Proteins-Structure Function and Genetics* **39** (2000), 252–260. [C536](#)
- [4] C. Dellago, P. G. Bolhuis, F. S. Csajka and D. Chandler, Transition path sampling and the calculation of rate constants. *Journal of Chemical Physics* **108** (1998), 1964–1977. [C536](#)
- [5] C. Dellago, P. G. Bolhuis and D. Chandler, Efficient transition path sampling: Application to Lennard- Jones cluster rearrangements. *Journal of Chemical Physics* **108** (1998), 9236–9245. [C536](#)
- [6] C. Dellago, P. G. Bolhuis and D. Chandler, On the calculation of reaction rate constants in the transition path ensemble. *Journal of Chemical Physics* **110** (1999), 6617–6625. [C536](#)
- [7] D. Passerone, M. Ceccarelli and M. Parrinello, A concerted variational strategy for investigating rare events. *Journal of Chemical Physics* **118** (2003), 2025–2032. [C536](#), [C548](#)
- [8] D. Passerone and M. Parrinello, Action-derived molecular dynamics in the study of rare events. *Physical Review Letters* **8710** (2001), art. no.-108302. [C536](#)
- [9] M. Parrinello, Action-derived molecular dynamics in the study of rare events. *Abstracts of Papers of the American Chemical Society* **221** (2001), 140-PHYS. [C536](#)