Sampling in Path Space - A Physical Application

Frank Pinski (Physics, Cincinnati)

Andrew Stuart (Warwick)

Patrick Malsom (Cincinnati)

Warwick Mathematics Institute

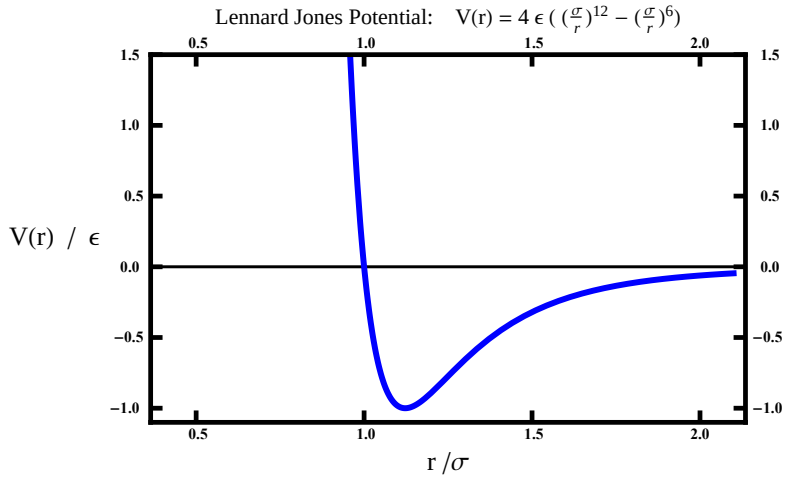
April 4, 2013

The General Problem

Consider molecules, or clusters of atoms, whose motions are governed by Brownian dynamics. The Free-Energy Landscape has many wells which are separated by barriers. Some may be large and all may shift with temperature or external field.

Such transitions are rare when the barrier is large compared to the available thermal energy. **How do we find the paths that describe the transitions to the new equilibrium state when such events are rare?** A possible solution: constrain paths to make the desired transition, sample these paths in a thermodynamically significant manner.

We use a Hybrid Monte Carlo Method in **infinite-dimensional** *path space* to sample transitions in small clusters interacting via Lennard-Jones potentials.



Lennard-Jones: 13 atoms

(Loading Video...)

Jellinek, Beck and Berry, J. Chem. Phys., 1987.

Outline

1. The General Problem
2. Sampling in Finite Dimensions
   - Monte Carlo Methods
   - Molecular Dynamics
   - Hybrid Methods
3. Brownian Dynamics and Path Space
   - Onsager-Machlup Functional
   - The Continuum Limit
   - Ito, Girsanov and the Radon-Nikodym Derivative
4. Sampling in Infinite Dimensions
   - Hybrid Monte Carlo Method in Path Space
   - Numerically Sampling the Measure
5. Results
6. Conclusion and Discussion

Sampling in Finite Dimensions

For simplicity I will frame the discussion in terms of thermodynamics.

Consider N coordinates q, the thermal distribution is given in terms of the Boltzmann factor:

$$P(q) = \frac{1}{Z} \exp\Big(- V(q)/\epsilon \Big)$$

where ϵ is the temperature of the physical system.

To understand the physical properties, one needs the expectations of measurable quantities:

$$\mathbb{E}^{\epsilon} O_p = \frac{1}{Z} \int dq \; O_p(q) \; \exp\Big(- V(q)/\epsilon \Big)$$

One then numerically samples the distribution P to calculate the thermodynamics of such a system.

Monte Carlo - Importance Sampling

In this talk, I only have time to describe the method, based on a Markov Chain:

1. Start with some initial set q_0
2. Generate new set q_1 at random
3. Accept or reject q_1 according to the Metropolis criterion

$$\text{If } \frac{P(q_1)}{P(q_0)} > \eta, \text{ then accept (otherwise reject)}$$

where η is a uniform random number between zero and one.

If a bias is introduced when generating q_1, then the accept/reject criterion must be modified to preserve the (detailed) balance. This is usually called the **Metropolis-Hasting** criterion.

Molecular Dynamics

Consider the molecules in this room. If we simply follow Newton's laws, we should be able to understand their thermodynamical properties.
This then corresponds to exploring the joint distribution:

$$\mathbb{P}(q, p) = P(q)P_G(p)$$

where P_G is a Gaussian distribution of the momenta, usually called the Maxwell Boltzmann distribution.
We then have a underlying Hamiltonian:

$$H = \frac{1}{2}\ p \cdot M^{-1} \cdot p\ +\ V(q)$$

$$\text{with} \qquad \frac{dx}{dt} = \frac{\partial H}{\partial p} \qquad \text{and} \qquad \frac{dp}{dt} = -\frac{\partial H}{\partial x}$$

Rely on the chaotic characteristics of the dynamical system to explore all available phase space.

Advantages and Disadvantages

Monte-Carlo methods are exact but may not be very efficient.

Conservation laws may be violated in Molecular-Dynamics calculations due to errors in the deterministic integration. Use methods that are better than leap-frog. (But at what cost?)

Leap-frog (velocity-Verlet) for time step h: $\{q_0, p_0\} \to \{q_1, p_1\}$

$$p_{1/2} = p_0 + \frac{h}{2} F(q_0) \qquad q_1 = q_0 + h\, M^{-1} \cdot p_{1/2} \qquad p_1 = p_{1/2} + \frac{h}{2} F(q_1)$$

This map is volume conserving and reversible (symplectic).

Combine these two methods: **Hybrid Monte Carlo**

Hybrid Monte Carlo

This then has the potential of being an exact method. Why?
Because it uses the accept/reject step to correct errors introduced by the approximate integration schemes.
The Mass matrix can be chosen to improve the sampling efficiency.
The steps $\{q_0\} \to \{q_1\}$

1. Pick the momentum $\{p_0\}$ according to the correct Gaussian distribution
2. Evolve $\{q_0, p_0\} \to \{q_1, p_1\} \to \ \ldots \ \to \{q_n, p_n\}$ via deterministic integration
3. Accept/reject according the Metropolis-Hasting criterion

$$\text{if scheme is sympletic} \qquad \frac{\mathbb{P}(q_n, p_n)}{\mathbb{P}(q_0, p_0)} > \eta$$

4. Repeat

Three ingredients:
(1) Choice of M (2) the splitting and (3) type of integrator.

Brownian Dynamics

Overdamped limit of the Langevin equation

$$dq = F(q)\,dt + \sqrt{2\,\epsilon}\;\,dW$$

Figure: Norbert Wiener - 1923

$$q(t+h) - q(t) = \int_t^{t+h} F(q)\,dt + \sqrt{h\,2\,\epsilon}\,\left(W_{t+h} - W_t\right)$$

Definition A standard (one-dimensional) Wiener process (also called Brownian motion) is a stochastic process $\{W_t\}_{t\geq 0}$ indexed by nonnegative real numbers t with the following properties:
(1) $W_0 = 0$.
(2) With probability 1, the function $t \to W_t$ is continuous in t.
(3) The process $\{W_t\}_{t\geq 0}$ has stationary, independent increments.
(4) The increment $W_{t+h} - W_h$ has the $NORMAL(0, 1)$ distribution.

Onsager-Machlup Functional

$$q(t+h) - q(t) = \int_t^{t+h} F(q)\, dt \; + \; \sqrt{h\, 2\, \epsilon}\, \Big(W_{t+n} - W_t \Big)$$

Computationally (Ito Calculus, Euler-Maruyama method)

$$q(t+h) - q(t) = h\, F\Big(q(t)\Big) + \sqrt{h\, 2\, \epsilon}\, \xi \qquad P_G(\xi) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}\, \xi^2)$$

String the steps together to make a path: $\{q_0, q_1, q_2, ..., q_n\}$
Onsager-Machlup functional gives the probability of the path:

$$P_{path} = \exp\left(-\frac{1}{2\, \epsilon} \sum \frac{h}{2} \Big| \frac{\Delta q}{h} - F \Big|^2 \right)$$

Quadratic Variation: sum over a number of steps that corresponds to a length of time T is given by $\qquad QV_\alpha = \sum \Big| \Delta q_\alpha \Big|^2 = 2\, T\, \epsilon.$

How does one proceed to continuum limit? What are the corresponding expressions in infinite dimensions?

Continuum Limit

To examine transitions, we start the path in one well, $q(0) = q^-$, and then constrain the end of the path, $q(T) = q^+$

$$P_{path} \propto \exp\Big(-I\Big) \qquad \text{with} \qquad I = \frac{1}{2\,\epsilon} \sum \frac{h}{2} \Big| \frac{\Delta q}{h} - F \Big|^2$$

Then we look at the limit as $h \to 0$. Informally we can write

$$I = \frac{1}{2\,\epsilon} \int_0^T dt \left(\frac{1}{2} \Big| \frac{dx}{dt} \Big|^2 + G(x_t) \right) \qquad \text{with} \quad G = \frac{1}{2} \Big| F \Big|^2 - \epsilon \triangle V$$

Why is the adjective "informally" used? Brownian paths are almost nowhere differentiable. The meaning of I has to be put on firmer ground.

Infinite Dimensions

$$q(0) = x^- \qquad dq = F(q)\,dt + \sqrt{2\,\epsilon}\;dW$$

Brownian Measure $\pi_0(q)$

Radon-Nikodym derivative, using the Girsanov equation

$$\frac{d\pi}{d\pi_0} = \exp\left(\frac{1}{2\,\epsilon}\left(\int_0^T dt\; \frac{1}{2}\left|F\right|^2 \;\; - \;\; \left\langle F(q), dq \right\rangle \right) \right)$$

The $< \ldots >$ denote inner product. In particular, need to use the Ito formula to evaluate last term which is a stochastic integral

$$\frac{d\pi}{d\pi_0} = \exp\left(-\frac{1}{2\,\epsilon}\int_0^T dt\; \left(\frac{1}{2}\left|F\right|^2 - \epsilon \triangle V \right) \right) = \exp\left(-\frac{1}{2\,\epsilon}\int_0^T dt\; G \right)$$

Now constrain the path to end at a fixed configuration: $q(T) = q^+$

The double-ended path probability is given by a proportionality:

$$\frac{d\pi}{d\pi_0} \propto \exp\left(-\frac{\Phi}{2\,\epsilon}\right) \qquad \text{with} \qquad q(0) = q^- \quad q(T) = q^+$$

Sampling in Path Space

First look at the finite dimensional version. Remember, that informally the path probability can be written as

$$\pi(q) \propto \exp\left(-\frac{1}{2\,\epsilon} \int_0^T dt \, \left(\frac{1}{2} \left| \frac{dq}{dt} \right|^2 + G \right) \right)$$

Approximate using finite sums and define $L = d^2/dt^2$

$$\log \pi(q) \approx \mathbb{C} - \frac{\Delta t}{2\,\epsilon} \sum_n \left(-\frac{1}{2}\, q_n\, L q_n + G(q_n) \right)$$

Define an effective Hamiltonian, adding auxiliary variables p,

$$H_{eff} = -\frac{1}{2} \Big\langle q, Lq \Big\rangle + \frac{1}{2} \Big\langle p, M^{-1} p \Big\rangle + \Phi$$

The effect of the interactions are contained within $\Phi = \langle 1,\, G \rangle$.

Want the first two terms to be a mixture of Brownian Bridges. This forces $M = -L$.

$$H_{eff} = -\frac{1}{2} \Big\langle q, Lq \Big\rangle - \frac{1}{2} \Big\langle v, Lv \Big\rangle + \Phi$$

Equations of Motion

Effective Hamiltonian (with $L = d^2/dt^2$)

$$H = -\frac{1}{2}\Big\langle q, Lq \Big\rangle - \frac{1}{2}\Big\langle v, Lv \Big\rangle + \Phi$$

Evolution of path (Hamilton's equations)

$$\frac{\partial q}{\partial \tau} = v \qquad \text{and} \qquad \frac{\partial v}{\partial \tau} = -q + L^{-1}\nabla G$$

Splitting and Integration scheme

$$w_0 = v_0 + \frac{h}{2} L^{-1}\nabla G_0$$

$$\begin{pmatrix} q_1 \\ w_1 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} q_0 \\ w_0 \end{pmatrix}$$

$$v_1 = w_1 + \frac{h}{2} L^{-1}\nabla G_1$$

Comments

The grid along the path must be fine enough as to resolve the Quadratic Variation, that is given by $QV = \sum \Delta q^2 = 2\,T\,\epsilon$.

Now take another look at

$$H_{eff} = -\frac{1}{2}\left\langle q, Lq \right\rangle - \frac{1}{2}\left\langle v, Lv \right\rangle + \Phi$$

Note the first two terms are proportional to the quadratic variation.

$$-\frac{1}{2}\left\langle q, Lq \right\rangle = \frac{\epsilon\,T}{\Delta t} \qquad\qquad \rightarrow \infty$$

Also remember that during the deterministic integration, energy sloshes between kinetic and potential energy.

To preserve the quadratic variation, the last term in H_{eff} must be much smaller than the first two terms.

$$\frac{1}{\Delta t} >> \frac{1}{\epsilon}\left|\overline{G}\right| \approx \frac{1}{2}\overline{\triangle V}$$

More Comments

Now take yet another look at

$$H_{eff} = -\frac{1}{2}\left\langle q, Lq \right\rangle - \frac{1}{2}\left\langle v, Lv \right\rangle + \Phi$$

Remember the integration employs a particular form of "splitting." Because the "full step" is exact:

(1) the quadratic variation is preserved.
(2) the curse of large dimensions is negated.

A final comment about the conservation of the effective energy. The effective energy is large and is dominated by the quadratic variation. The interactions, Φ, contribute little. One should not calculate the drift of the effective energy by direct subtraction.

We have derived an explicit expression for the energy drift. This expression is evaluated after each step of the deterministic integration. The expression contains terms that telescope over the length of this integration. Thus the error in the energy oscillates over the course of the deterministic integration.

Lennard-Jones: building the 13-atom cluster

(Loading Video...)

(Loading Video...)

(Loading Video...)

(Loading Video...)

Lennard-Jones Clusters: $\epsilon = 0.13$

For the 13-atom cluster, we considered the transition from its ground state to a conformation where one atom sits on the surface, and a "dimple" exists on the opposite side.
For the 14-atom cluster, we considered the process that starts with the "extra" atom on one side of the cluster and ends with the "extra" atom on the other side.

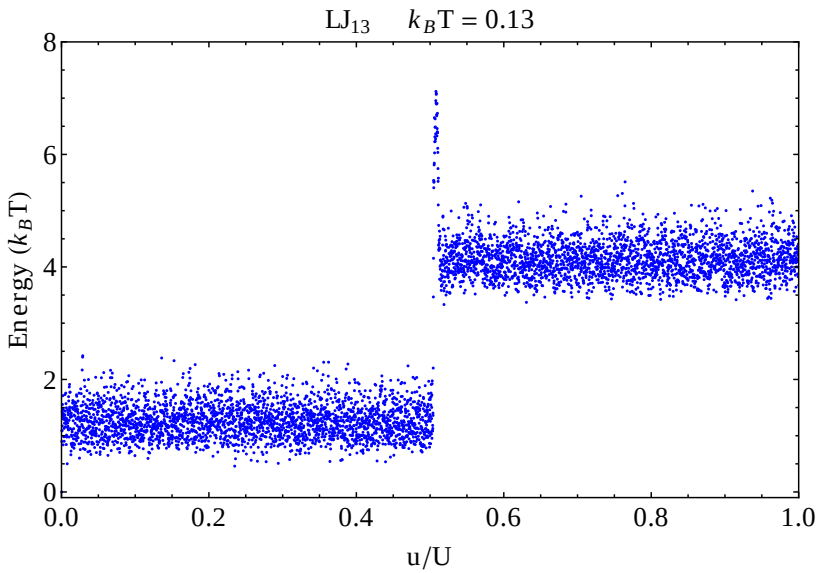


Figure: LJ_{13}: Energy along the path.

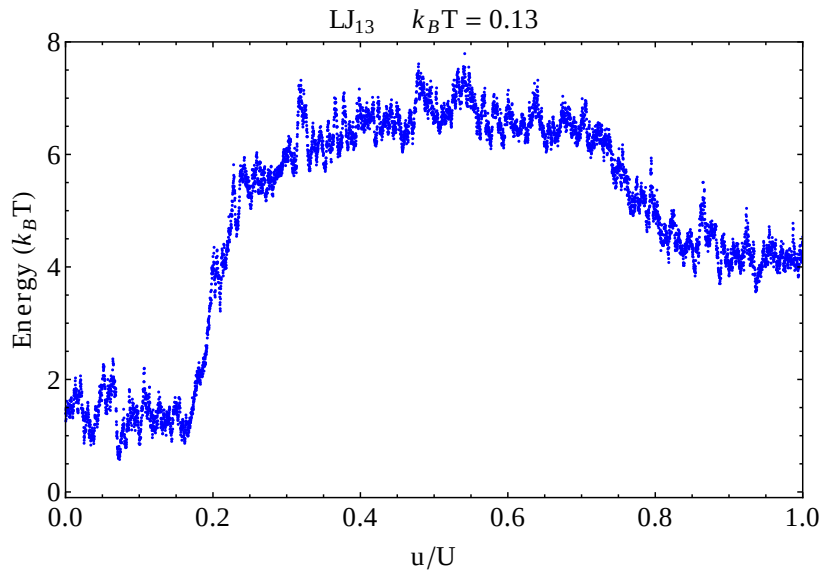


Figure: LJ_{13}: Energy along the path.

Lennard-Jones cluster: inspecting the transition

FRONT VIEW

(Loading Video...)

BACK VIEW

(Loading Video...)

HMC - Some Numerical Aspects

The effective Hamiltonian for N_p in N_d dimensions can be written as $H = \sum_{i\,\alpha}\left(-\frac{1}{2}\langle q_{i\,\alpha},\ Lq_{i\,\alpha}\rangle - \frac{1}{2}\langle v_{i\,\alpha}\,, Lv_{i\,\alpha}\rangle\right) + \langle 1,\ G(q)\,\rangle$. On a finite grid, using the quadratic variation of the conditioned Brownian Bridges, both the first and second terms become $N_u \times \epsilon$, where ϵ is the temperature and N_u is the number of divisions along the path. To keep the quadratic variation of the path nearly constant, the last term must be small compared to the other two. If one uses $\overline{G}$ as an average value of G, then $\Delta t = \frac{U}{N_u} >> \frac{\overline{G}}{N_p N_d \epsilon}$. The value of $\overline{G}$ can be approximated by its equilibrium average. Thus for these clusters we take $\Delta t \approx 0.0001$.

For the time step h used in the deterministic integration, we have chosen $h = 0.00024$ which gives us an acceptance rate of over 90%. We then integrate over N_{MD} steps. We chose $N_{MD} \approx 15000$ to give $N_{MD}h$ a value somewhere between 1 and 3.

Messages to take Home

- Explained how to implement the Hybrid Monte Carlo Method and how it can be used to sample Path Space
- Studied small Lennard-Jones clusters
- Even for such simple systems: HMC in path space provides physical insights

And I should mention some who have shown patience and have taught me a lot.

- J. Voss
- F. Theil
- A. Beskos
- J.-M. Sanz-Serna
- H. Weber
- G. Simpson
- M. Ottobre
- N.S. Pillai

Lennard-Jones 14 Cluster: $\epsilon = 0.13$

For the 13-atom cluster, we considered the transition from its ground state to a conformation where one atom sits on the surface, and a "dimple" exists on the opposite side.
For the 14-atom cluster, we considered the process that starts with the "extra" atom on one side of the cluster and ends with the "extra" atom on the other side.

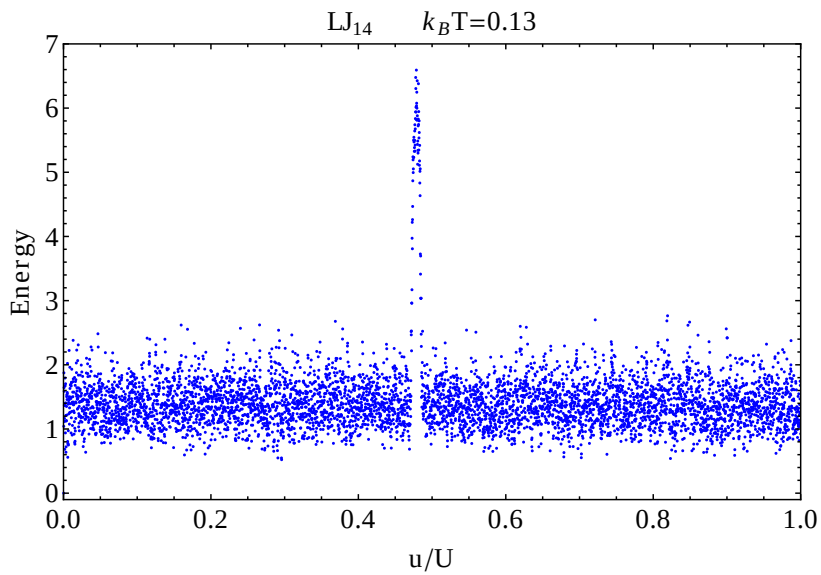


Figure: LJ_{14}: Energy along the path.

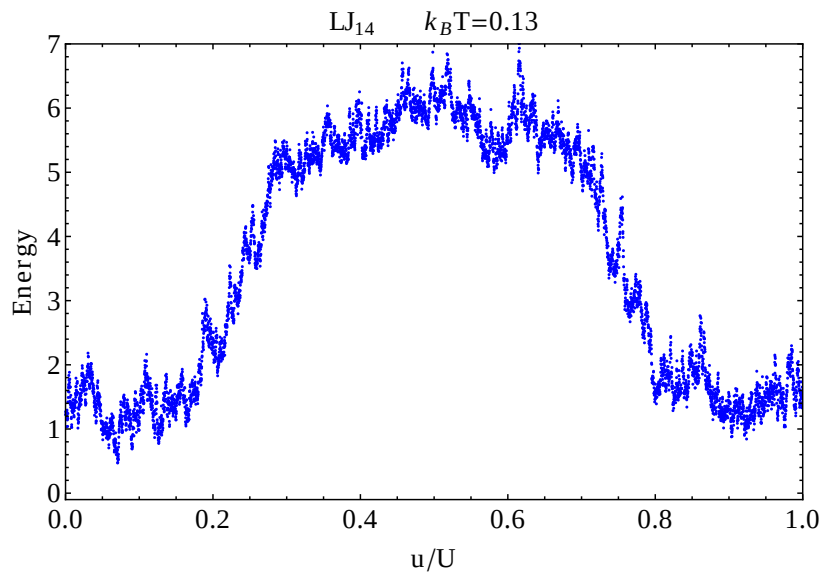


Figure: LJ_{14}: Energy along the path.

Lennard-Jones 14 cluster: inspecting the transition

FRONT VIEW

(Loading Video...)

BACK VIEW

(Loading Video...)

Lennard-Jones cluster: inspecting the transition

FRONT VIEW

(Loading Video...)

BACK VIEW

(Loading Video...)