A fresh look at the Onsager Machlup functional

Patrick Malsom

Department of Physics, University of Cincinnati

April 16, 2014

Outline

1. Introduction to Stochastic Sampling
   - Overdamped (Brownian) Dynamics
   - Metropolis Hastings Algorithm
2. Onsager Machlup Functional
   - A Different Way to Sample: Path Space
   - Results
3. A Different Route to the OM Functional
   - Hybrid Monte Carlo
   - Different Perspective
4. Discussion

Starting point

Let us begin this discussion with a damped equation of motion

$$\frac{dp}{d\tau} = F - \gamma p + \text{Noise}$$

- Brownian regime: $\frac{dp}{d\tau} \to 0$
- Scale the time: $\tau \to \gamma m t$

$$\frac{dx}{dt} = F(x) + \text{Noise}$$

Research Objective

Find an *efficient* numerical method of sampling transitions in this Brownian regime.

Sampling in the overdamped regime

Writing this stochastic differential equation (SDE)

$$dx = F(x)dt + \sqrt{2\varepsilon}\, dW = \varepsilon \nabla \log P_B dt + \sqrt{2\varepsilon}\, dW$$

- ε is temperature
- $F(x)$ is the force $\qquad F(x) = -\nabla U(x)$
- W represents the Wiener process (White noise)

Classical dynamics $\Rightarrow$ Boltzmann:

$$P_B \propto \exp\left(\frac{-U(x)}{\varepsilon}\right)$$

UNIVERSITY OF Cincinnati

Discrete sampling

$$dx = F(x)dt + \sqrt{2\varepsilon}\, dW$$

$$x_1 - x_0 = F_0\, \Delta t + \sqrt{2\varepsilon \Delta t}\, \xi_0$$

- ξ is a Gaussian random number $(\mu = 0\ , \quad \sigma = 1)$
- If ε is much smaller than the barrier height, transition events become exponentially rare.
- The high-frequency modes are dominated by noise.

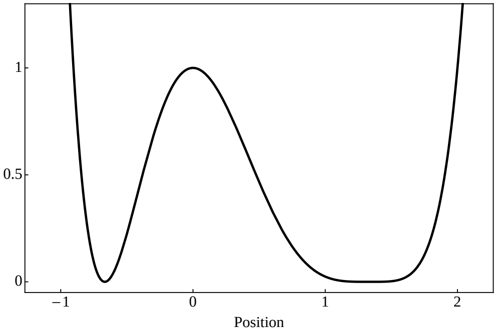
Quadratic variation sum rule:

$$\sum (x_{i+1} - x_i)^2 = 2\varepsilon N \Delta t = 2\varepsilon\ T$$

- Discrete time step size is an important consideration (time step size too large will lead to incorrect sampling)

A small detour: 1D Potential

I will use a simple one dimensional potential to illustrate my results.



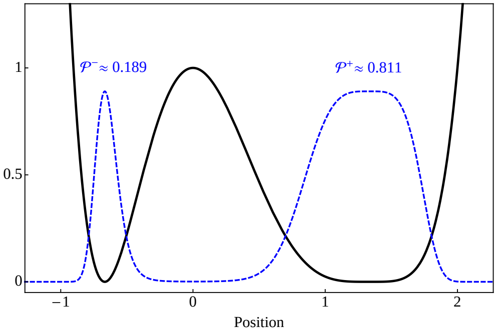
$$U(x) = \frac{(4x-3)^4\ (2x+3)^2}{1024}$$

- Minima: $-\frac{2}{3}$ and $\frac{4}{3}$
- Barrier Height: $U_{\text{bar}} = 1$
- $\varepsilon = 0.15$

UNIVERSITY OF Cincinnati

A small detour: 1D Potential

I will use a simple one dimensional potential to illustrate my results.

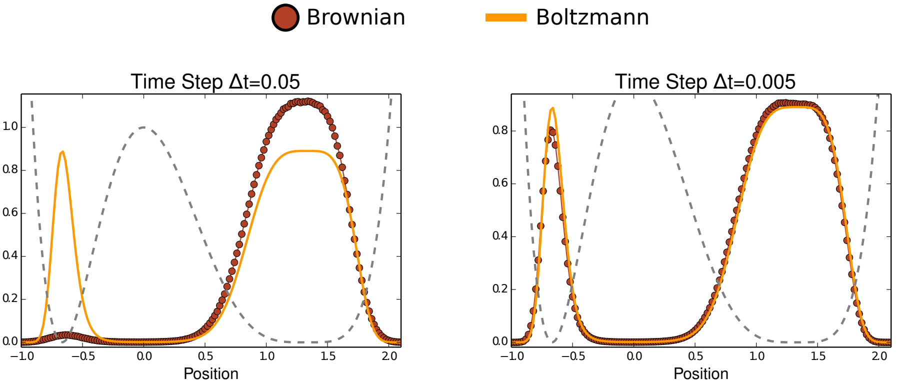


$$U(x) = \frac{(4x-3)^4\ (2x+3)^2}{1024}$$

- Minima: $-\frac{2}{3}$ and $\frac{4}{3}$
- Barrier Height: $U_{\text{bar}} = 1$
- $\varepsilon = 0.15$

Boltzmann probability: entropy drives the particle to spend more time in the broad well.

Difficulties with this simple discrete sampling



Metropolis-Hastings on a Markov chain

Markov chain $\rightarrow$ **memoryless**

Metropolis-Hastings criterion:

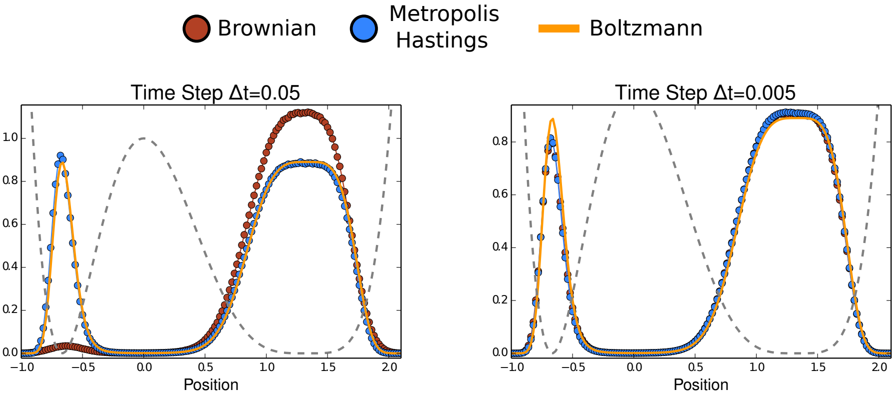
$$\frac{\pi(B \to A)}{\pi(A \to B)} \ \exp\left(\frac{-\Delta U(x)}{\varepsilon}\right) > \eta$$

Detailed balance drives the solution to a stationary probability

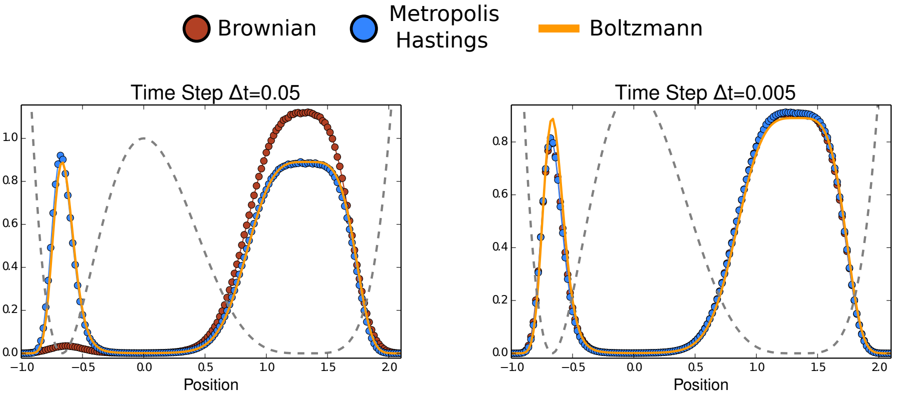
$$\mathcal{P}(A)\pi(A \to B) = \mathcal{P}(B)\pi(B \to A)$$

- $P(A)$: equilibrium probability at A
- $\pi(A \to B)$: given A, probabilty of going to B

Adding in the Metropolis Hastings criterion



Adding in the Metropolis Hastings criterion



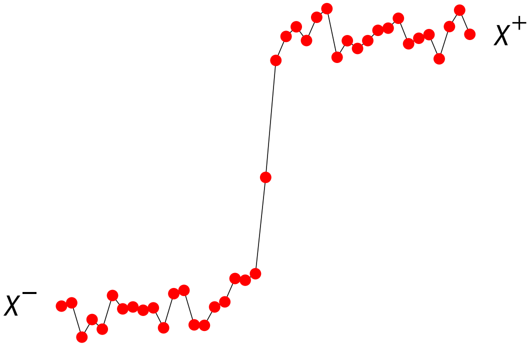
Very few transitions are seen at $\varepsilon << U_{\mathbf{bar}}$

$\sim$ 15000 steps for one transtion ($\Delta t = 0.005$)

Only calculating the interesting parts

The focus of this work is to sample transitions (*very* rare at low temperatures). The majority of the sampling resides in the wells.

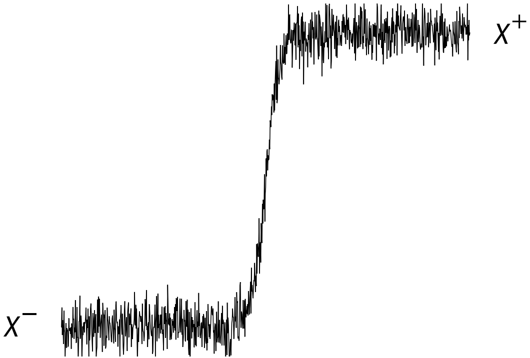
Can we develop a method that considers the transition path as an object, and sample these paths directly?



Only calculating the interesting parts

The focus of this work is to sample transitions (*very* rare at low temperatures). The majority of the sampling resides in the wells.

Can we develop a method that considers the transition path as an object, and sample these paths directly?



The probability of a path

Use the noise history to express the path probability.

Onsager-Machlup functional

$$\mathcal{P}_{\mathrm{OM}} \propto \prod_i \exp\left(-\frac{\xi_i^2}{2}\right)$$

Eliminate the random variables and express the path probability in terms of the path variables themselves.

$$x_1 - x_0 - F(x_0)\Delta t = \sqrt{2\varepsilon\,\Delta t}\;\xi_0$$

Onsager Machlup

The path probability:

$$\mathcal{P}_{\mathrm{OM}} \propto \exp\left(-\frac{\Delta t}{2\varepsilon}\sum_i\left[\frac{1}{2}\left(\frac{\Delta x}{\Delta t}\right)^2 + G\right]\right)$$

The '*path potential*':

$$G = \frac{1}{2}|F|^2 - \varepsilon\nabla^2 U$$

Onsager Machlup as an action

The OM functional can be thought of as an action in the classical sense. Sampling this action provides a collection of paths weighted by the correct thermodynamic factor.

$$S = \frac{1}{2\varepsilon} \int_0^T dt \left[\frac{1}{2} \left(\frac{dx}{dt} \right)^2 + G \right]$$

- $T = N \, \Delta t$ is the length of the path
- Short paths may lead to unphysical results
- Long paths should reproduce the Boltzmann distribution
- The integral is undefined (infinite)

We impose *boundary conditions* at each end.

The language of paths

$$P_{\mathrm{OM}} \propto \exp\left(-\frac{1}{2\varepsilon}\sum \frac{1}{2} v\, M\, v\right)\, \exp\left(-\frac{1}{2\varepsilon}\sum\left[\frac{1}{2}\left(\frac{\Delta x}{\Delta t}\right)^2 + G\right]\right)$$

$$H_{\mathrm{eff}} = \frac{1}{2}\langle v, L\ v\rangle + \frac{1}{2}\langle x, L\ x\rangle + \langle 1, G\rangle$$

- $L = -\frac{d^2}{dt^2}$ (positive definite)
- v are *auxiliary* variables
- the 'mass matrix' is chosen to be L.
- $\langle \cdots \rangle$ are inner products

Both positions and auxiliary variables are Brownian bridges. In the absence of interaction we are simply mixing these bridges.

The effective Hamiltonian

$$H_{\text{eff}} = \frac{1}{2}\langle v, L\ v\rangle + \frac{1}{2}\langle x, L\ x\rangle + \langle 1, G\rangle$$

The equations of motion

$$\frac{\partial x}{\partial \tau} = \frac{\partial H_{\text{eff}}}{\partial v} = Lv$$

$$\frac{\partial v}{\partial \tau} = -\frac{\partial H_{\text{eff}}}{\partial x} = -Lx - \nabla G$$

Combining these equations yields the (partial) differential equation of motion

$$\frac{\partial^2 x}{\partial \tau^2} = -x - L^{-1}\nabla G(x)$$

Following Hamilton's equations

$$\frac{\partial^2 x}{\partial \tau^2} = -x - L^{-1}\nabla G(x)$$

- velocity at half step: $w_i = v_i - \frac{h}{2}L^{-1}\nabla G_i$
- position at full step:

$$\begin{pmatrix} x_{i+1} \\ w_{i+1} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_i \\ w_i \end{pmatrix}$$

- velocity at full step: $\quad v_{i+1} = w_{i+1} - \frac{h}{2}L^{-1}\nabla G_{i+1}$

This scheme handles the high frequency modes exactly.

When implementing this scheme numerically: $\theta \to h$

Accumulation of error from the integration

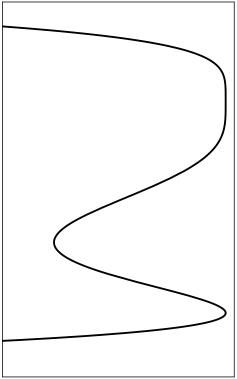
Directly calculate the Molecular Dynamics (MD) integration errors to avoid the subtration of large numbers (the *action* is possibly infinite in the continuum limit). This error will accumulate as MD steps are generated.

$$
\begin{aligned}
\Delta H_{\text{eff}} = \; & \langle 1, G_{i+1} \rangle - \langle 1, G_i \rangle \\
& + \frac{h^2}{8} \Big(\langle \nabla G_{i+1}, L^{-1} \nabla G_{i+1} \rangle - \langle \nabla G_i, L^{-1} \nabla G_i \rangle \Big) \\
& - \frac{h}{2 \sin \theta} \Big(\langle \nabla G_{i+1}, x_{i+1} - x_i \rangle - \langle \nabla G_i, x_i - x_{i+1} \rangle \Big)
\end{aligned}
$$

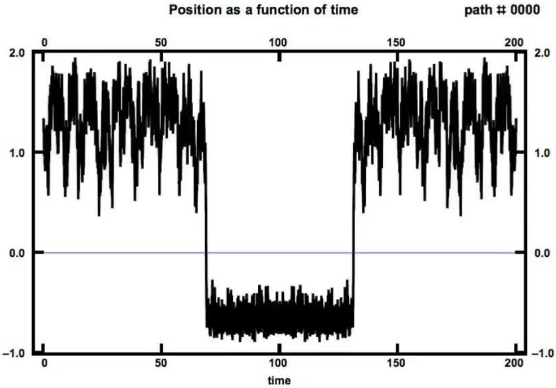
Note that all of the second term and part of the third term telescopes which can lead to an oscillation in ΔH_{eff} over many MD steps.

Results

Lets use this machinery to sample the double ended path



Potential



UNIVERSITY OF Cincinnati

NEW PHYSICS!...

Sampling of this measure is <u>not</u> consistent with Boltzmann statistics!

Clearly such paths cannot represent physical transitions. For very long paths, we expect the positions will be visited in a manner compatible with Boltzmann.

Onsager and Machlup published their work 61 years ago.
Over 35 years ago, Graham pushed the idea of using the OM functional as an action.
Many people have used this formalism, none have identified this problem.

What is wrong?

Restart this analysis with a well understood algorithm

Hybrid Monte-Carlo (HMC) sampling of configurations.

- Choose velocity: $v_0 = \sqrt{\varepsilon}\,\xi_0$ (Markov chain)
- Leap-Frog integrator (symplectic method)

$$x_1 = x_0 + h v_0 + \frac{h^2}{2} F(x_0) \qquad v_1 = v_0 + h\left(\frac{F(x_1) + F(x_0)}{2}\right)$$

- SDE (remember)

$$x_1 = x_0 + \sqrt{2\varepsilon\,\Delta t}\,\xi_0 + \Delta t\, F(x_0)$$

- Identify $h^2 = 2\Delta t$

$$\delta e = \Delta KE + U(x_1) - U(x_0)$$

Equivalence with the Onsager Machlup functional

Manage errors with Metropolis-Hastings

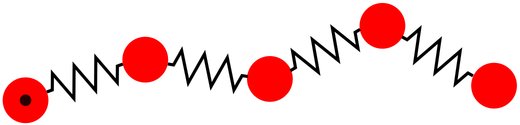
$$\exp\left(\frac{-\Delta KE - \Delta U(x)}{\varepsilon}\right) > \eta$$

$$\mathcal{P}_{\text{HMC}} \propto \exp\left(-\frac{\Delta t}{2\varepsilon}\sum\left[\frac{1}{2}\left(\frac{\Delta x}{\Delta t}\right)^2 + \frac{1}{2}|F|^2 - \varepsilon\nabla^2 V - \left|\frac{\delta e}{\Delta t}\right|\right]\right)$$

When δe is small, this new functional is equivalent to the OM functional.

The size of Δt is small enough that the quadratic variation sum rule is satisfied. In this regime, δe is also small and thus is of little consequence.

The single ended path



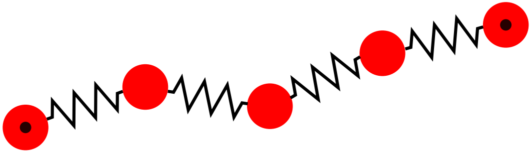
$$x^{-} \to x_0 \to x_1 \to x_2 \to x_3 \cdots$$

step x_{i+1} needs knowledge of step x_i only **(Markov chain)**

$$\mathbb{E}\Big[\left(x_{i+1} - x_i\right) \left(x_i - x_{i-1}\right)\Big] = 0$$

The single ended path moving 'forward in time' samples the Boltzmann distribution

The double ended path



$$x^{-} \rightarrow x_0 \rightarrow x_1 \rightarrow x_2 \rightarrow x_3 \cdots x_{N-1} \rightarrow x^{+}$$

This constraint adds something important. x_{i+1} depends on the constraints as well as x_i

$$\mathbb{E}\left[\left(x_{i+1} - x_i\right)\left(x_i - x_{i-1}\right)\right] \neq 0$$

The Markov property of the chain is lost and all bets are off.

University of Cincinnati

Concluding remarks

- Overview of a simple stochastic simulation
  - Generates the correct distibution
  - Inefficient when trying to sample transitions
- Shown what the OM functional is and how it can be used
  - Ensemble of transitition paths
  - Leads to broken physics!
- Recast the diffusion process in terms of HMC
  - Illuminates what went wrong
  - 'Simple' method that is well understood
  - Generates (almost) the same probability
  - Correlation destroys the Markov chain

Lesson to take home

Using the Onsager Machlup functional as an effective action to sample **double ended** paths leads to unphysical sampling.

University of Cincinnati