Fluctuations in a Thermodynamic System: The Limitations of the Onsager-Machlup (OM) Functional

Frank J. Pinski *(co-work with Patrick Malsom)*

Department of Physics, University of Cincinnati

May 26, 2014

The General Problem

Consider molecules, or clusters of atoms: The Free-Energy Landscape has many wells or basins which are separated by energy barriers.

Some of the barriers may be large, some may be small, and some may shift with temperature or external field.

Such transitions will be 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 and then sample these paths in a thermodynamic significant manner.

The most common way of doing this: using the Onsager-Machlup (OM) functional as a "Thermodynamic Action" which defines not only the most probable path but includes fluctuations.

Outline

In this talk I will look at the OM functional from the traditional point of view and explain some of the sophisticated sampling methods used to obtain an ensemble of paths.

A new perspective will be used to explore the limitations of this approach. And I will end with an unexpected result.

Starting Point - Brownian Dynamics

Sample Boltmann Distribution: $P_B \propto \exp(-U/\epsilon)$

$$dx = F\,dt + \sqrt{2\,\epsilon}\,dW_t = \epsilon \nabla \log P_B\,dt + \sqrt{2\,\epsilon}\,dW_t$$

x is the position of the particle
F is the force: $F = -\nabla U$
ϵ is the temperature
t is the time **along the path**
dW_t is the standard Wiener Process (White Noise)

If a large energy barrier exists, the transition is a **rare event**.

Thrust of this work: find an efficient way of sampling the transition paths themselves in a thermodynamically significant manner.

Discrete time step Δt

$$x_{n+1} = x_n + F(x_n)\,\Delta t \, + \sqrt{2\epsilon\,\Delta t}\;\xi_n$$

x_n is the position of the particle at time $t = n\,\Delta t$
$F(x_n)$ is the force
ϵ is the temperature
t_n is the time **along the path**
N is the number of steps in the process
$T \;=\; N\,\Delta t$ is the total time of the process
ξ_n is an Gaussian-distributed random number (mean 0; variance 1)

$$P_G(\xi)\; d\xi = \frac{1}{\sqrt{2\,\pi}} \exp\left(-\frac{1}{2}\xi^2\right)\; d\xi$$

Quadratic Variation

$$\sum \Delta x^2 \;=\; 2\,\epsilon\,N\,\Delta t = \;2\,\epsilon\;T$$

Definition of Path

$$x_{n+1} = x_n + F(x_n)\,\Delta t + \sqrt{2\epsilon\,\Delta t}\;\xi_n$$

Iterate to form

$\{x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7, \dots x_N\}$

Such a sequence of positions is a **path**.

Onsager-Machlup Functional

Replace the noise history:

$$\xi_n = \Big(x_{n+1} - x_n - F(x_n)\,\Delta t\Big)/\sqrt{2\,\epsilon\,\Delta t}$$

Path Probability

$$\log P_{path} = -I = -\frac{\Delta t}{4\epsilon}\sum\left|\frac{x_{n+1}-x_n}{\Delta t} - F(x_n)\right|^2$$

Path Probability

$$I = -\log P_{path} = \frac{\Delta t}{4\epsilon} \sum \left| \frac{x_{n+1}-x_n}{\Delta t} - F(x_n) \right|^2$$

$$I = \frac{\Delta t}{2\epsilon} \sum \left\{ \frac{1}{2} \left| \frac{x_{n+1}-x_n}{\Delta t} \right|^2 + \frac{1}{2} \left| F(x_n) \right|^2 - F(x_n) \cdot \left(\frac{x_{n+1}-x_n}{\Delta t} \right) \right\}$$

$$I = \frac{\Delta t}{2\epsilon} \sum \left\{ \frac{1}{2} \left| \frac{x_{n+1}-x_n}{\Delta t} \right|^2 + \frac{1}{2} \left| F(x_n) \right|^2 + \frac{1}{2} \frac{F(x_{n+1})-F(x_n)}{x_{n+1}-x_n} \cdot \frac{(x_{n+1}-x_n)^2}{\Delta t} - \frac{F(x_{n+1})+F(x_n)}{2} \cdot \left(\frac{x_{n+1}-x_n}{\Delta t} \right) \right\}$$

Measure relative to Brownian Bridge Measure

$$I - I_0 = \frac{U(x_N)-U(x_0)}{2\,\epsilon} + \frac{1}{2\epsilon} \int_0^T dt \; G(x_t) \qquad G = \frac{1}{2}|F|^2 - \epsilon \nabla^2 U$$

Path Probability

$$I = -\log P_{path} = \frac{\Delta t}{4\epsilon} \sum \left| \frac{x_{n+1} - x_n}{\Delta t} - F(x_n) \right|^2 \qquad G = \frac{1}{2}|F|^2 - \epsilon \nabla^2 U$$

$$I = \frac{U(x_N) - U(x_0)}{2\,\epsilon} + \frac{\Delta t}{2\epsilon} \sum \left\{ \frac{1}{2} \left| \frac{x_{n+1} - x_n}{\Delta t} \right|^2 + \quad G(x_n) \right\}$$

If we fix $x_0 = x^-$ and $x_N = x^+$ the first term is a constant.

Most Probable Path (MPP)

Look at the minimum of I. Look for heter0clinic orbits to

$$H = \frac{1}{2}\, p^2 - G(x)$$

Solution is smooth; it is not a path itself.

Should look at the solution as the center of a ball that contains the most likely paths.

Monte Carlo Sampling

We want to sample paths from the measure,

$$\pi_{path} \propto \exp\left(-\frac{\mathbb{I}}{2\,\epsilon}\right) \qquad \text{and} \qquad \mathbb{I} = \frac{1}{2}\langle x,\, L\,x\rangle + \langle 1,\, G(x)\,\rangle$$

(math notation: $\langle \ldots \rangle$ denotes an inner product.)

L is a positive definite operator $\qquad L = -d^2/dt^2$

Augment $\mathbb{I}$ to include "Kinetic Energy" thereby forming H_{eff}:

$$H_{eff} = \frac{1}{2}\langle p\, M^{-1} p\rangle + \frac{1}{2}\langle x,\, L\,x\rangle + \langle 1,\, G(x)\,\rangle$$

where M is the mass matrix.

The auxiliary variables p are conjugate to the path variables x. This term does not alter the stationary distribution of paths.

1. Choice of mass Matrix M. $\qquad M = L$.
2. Pick p from its known distribution.
3. Use a splitting to form an integrator that is reversible and volume conserving (Symplectic).
4. Avoid the subtraction of large numbers when implementing the Metropolis-Hasting step.

Deterministic Integrator

Consider the second order equation and convert it to

$$v = \frac{\partial x}{\partial \tau} \qquad \textit{and} \qquad \frac{\partial v}{\partial \tau} = -x - L^{-1} D G$$

Splitting of the Verlet integrator (τ is Algorithm (MC) time)

1. Half step $\quad w_i = v_i - \frac{h}{2}\, L^{-1} D G_i$
2. Full step – Rotation

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

3. Half step $\quad v_{i+1} = w_{i+1} - \frac{h}{2}\, L^{-1} D G_{i+1}$

$$\cos\theta = \cos h \quad \textit{or} \quad \frac{4 - h^2}{4 + h^2} \qquad \sin\theta = \sin h \quad \textit{or} \quad \frac{4\,h}{4 + h^2}$$

Integration scheme is Reversible and Volume Conserving.
For finite representations, this Verlet-like splitting preserves the Quadratic Variation of the evolving path.

Metropolis-Hastings Criterion

The value of H_{eff} is almost surely infinite in the continuum limit.

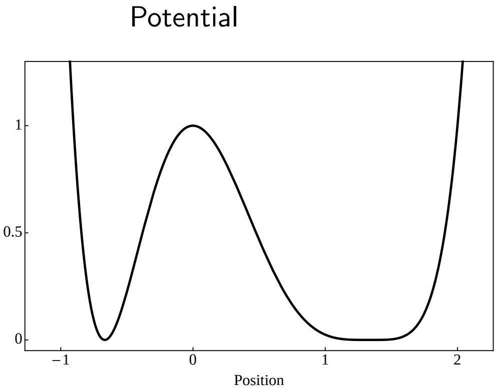
Must devise a method to calculate differences in H_{eff} as the path evolves without subtracting large (possibly infinite) numbers.

At the end of every MD step, ΔH_{eff} can be tracked.

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

Accumulate the changes as one performs MD integration. If step size, h, is small, drift in H_{eff} is minimal, the evolved path will be accepted. For large step sizes, the integration error will be substantial, and the entire sequence of paths will be rejected.

Start out with 1-dimensional Potential

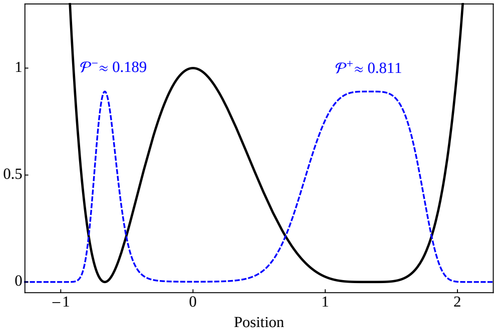


$$U(x) = \frac{(3x-4)^4\,(3x+2)^2}{1024}$$

- Degenerate Minima at $x = -\frac{2}{3}$ and $x = \frac{4}{3}$
- Barrier Height: $U_B = 1$
- $\varepsilon = 0.15$

Start out with 1-dimensional Potential

Potential

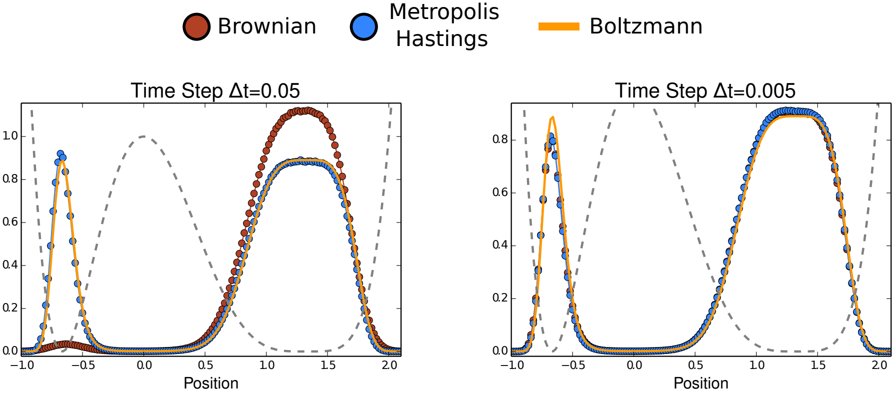


$$U(x) = \frac{(3x-4)^4\ (3x+2)^2}{1024}$$

- Degenerate Minima at $x = -\frac{2}{3}$ and $x = \frac{4}{3}$
- Barrier Height: $U_{\mathrm{B}} = 1$
- $\varepsilon = 0.15$

Boltzmann probability: entropy drives the particle to spend less time in the skinny well.

SDE with and without the Metropolis step

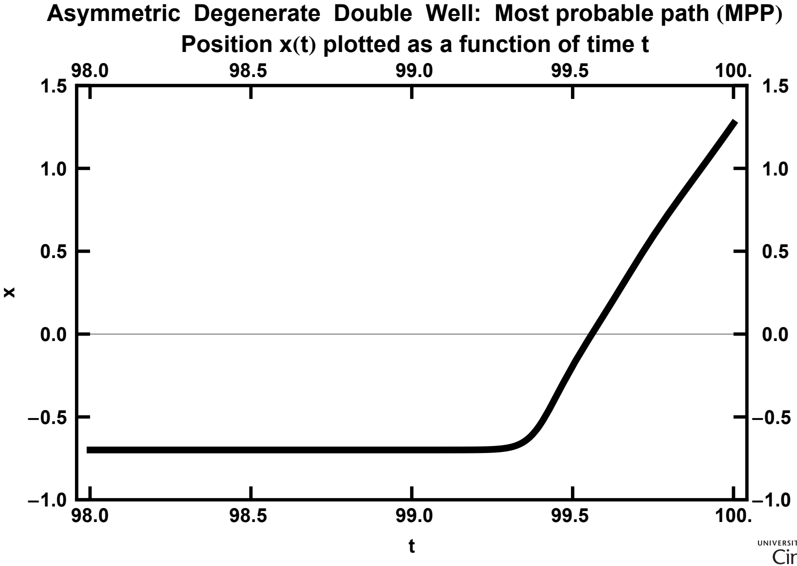


In the regime where the quadratic variation is satisfied, the Metropolis step makes only small adjustments.

UNIVERSITY OF Cincinnati

MAP estimator

Find the maximum a posteriori probability (MAP) estimate or the most probable path (MPP), fixing the path length (time) $T = 100$.



UNIVERSITY OF Cincinnati

Some paths

(Loading Video...)

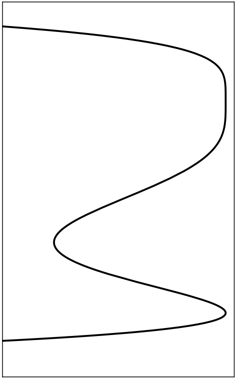
Some paths

(Loading Video...)

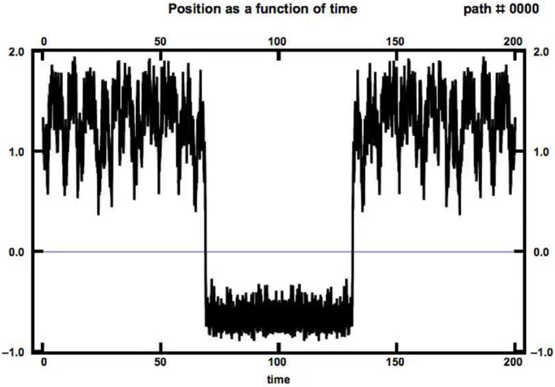
All have the same path probability ! ! ! ! ! !

Path sampling for a one dimensional well $U(x)$

Let's use this machinery to sample double ended paths.
Start with a path that was ripped from a SDE calculation.
Note that the initial path starts and ends in the same well.



Potential



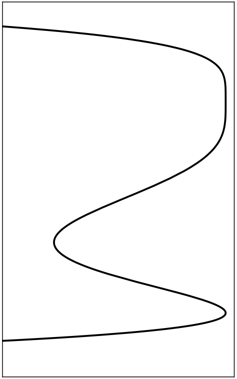
t (time along the path)

The path # denotes the evolution in τ the algorithmic time.

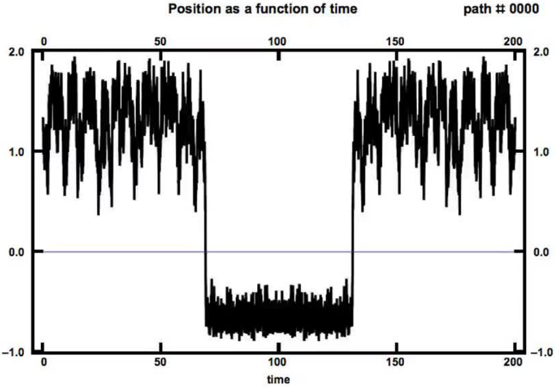
UNIVERSITY OF Cincinnati

Path sampling for a one dimensional well $U(x)$

Let's use this machinery to sample double ended paths.
Start with a path that was ripped from a SDE calculation.
Note that the initial path starts and ends in the same well.



Potential



t (time along the path)

These results are not consistent with Boltzmann distribution!

Take a new perspective

Start with a Monte-Carlo method instead of a SDE. Why?
The method samples the Boltzmann distribution and is understood.

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

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

-
-
-

UNIVERSITY OF Cincinnati

Take a new perspective

Start with a Monte-Carlo method instead of a SDE. Why?
The method samples the Boltzmann distribution and is understood.

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

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

- SDE (remember)

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

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

- Define error in the Energy

$$\delta e = \frac{1}{2}v_1^2 + U(x_1) - \frac{1}{2}v_0^2 - U(x_0)$$

Equivalence with the Onsager-Machlup functional

Manage errors with Metropolis-Hastings $\exp\left(-\delta e/\epsilon\right) > \eta$

$M(x) = Min[1, e^{-x}]$ $-\log(M(x)) = Max[0, x] = \Big(x + |x|\Big)/2$

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

When δe is small,

this new functional is equivalent to the OM functional.

Remember that the size of Δt was chosen to be small enough to ensure that the quadratic variation sum rule is satisfied.

In this regime, δe is also small and has little effect.

What went wrong ? ? ? ? ?

Hamiltonian $$H = \frac{p^2}{2\,m} + U(x)$$

Velocities are distributed according to the Maxwell-Boltzmann distribution. And the same distribution at every position x.

MAP estimator

In constructing the MPP one introduces correlations between the velocities and the positions. One wants to find the "optimal" velocities. Thus the MPPs are not consistent with the Boltzmann distribution: $\exp\Big(-U(x)/\epsilon\Big)$.

We need to ensure that the velocities are Gaussian distributed along the path.

What went wrong ? ? ? ? ?

Take a closer look at the sampling algorithm.

The first step is to generate a Brownian Bridge that will be used as the velocities in algorithmic time. These velocities govern the evolution of the paths, thereby proving a sample of paths.

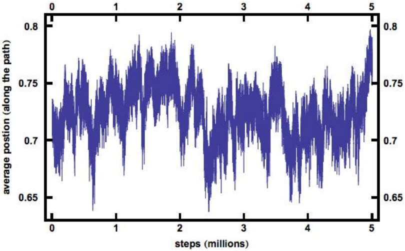
However, during the deterministic integration, "energy" is allowed to be redistributed among the modes. This is where the problems originate. The "energy" flows out of the low-frequency modes and into the many thousands of high-frequency modes.

The OM functional alone does not define the action. The action is defined by OM functional with the requirement that every frequency channel must be noisy. The deterministic integration defeats this requirement by allowing the low-frequency modes to become optimal. This then is the ingredient that forces paths to look similar to MPPs but with some high frequency noise. This latter feature is needed to ensure that the quadratic variation is satisfied.

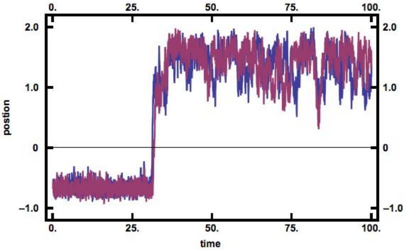
The fix

Simply do not perform the deterministic integration steps.
Stick with the Metropolis Adjusted Langevin Algorithm (MALA).

Evolution of the average position



Beginning path (blue)
Path after 5 million steps (red)



Using MALA: after 5 million moves, **the path has not collapsed.**

Concluding remarks

- Overview of Brownian dynamics
  - Generates the correct distribution (Boltzmann)
  - Inefficient when trying to sample transitions
- Derived the OM functional
  - Ensemble of transition paths
  - **Not** consistent with the Boltzmann distribution!
- **New Perspective**
  Recast the diffusion process in terms of a MCMC process
  - Method where the errors are well understood
  - Generates (almost) the same measure as the OM measure

Lesson to take home

The **double ended** paths that are generated by using the Onsager-Machlup functional alone are unphysical.
Need to require a Gaussian distribution of velocities along the path.