Gaussian Tubes in Path Space: the underlying Ornstein-Uhlenbeck Processes

Frank Pinski, University of Cincinnati

Joint work with Andrew Stuart (Warwick)

September 18, 2012

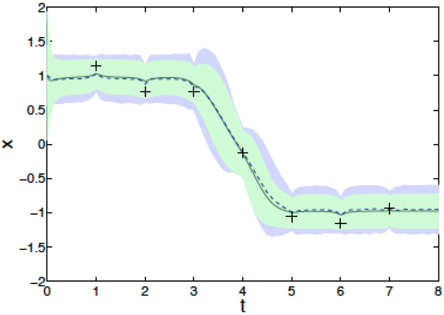
The General Problem

Consider molecules, or clusters of atoms: 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 thermodynamic significant manner.

Tubes in Path Space

Imagine the collection of such transition paths. The distribution of paths then looks like a tube. The mean of the distribution is at center of the tube with a width of the tube characterized by the fluctuations. This then is the qualitative picture of what we call Gaussian tubes.



This idea is not new. A long list of people have worked on ideas including: Eyink, Friston, Opper, Cornford, Archambeau, ...

What is new in this talk?

Here we concentrate on determining the Ornstein-Uhlenbeck (OU) Processes that are responsible for the Gaussian distribution.

By uncovering the connection between the measure in path space and the physical forces, we can developing an understanding of other properties including Free Energy, and Entropy production, and and understanding how to include external time-dependent forces, as needed for example in Optimal Control Theory.

Outline

1. Introduction
2. Variations on a Theme
   - Jensen inequality (Convex Functions)
   - Gibbs-Bogoliubov bounds (Free Energy)
   - Kullback-Leibler distance (between measures)
   - Illustrative Example (Lennard-Jones Dimer)
3. Diffusion and Bridge Diffusion
   - Underlying Brownian dynamics
   - Rare Events and Bridge diffusion
   - Generalization to time-dependent external forces
4. Gaussian Tubes
5. The KL distance
6. Possible Optimization algorithms
7. Some Numerical Considerations
8. Discussion

Variations on a Theme

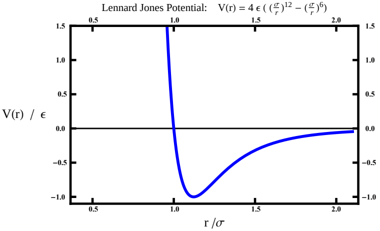
Jensen: For probability distribution with average $\mathbb{E}(...)$ and for a continuous and convex function $f(x)$, where x refers to a set of several variables, then the inequality: $\mathbb{E}\Big(f(x)\Big) \geq f(\ \mathbb{E}\, x\ )$.

Gibbs-Bogoliubov: For a system of particles interacting via a potential $\mathbb{V}(x)$ at an inverse temperature β, the probability distribution function is $P = \exp\Big(-\beta\,\mathbb{V}(x)\Big)/Z$ where Z is the partition function, and the Free Energy is $F = -\beta^{-1}\log Z$.
Using a reference system, then $Z = Z_0\ \ \mathbb{E}^0 \exp\Big(-\beta(\mathbb{V}-\mathbb{V}_0)\Big)$.
The free-energy bound: $\Phi = F_0 + \mathbb{E}^0\Big(\mathbb{V}-\mathbb{V}_0\Big) \leq F$

Kullback-Leibler: The KL-distance is a non-symmetric measure of the difference between two probability measures P and P_0.

$$D_{kl}(P||P_0) = -\int_X \log\frac{dP}{dP_0}\, dP_0 \ \geq\ 0 \qquad\qquad \Big(F - \Phi \geq 0\Big)$$

Lennard-Jones Dimer



Partition function is not integrable: $\int dr_1\, dr_2\, \exp\left(-\beta\, \mathbb{V}_{LJ}\right)$

— — — — — — — — —

Expand about $r_0 = 2^{1/6}\sigma$,

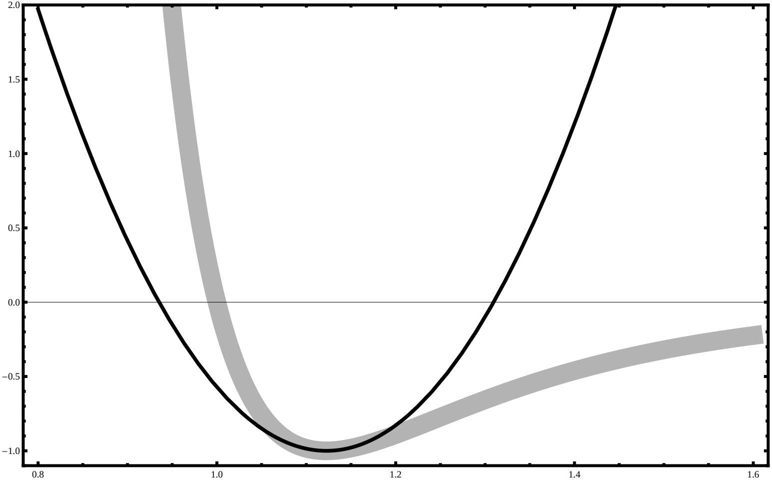
$$\widetilde{\mathbb{V}} = \mathbb{V}_0 + \frac{1}{2}(r - r_0) \cdot A \cdot (r - r_0)$$

where A is the Hessian evaluated at r_0.

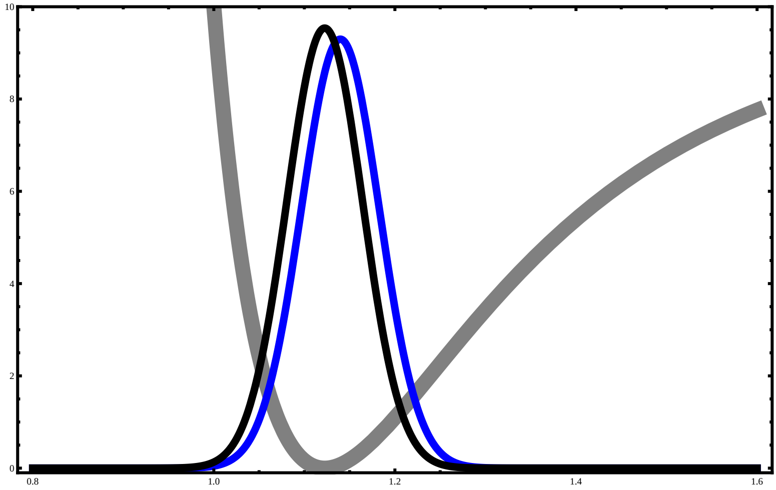
— — — — — — — — —

Vary r_0 and A to minimize $\Phi \;=\; \widetilde{F} + \mathbb{E}^G\left(\mathbb{V}_{LJ} - \widetilde{\mathbb{V}}\right)$

Lennard-Jones Potential (Gray) and its Expansion (Black)

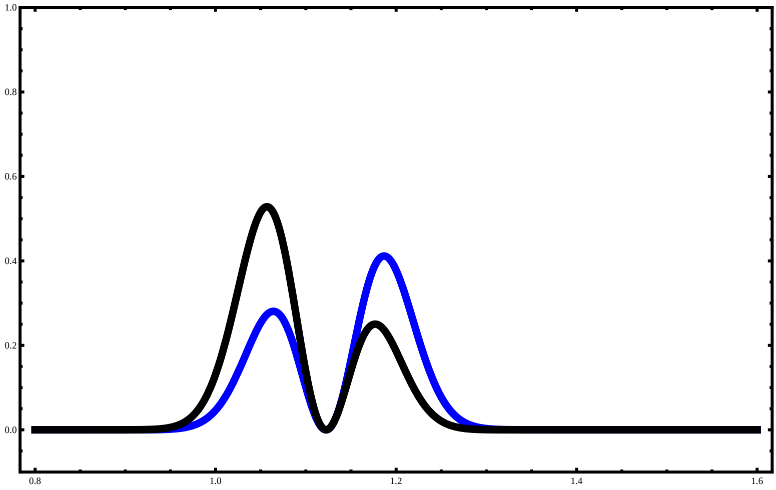


Lennard-Jones Potential (Gray)
Gaussian Distributions $\beta^{-1} = 0.1\,\epsilon_{LJ}$
from the potential expansion (Black) and optimal (Blue)



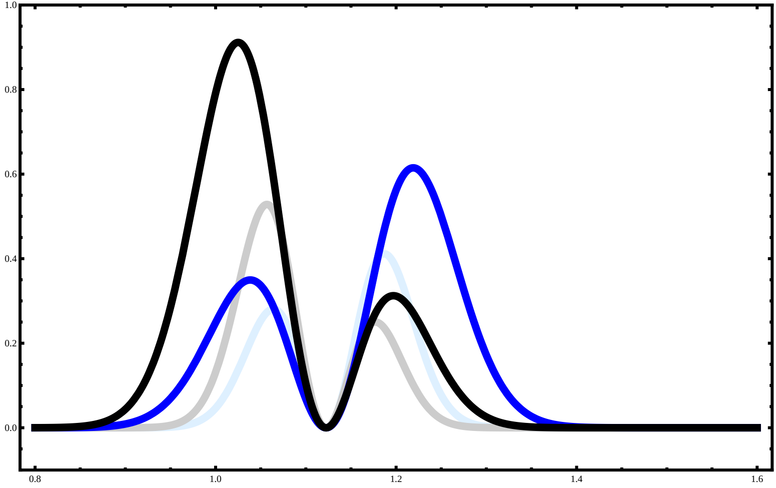
Some Analysis $\beta^{-1} = 0.1\,\epsilon_{LJ}$

Plot of the product of the potential times the Gaussian distribution



Some Analysis $\beta^{-1} = 0.2\,\epsilon_{LJ}$

Plot of the product of the potential times the Gaussian distribution



Brownian Dynamics

Add damping and random forces (thermal effects) to Newton's Law

$$\frac{d\,p}{d\,t} = F - \gamma\, p + \text{ noise}$$

Large γ limit

$$0 = F - \gamma\, p + \text{ noise}$$

Absorb γ and m into the time scale
(with β being the inverse temperature)

$$\frac{dx}{dt} = F + \sqrt{2/\beta}\;\frac{dW}{dt} = \frac{1}{\beta}\frac{\partial \log(P_B)}{\partial x} + \sqrt{2/\beta}\;\frac{dW}{dt}$$

Finite representation

$$x_{i+1} = x_i \;+\; F_i\,\Delta t + \sqrt{2\Delta t/\beta}\;\xi_i \qquad \text{with} \qquad \mathbb{P}(\xi) \propto \exp(-\xi^2/2)$$

Path Probability

Finite representation

$$x_{i+1} = x_i + F_i \,\Delta t + \sqrt{2\Delta t/\beta}\;\xi_i \qquad \text{with} \qquad \mathbb{P}(\xi) \propto \exp(-\xi^2/2)$$

$$\mathbb{P}_{path} = \prod_i \mathbb{P}(\xi_i) = C_0 \exp(-\frac{1}{2}\sum_i \xi_i^2)$$

Onsager-Machlup Functional

$$\log \mathbb{P}_{path} = C_1 - \frac{\beta}{2}\sum_i \frac{\Delta t}{2}\left(\frac{\Delta x_i}{\Delta t} - F_i\right)^2$$

Continuum limit

$$\log \mathbb{P}_{path} = C_2 - \frac{\beta}{2}\int_0^T dt\left(\frac{1}{2}\left(\frac{\partial x}{\partial t}\right)^2 + G\right)$$

with

$$G = \frac{1}{2}\left|F\right|^2 - \frac{1}{\beta}\triangle\mathbb{V}$$

Fix the boundary conditions so that all paths start in one basin and end in the other. Questions: Advantages? Mathematical rigor?

Molecular Motions

Consider a particle moving in a time-independent potential $\mathbb{V}(x)$.
Force: $F(x) = -\frac{\partial \mathbb{V}}{\partial x}$. and Brownian dynamics

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

Boundary values: $x(0) = x^-$ and $x(T) = x^+$.
The probability of such paths generates the measure ν on X.

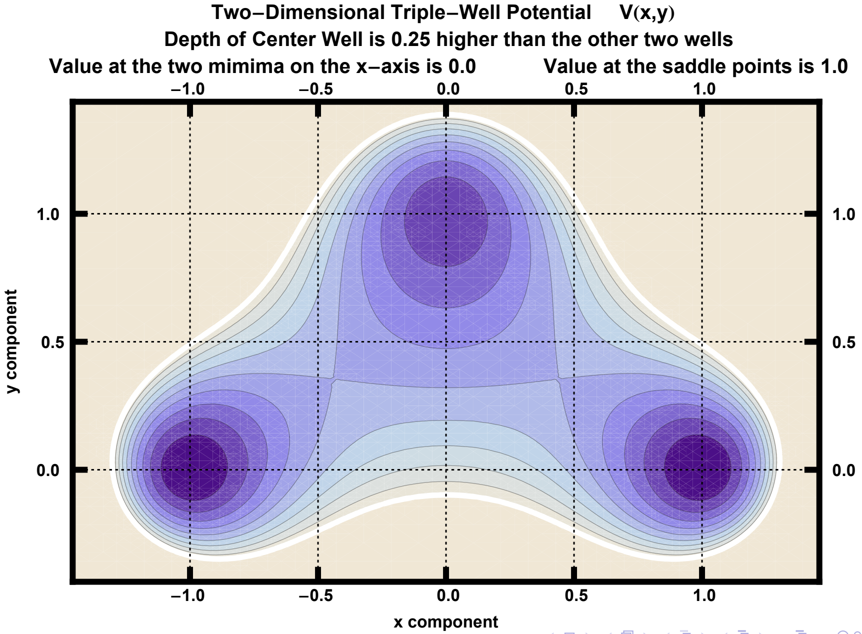
We need a result that is an application of the Ito formula:

$$d\mathbb{V} = - < F(x), dx > + \epsilon \triangle \mathbb{V} dt$$

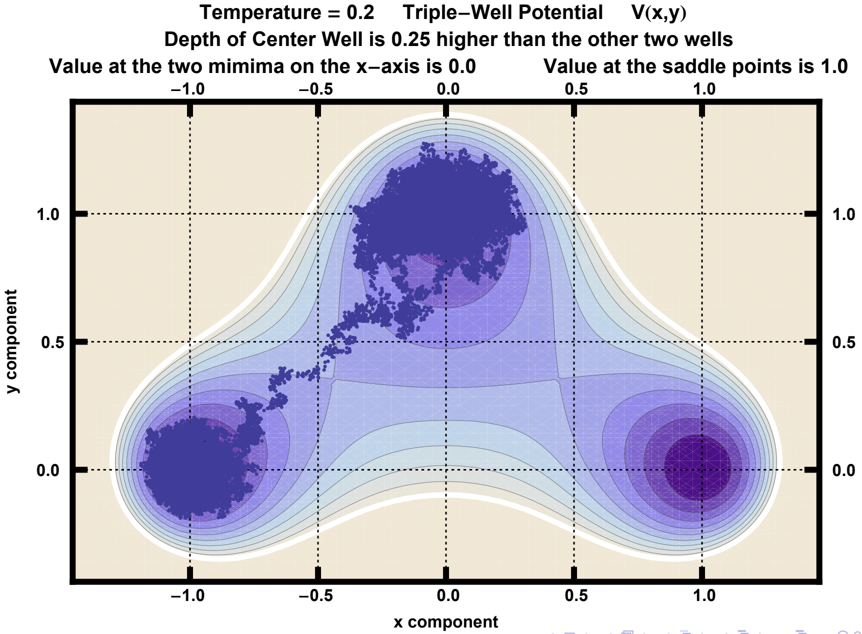
Using the Girsanov theorem, the Radon-Nikodym derivative is

$$\frac{d\nu}{d\mu_0} \propto \exp\left(-I\right) \qquad \text{with} \qquad I = \frac{1}{2\epsilon}\int_0^T dt\,\left(\frac{1}{2}\Big|F\Big|^2 - \epsilon\,\Delta\mathbb{V}\right)$$

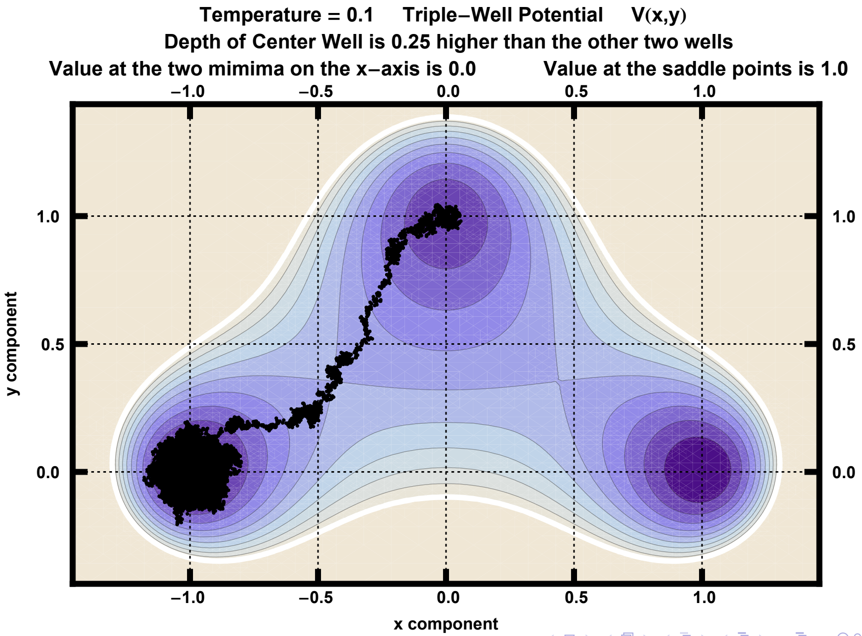
Example in Two Dimensions



Example in Two Dimensions



Example in Two Dimensions



Diffusion and Bridge Diffusion

Consider a particle moving in a time-dependent potential $\mathbb{V}_0(x,t)$.
The force on the particle is then $F_0(x,t) = -\frac{\partial \mathbb{V}_0}{\partial x}$.
For a temperature $\beta^{-1} = \epsilon$, the Brownian dynamics for the movement of a particle is

$$dx = -F_0(x,t)\,dt + \sqrt{2\,\epsilon}\;dW$$

Constrain the diffusion so that $x(0) = x^-$ and $x(T) = x^+$.
The probability of such paths generates the measure μ on X.

We need a result that is an application of the Ito formula:

$$d\mathbb{V}_0 = - < F_0(x,t), dx > + \left(\epsilon \triangle \mathbb{V}_0 + \dot{\mathbb{V}}_0\right) dt$$

Using the Girsanov theorem, the Radon-Nikodym derivative is

$$\frac{d\mu}{d\mu_0} \propto \exp\left(-I_0\right) \quad \text{with} \quad I_0 = \frac{1}{2\epsilon}\int_0^T dt \left(\frac{1}{2}\left|F_0\right|^2 - \epsilon\,\Delta\mathbb{V}_0 - \dot{\mathbb{V}}_0 \right)$$

Gaussian Tubes - 1

Consider a time-dependent Ornstein-Uhlenbeck Process

$$\mathbb{V}_{ou}(x,t) = \mathbb{V}\big(m(t)\big) - x\,\frac{dm}{dt} + \frac{1}{2}\,(x-m)A(t)(x-m)$$

$m(t)$ is the mean path or the center of the tube
$A(t)$ is the restoring force "constant"

$$F(x,t) = \frac{dm}{dt} \; - \; A(t)(x-m)$$

$\dot{m}$ pushes the particle from x^- to x^+ in a time T
$A(t)$ is needed to pull it back on course if it strays too far

Gaussian Tubes - 2

Brownian Dynamics Temperature $= \epsilon$

$$dx = \frac{dm}{dt}\,dt \;-\; A(t)(x - m)\,dt + \sqrt{2\,\epsilon}\,dW$$

The Radon-Nikodym derivative

$$\frac{d\nu_{ou}}{d\mu_0} = \frac{1}{Z_{ou}} \exp\left(-\mathbb{I}_{ou}\right)$$

with $B = A^2 - \dot{A}$ and

$$\mathbb{I}_{ou} = \frac{1}{2\epsilon}\int_0^T dt\Big(\frac{1}{2}\,\left(x - m(t)\right)\cdot B(t)\cdot \left(x - m(t)\right)$$

$$+\frac{1}{2}\,\dot{m}\cdot\dot{m} + x\cdot\ddot{m} - \epsilon\,Tr(A) + \dot{m}\cdot F(m)\Big)$$

Gaussian Tubes - 3

$$\mathbb{I}_{ou} = \frac{1}{2\epsilon}\int_0^T dt\Big(\,\frac{1}{2}\,\big(x - m(t)\big)\cdot B(t)\cdot \big(x - m(t)\big)\,+\frac{1}{2}\,\dot{m}\cdot\dot{m} - \,\dot{m}\cdot\dot{x}\Big)$$

The full measure can be expressed informally in terms of

$$\tilde{I}_{ou} = \frac{1}{2\epsilon}\int_0^T dt\Big(\frac{1}{2}\Big|\dot{x} - \dot{m}\Big|^2 + \frac{1}{2}\,\big(x - m(t)\big)\cdot B(t)\cdot\big(x - m(t)\big)\Big)$$

Note that this is the form we expect. One **surprise**, the measure in path space contains $B = A^2 - \dot{A}$ and not simply A^2.

$m(t)$ is the center and $B^{-1/2}$ is the width of the Gaussian tube.

KL-Distance revisited

Now look at the KL divergence

$$D_{kl}(\nu||\nu_{ou}) = -\mathbb{E}^{ou}\log\frac{d\nu}{d\nu_{ou}} = -\mathbb{E}^{ou}\log\Big(\frac{d\nu}{d\mu_0}\ /\ \frac{d\nu_{ou}}{d\mu_0}\Big)\ \geq\ 0$$

where μ_0 is the Brownian Bridge measure and ν_{ou} is the OU measure described above, and ν is the original path measure.

The KL divergence can be written in a compact form as

$$D_{kl}(\nu||\nu_{ou}) = \mathbb{E}^{ou}(I - \mathbb{I}_{ou})\ +\ \log\Big(\mathbb{E}^{ou}\exp\Big(-(I - \mathbb{I}_{ou})\Big)\ \Big) \geq\ 0$$

with

$$I = \frac{1}{2\epsilon}\int_0^T dt\ G = \frac{1}{2\epsilon}\int_0^T dt\left(\frac{1}{2}\Big|F\Big|^2 - \epsilon\,\Delta\mathbb{V}\right)$$

and

$$\mathbb{I}_{ou} = \frac{1}{2\epsilon}\int_0^T dt\Big(\ \frac{1}{2}\ (x - m(t))\cdot B(t)\cdot(x - m(t)) -\ \dot{m}\cdot\dot{x}\Big)$$

Possible Optimization algorithms

Find the OU parameters that minimize D_{kl}.

$$D_{kl}(\nu||\nu_{ou}) = \mathbb{E}^{ou}(I - \mathbb{I}_{ou}) + \log\Big(\mathbb{E}^{ou} \exp\Big(-(I - \mathbb{I}_{ou})\Big)\Big) \geq 0$$

The Gradients

$$2\,\epsilon\,\frac{\partial D_{kl}}{\partial m} = \mathbb{E}^{ou}\Big((I - \mathbb{I}_{ou})\,(B\cdot(x-m) - \ddot{x} + \ddot{m})\Big)$$

$$-\mathbb{E}^{ou}(I - \mathbb{I}_{ou})\;\;\mathbb{E}^{ou}\Big(\,(B\cdot(x-m) - \ddot{x} + \ddot{m})\Big)$$

and

$$2\,\epsilon\,\frac{\partial D_{kl}}{\partial B_{\alpha\gamma}} = -\mathbb{E}^{ou}\Big((I - \mathbb{I}_{ou})(x-m)_\alpha(x-m)_\gamma\Big)$$

$$+\mathbb{E}^{ou}(I - \mathbb{I}_{ou})\;\;\mathbb{E}^{ou}\Big((x-m)_\alpha(x-m)_\gamma\Big)$$

Numerical Considerations

$$\mathbb{I}_{ou} = \frac{1}{2\epsilon}\int_0^T dt\Big(\frac{1}{2}\, (x - m(t)) \cdot B(t) \cdot (x - m(t)) \, - \, \dot{m} \cdot \dot{x}\Big)$$

At the beginning of the path, the particle is confined to a one of the wells. Both $\dot{m}$ and $\dot{A}$ are zero. Thus $B = A^2$ in such a region.

Extract m and A from traditional methods in these regions.

Use path sampling to concentrate on the transitional regions where $\dot{m}$ and $\dot{A}$ are not zero.

Note that once B has been found, one still needs to find A by solving a differential equation: $B = A^2 - \dot{A}$.

Lennard-Jones Clusters: $1/\beta = 0.13$

$$\mathbb{I}_{ou} = \frac{1}{2\epsilon}\int_0^T dt\Big(\frac{1}{2}\,(x - m(t)) \cdot B(t) \cdot (x - m(t)) \; - \; \dot{m} \cdot \dot{x}\Big)$$

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.

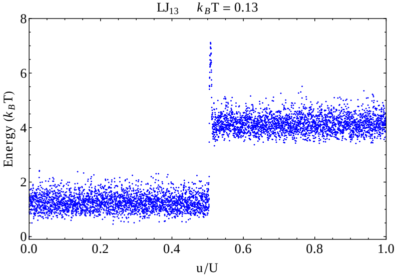


Figure: LJ_{13}: Energy along the path.

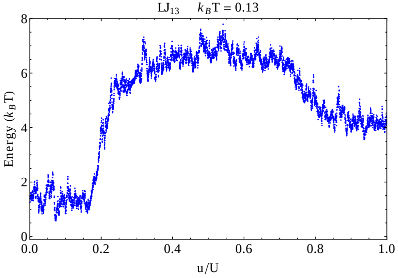


Figure: LJ_{13}: Energy along the path.

Discussion

Summary

- Tried to motivate the concept of "optimized" Gaussian distributions
- Introduced the idea of Gaussian tubes
- Started with time-dependent OU processes
- Connected the OU processes to the Gaussian tubes
- Discussed a possible algorithm for finding the optimized OU processes
- Looked briefly at some numerical considerations

Future

- Free Energy – endpoints
- Integration along the path

  Free Energy Barriers and Entropy production
- Optimal Control - imposed time-dependent external force