Rare Events and the Thermodynamic Action

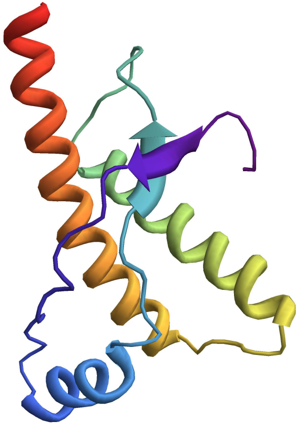
Thesis Defense

Patrick Malsom

Department of Physics, University of Cincinnati

June 25, 2015

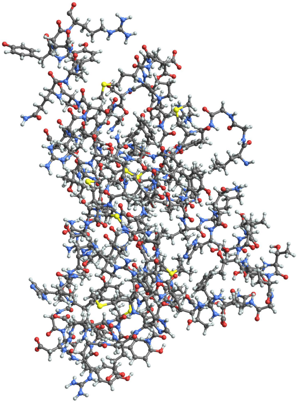
The Big Picture - Protein Folding



Original motivation was to study changes in complex chemical/biological systems.

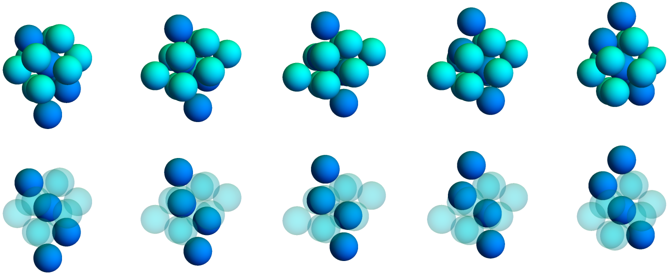
UNIVERSITY OF Cincinnati

The Big Picture - Protein Folding



Proteins are horribly complex structures. The development of novel sampling methods must be performed on simpler systems.

A Smaller Picture - Lennard-Jones Clusters



- Move to a cluster of Lennard-Jones (6-12) particles.
- Implementing new path sampling methods led to more questions than answers.

The Foundation Crumbles

- Many previous works use the Ito-Girsanov formalism
- Using above formalism, LJ cluster study lead to more confusion
- Switch to simple one-dimensional external potential and **the bottom dropped out**

What I hope to convey in this talk

- Fundamental misunderstanding of how the Ito-Girsanov expression has been used
- Identified the error
- Developed novel and more accurate formalism

UNIVERSITY OF Cincinnati

Outline

1. Sampling Paths in the Continuous-Time Limit
   - Brownian Dynamics and Onsager Machlup
   - Continuous-Time Path Probability (Ito-Girsanov)
   - Path Space Hybrid Monte-Carlo

2. Results and Implications
   - Results of 1D Example
   - Path Equivalence
   - Interpretation of the Ito-Girsanov Probabiliy

3. Lens of the Metropolis Algorithm
   - Advantages of Well Understood Method
   - OM-like functionals which produce higher-fidelity paths

Preliminaries

The goal of this work is to create an algorithm to study conformational changes which are so rare that they are inaccessible to experiment or direct computation.

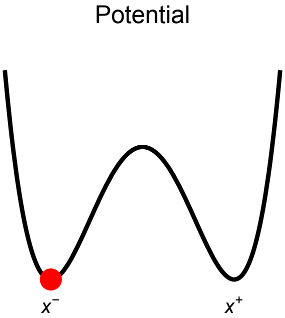
Waiting times: long when the energy barrier is large compared to the available thermal energy.

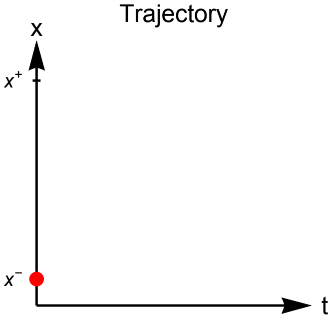
The starting point (Boltzmann and Newton)

- ε is temperature
- $F(x)$ is the force $\qquad F(x) = -\nabla V(x)$
- Classical thermodynamics

$$\mathbb{P}_B \propto \exp\left(\frac{-V(x)}{\varepsilon}\right)$$

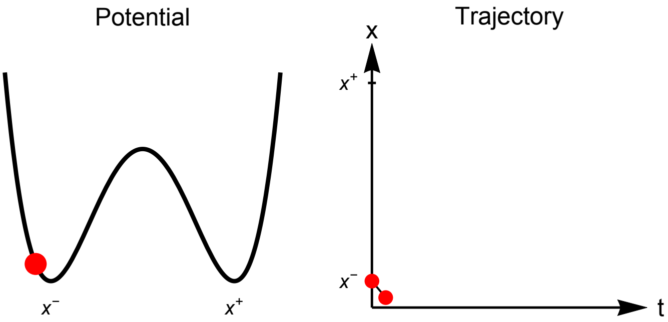
What is a Trajectory?





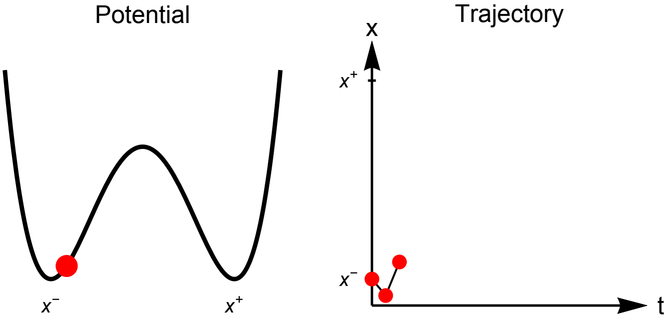
- Advance forward in time (t)
- Transitions rarely happen

What is a Trajectory?



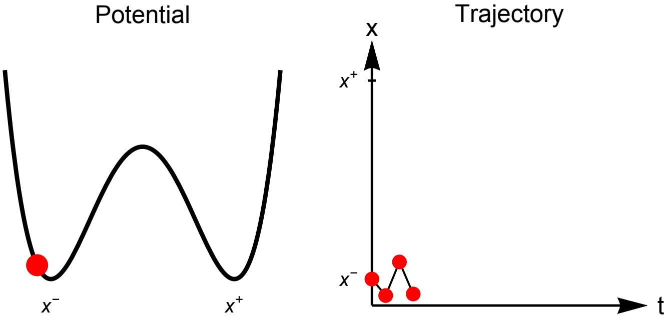
- Advance forward in time (t)
- Transitions rarely happen

What is a Trajectory?



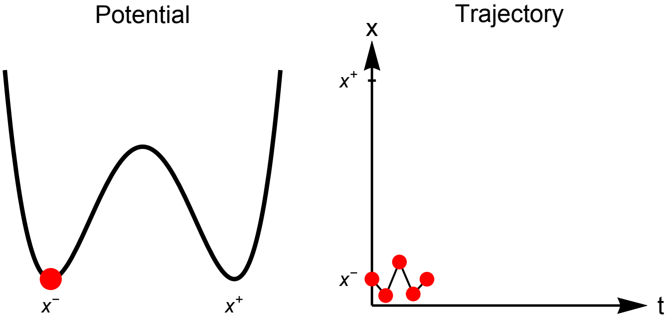
- Advance forward in time (t)
- Transitions rarely happen

What is a Trajectory?



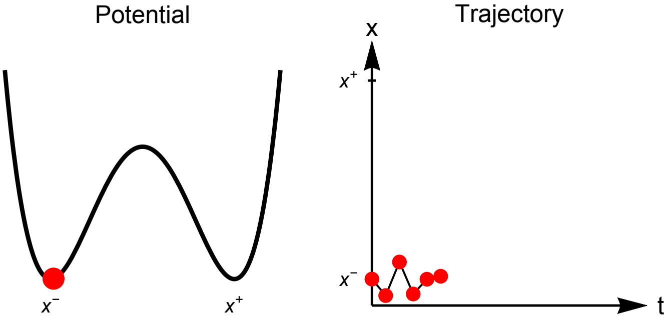
- Advance forward in time (t)
- Transitions rarely happen

What is a Trajectory?



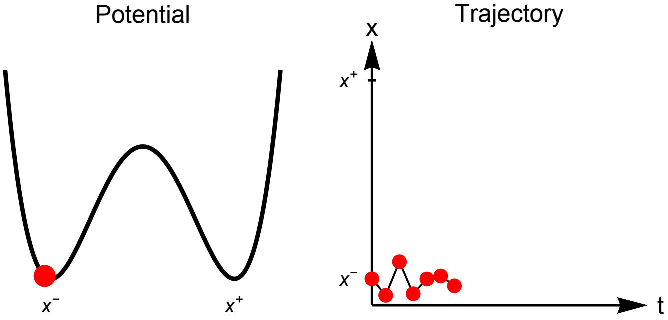
- Advance forward in time (t)
- Transitions rarely happen

What is a Trajectory?



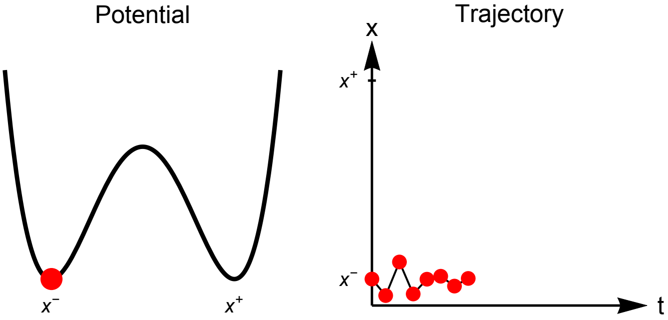
- Advance forward in time (t)
- Transitions rarely happen

What is a Trajectory?



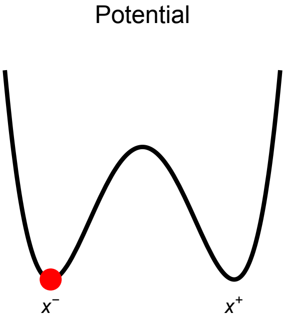
- Advance forward in time (t)
- Transitions rarely happen

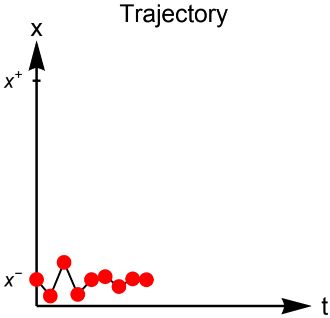
What is a Trajectory?



- Advance forward in time (t)
- Transitions rarely happen

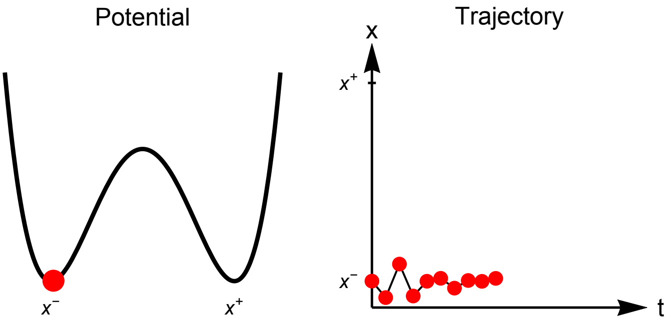
What is a Trajectory?





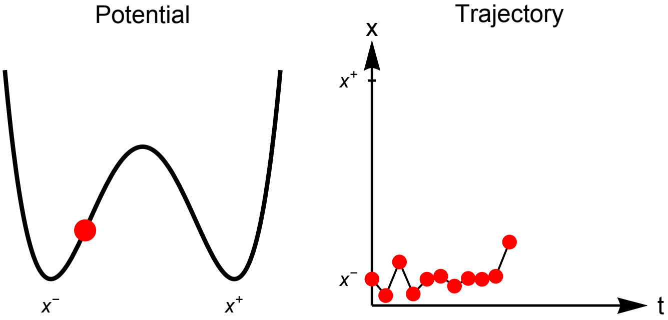
- Advance forward in time (t)
- Transitions rarely happen

What is a Trajectory?



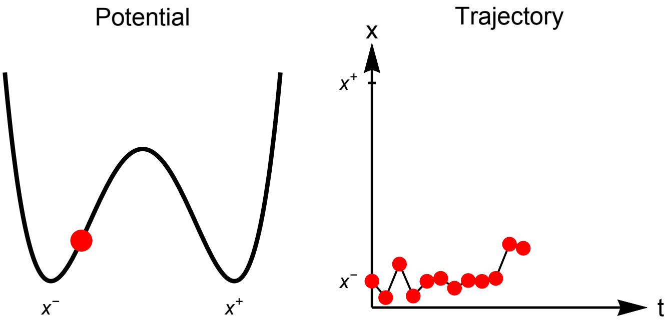
- Advance forward in time (t)
- Transitions rarely happen

What is a Trajectory?



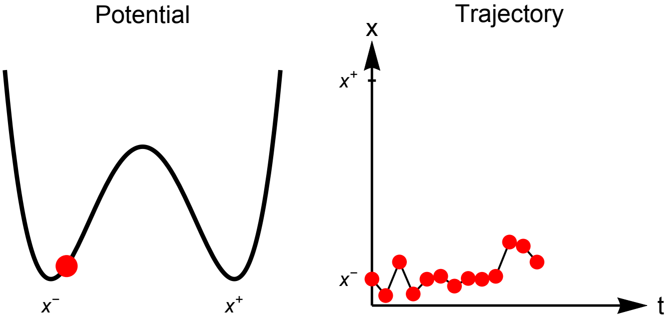
- Advance forward in time (t)
- Transitions rarely happen

What is a Trajectory?



- Advance forward in time (t)
- Transitions rarely happen

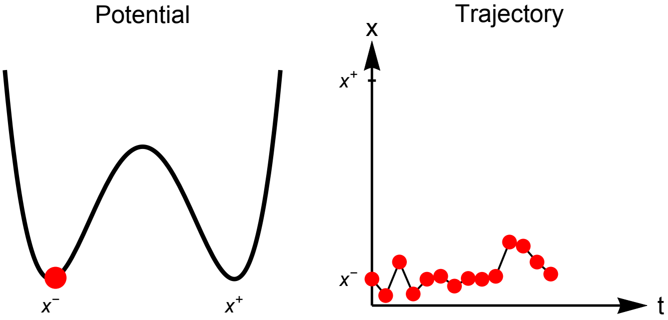
What is a Trajectory?



- Advance forward in time (t)
- Transitions rarely happen

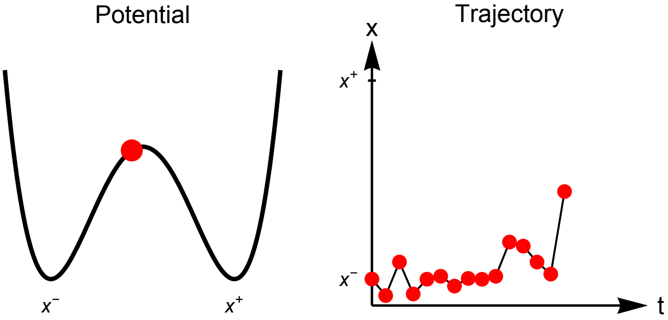
UNIVERSITY OF Cincinnati

What is a Trajectory?



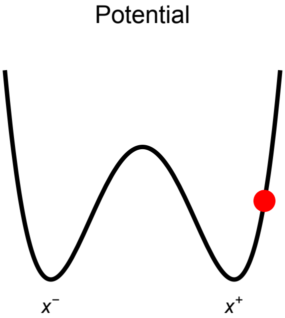
- Advance forward in time (t)
- Transitions rarely happen

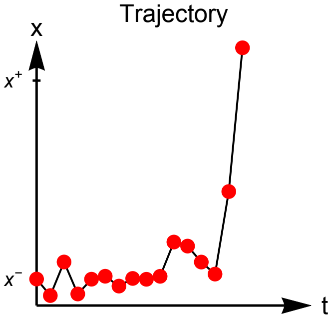
What is a Trajectory?



- Advance forward in time (t)
- Transitions rarely happen

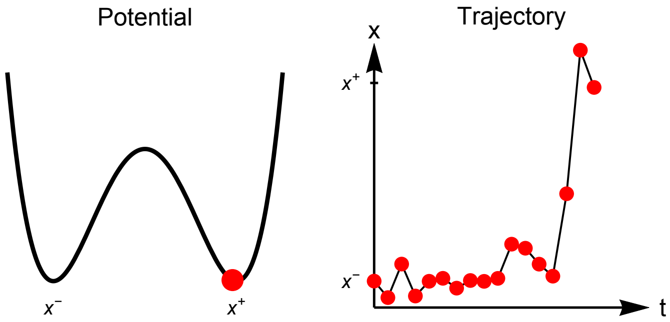
What is a Trajectory?





- Advance forward in time (t)
- Transitions rarely happen

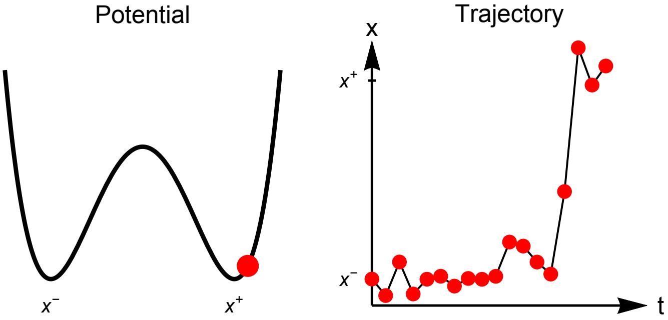
What is a Trajectory?



- Advance forward in time (t)
- Transitions rarely happen

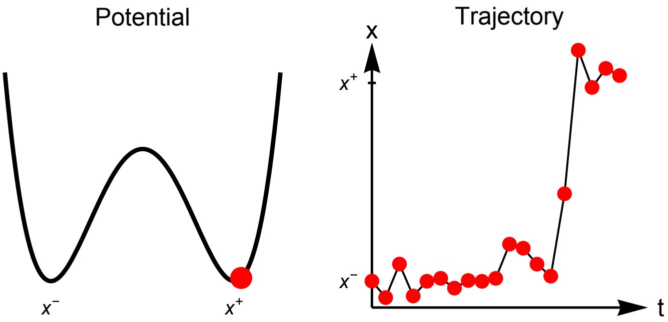
University of Cincinnati

What is a Trajectory?



- Advance forward in time (t)
- Transitions rarely happen

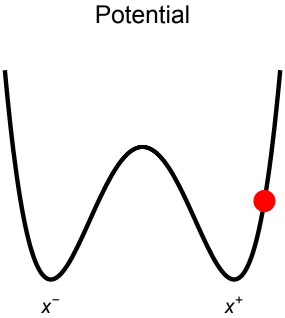
What is a Trajectory?

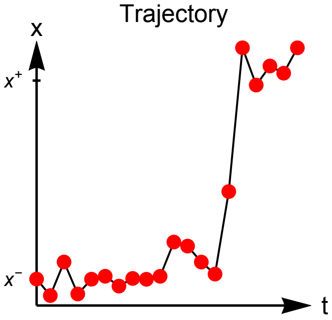


- Advance forward in time (t)
- Transitions rarely happen

UNIVERSITY OF Cincinnati

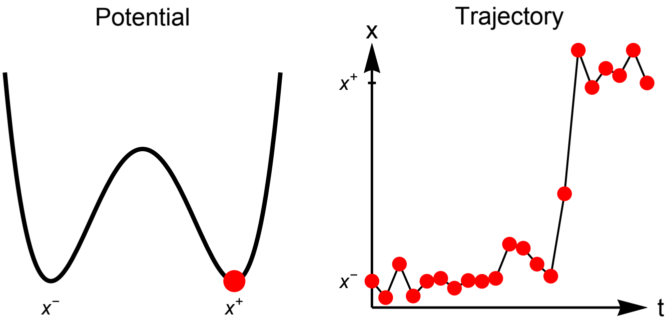
What is a Trajectory?





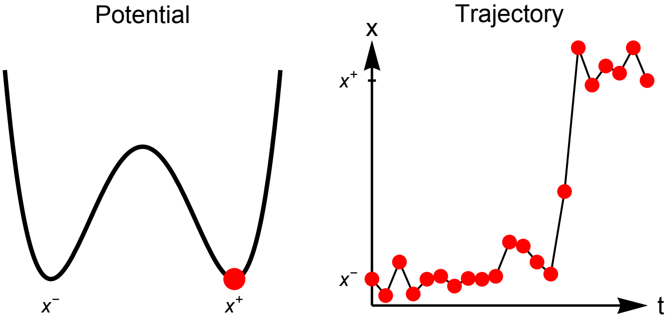
- Advance forward in time (t)
- Transitions rarely happen

What is a Trajectory?



- Advance forward in time (t)
- Transitions rarely happen

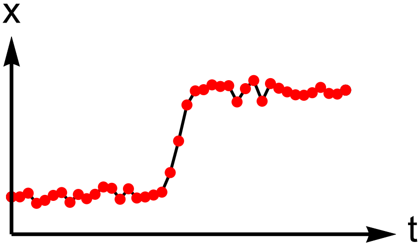
What is a Trajectory?



- Advance forward in time (t)
- Transitions rarely happen

What is a Path?

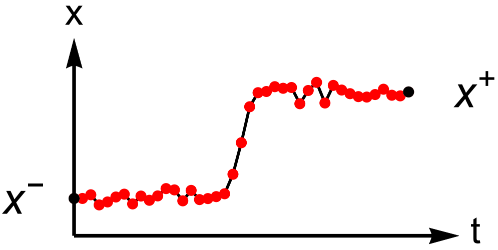
Similar to trajectories, but positions are defined for all time steps.



- Positions are a function of time (t)
- Double ended: $x(0) = x^{-}$ and $x(T) = x^{+}$
- Paths evolve over time τ

What is a Path?

Similar to trajectories, but positions are defined for all time steps.

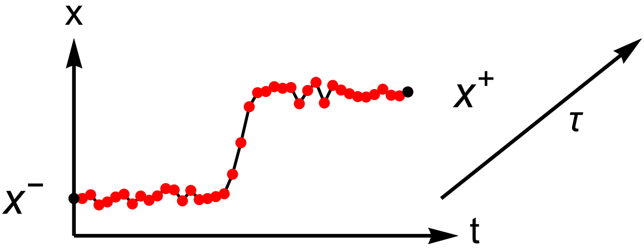


- Positions are a function of time (t)
- Double ended: $x(0) = x^-$ and $x(T) = x^+$
- Paths evolve over time τ

UNIVERSITY OF Cincinnati

What is a Path?

Similar to trajectories, but positions are defined for all time steps.

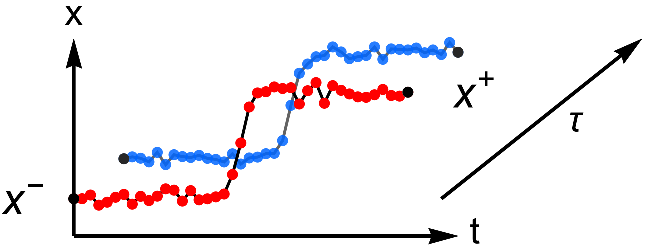


- Positions are a function of time (t)
- Double ended: $x(0) = x^-$ and $x(T) = x^+$
- Paths evolve over time τ

UNIVERSITY OF Cincinnati

What is a Path?

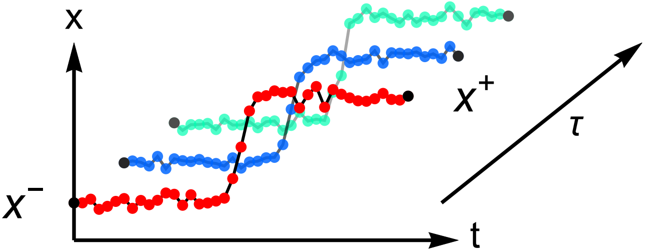
Similar to trajectories, but positions are defined for all time steps.



- Positions are a function of time (t)
- Double ended: $x(0) = x^-$ and $x(T) = x^+$
- Paths evolve over time τ

What is a Path?

Similar to trajectories, but positions are defined for all time steps.

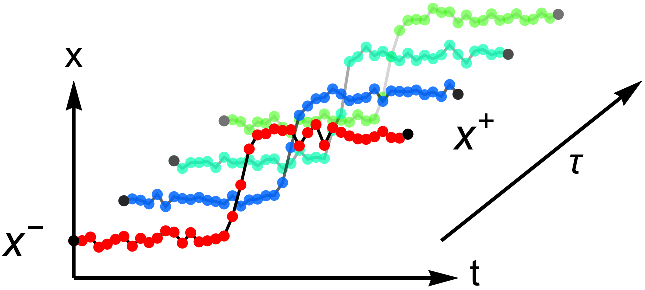


- Positions are a function of time (t)
- Double ended: $x(0) = x^-$ and $x(T) = x^+$
- Paths evolve over time τ

UNIVERSITY OF Cincinnati

What is a Path?

Similar to trajectories, but positions are defined for all time steps.

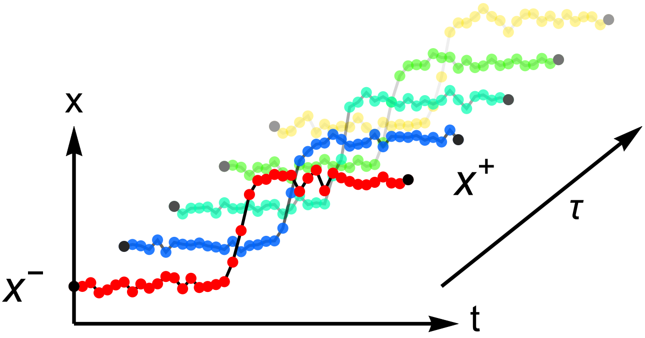


- Positions are a function of time (t)
- Double ended: $x(0) = x^-$ and $x(T) = x^+$
- Paths evolve over time τ

UNIVERSITY OF Cincinnati

What is a Path?

Similar to trajectories, but positions are defined for all time steps.



- Positions are a function of time (t)
- Double ended: $x(0) = x^-$ and $x(T) = x^+$
- Paths evolve over time τ

UNIVERSITY OF Cincinnati

Starting point

Let us begin this discussion with a damped equation of motion

$$\frac{dp}{dt'} = F - \gamma p + \text{Noise}$$

- Brownian regime: $\frac{dp}{dt'} \to 0$
- Scale the time $(t' \to \gamma m t)$: $\frac{dx}{dt} = F(x) + \text{Noise}$

Brownian Stochastic Differential Equation (SDE)

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

Discrete Brownian SDE

$$x_{i+1} - x_i = F(x_i)\, \Delta t + \sqrt{2\, \epsilon\, \Delta t}\, \xi_i$$

UNIVERSITY OF Cincinnati

Discrete sampling

$$x_{i+1} - x_i = F(x_i)\,\Delta t + \sqrt{2\,\epsilon\,\Delta t}\,\xi_i$$

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

Quadratic variation:

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

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

The probability of a path

$$x_{i+1} - x_i = F(x_i)\,\Delta t + \sqrt{2\,\epsilon\,\Delta t}\,\xi_i$$

The Onsager-Machlup Probability

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

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

Onsager-Machlup Functional

$$-\ln \mathbb{P}_{\mathrm{OM}} = \frac{\Delta t}{2\epsilon} \sum_i \frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} - F(x_i)\right)^2 + C$$

Continuous-Time Limit

$$-\ln \mathbb{P}_{\mathrm{OM}} = \frac{\Delta t}{2\epsilon} \sum_i \frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} - F(x_i) \right)^2$$

UNIVERSITY OF Cincinnati

Continuous-Time Limit

$$
\begin{aligned}
-\ln \mathbb{P}_{\mathrm{OM}} &= \frac{\Delta t}{2\epsilon} \sum_i \frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} - F(x_i) \right)^2 \\
&= \frac{\Delta t}{2\epsilon} \sum_i \left[\frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} \right)^2 + \frac{1}{2} F(x_i)^2 - \frac{\Delta x_i}{\Delta t} F(x_i) \right]
\end{aligned}
$$

UNIVERSITY OF Cincinnati

Continuous-Time Limit

$$-\ln \mathbb{P}_{\mathrm{OM}} = \frac{\Delta t}{2\epsilon} \sum_i \frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} - F(x_i) \right)^2$$
$$= \frac{\Delta t}{2\epsilon} \sum_i \left[\frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} \right)^2 + \frac{1}{2} F(x_i)^2 - \frac{\Delta x_i}{\Delta t} F(x_i) \right]$$

$$-\frac{\Delta x_i}{\Delta t} F(x_i) =$$

Continuous-Time Limit

$$-\ln \mathbb{P}_{\mathrm{OM}} = \frac{\Delta t}{2\epsilon} \sum_i \frac{1}{2}\left(\frac{\Delta x_i}{\Delta t} - F(x_i)\right)^2$$

$$= \frac{\Delta t}{2\epsilon} \sum_i \left[\frac{1}{2}\left(\frac{\Delta x_i}{\Delta t}\right)^2 + \frac{1}{2}F(x_i)^2 - \frac{\Delta x_i}{\Delta t}F(x_i)\right]$$

$$-\frac{\Delta x_i}{\Delta t}F(x_i) = \frac{1}{2}\left(\frac{\Delta x_i}{\Delta t}\left(F_{i+1} - F_i\right)\right) - \frac{1}{2}\left(\frac{\Delta x_i}{\Delta t}\left(F_{i+1} + F_i\right)\right)$$

UNIVERSITY OF Cincinnati

Continuous-Time Limit

$$
\begin{aligned}
-\ln \mathbb{P}_{\mathrm{OM}} &= \frac{\Delta t}{2\epsilon} \sum_i \frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} - F(x_i) \right)^2 \\
&= \frac{\Delta t}{2\epsilon} \sum_i \left[\frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} \right)^2 + \frac{1}{2} F(x_i)^2 - \frac{\Delta x_i}{\Delta t} F(x_i) \right]
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\Delta x_i}{\Delta t} F(x_i) &= \frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} (F_{i+1} - F_i) \right) - \frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} (F_{i+1} + F_i) \right) \\
&= \frac{1}{2} \left(\frac{(\Delta x_i)^2}{\Delta t} \left(\frac{F_{i+1} - F_i}{\Delta x_i} \right) \right) - \frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} (F_{i+1} + F_i) \right)
\end{aligned}
$$

Continuous-Time Limit

$$-\ln \mathbb{P}_{\mathrm{OM}} = \frac{\Delta t}{2\epsilon}\sum_i \frac{1}{2}\left(\frac{\Delta x_i}{\Delta t} - F(x_i)\right)^2$$

$$= \frac{\Delta t}{2\epsilon}\sum_i \left[\frac{1}{2}\left(\frac{\Delta x_i}{\Delta t}\right)^2 + \frac{1}{2}F(x_i)^2 - \frac{\Delta x_i}{\Delta t}F(x_i)\right]$$

$$-\frac{\Delta x_i}{\Delta t}F(x_i) = \frac{1}{2}\left(\frac{\Delta x_i}{\Delta t}\left(F_{i+1} - F_i\right)\right) - \frac{1}{2}\left(\frac{\Delta x_i}{\Delta t}\left(F_{i+1} + F_i\right)\right)$$

$$= \frac{1}{2}\left(\frac{(\Delta x_i)^2}{\Delta t}\left(\frac{F_{i+1} - F_i}{\Delta x_i}\right)\right) - \frac{1}{2}\left(\frac{\Delta x_i}{\Delta t}\left(F_{i+1} + F_i\right)\right)$$

$$\approx \frac{1}{2}\left(2\epsilon F'(x_i)\right) - \left(\frac{\Delta x_i}{\Delta t}\bar{F}_i\right) \qquad \boxed{\Leftarrow \textit{Quad. Var.}}$$

UNIVERSITY OF Cincinnati

Continuous-Time Limit

$$-\ln \mathbb{P}_{\mathrm{OM}} = \frac{\Delta t}{2\epsilon}\sum_i \frac{1}{2}\left(\frac{\Delta x_i}{\Delta t} - F(x_i)\right)^2$$

$$= \frac{\Delta t}{2\epsilon}\sum_i \left[\frac{1}{2}\left(\frac{\Delta x_i}{\Delta t}\right)^2 + \frac{1}{2}F(x_i)^2 - \frac{\Delta x_i}{\Delta t}F(x_i)\right]$$

$$-\frac{\Delta x_i}{\Delta t}F(x_i) = \frac{1}{2}\left(\frac{\Delta x_i}{\Delta t}(F_{i+1} - F_i)\right) - \frac{1}{2}\left(\frac{\Delta x_i}{\Delta t}(F_{i+1} + F_i)\right)$$

$$= \frac{1}{2}\left(\frac{(\Delta x_i)^2}{\Delta t}\left(\frac{F_{i+1} - F_i}{\Delta x_i}\right)\right) - \frac{1}{2}\left(\frac{\Delta x_i}{\Delta t}(F_{i+1} + F_i)\right)$$

$$\approx \frac{1}{2}\Big(2\epsilon F'(x_i)\Big) - \left(\frac{\Delta x_i}{\Delta t}\bar{F}_i\right) \qquad \boxed{\Leftarrow \textit{Quad. Var.}}$$

$$= -\epsilon\nabla^2 V(x) - \underbrace{\left(\frac{\Delta x_i}{\Delta t}\bar{F}_i\right)}_{\text{Power}}$$

Continuous-Time Limit

$$
\begin{aligned}
-\ln \mathbb{P}_{\mathrm{OM}} &= \frac{\Delta t}{2\epsilon} \sum_i \frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} - F(x_i) \right)^2 \\
&= \frac{\Delta t}{2\epsilon} \sum_i \left[\frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} \right)^2 + \frac{1}{2} F(x_i)^2 - \frac{\Delta x_i}{\Delta t} F(x_i) \right]
\end{aligned}
$$

The continuous-time probability measure (double-ended)

$$
-\ln \mathbb{P}_{\mathrm{Informal}} = \frac{1}{2\,\epsilon} \int_0^T dt \left[\frac{1}{2} \left(\frac{dx}{dt} \right)^2 + \frac{1}{2} \left| F(x) \right|^2 - \epsilon \nabla^2 V(x) \right]
$$

Continuous-Time Limit

$$-\ln \mathbb{P}_{\text{OM}} = \frac{\Delta t}{2\epsilon} \sum_i \frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} - F(x_i) \right)^2$$

$$= \frac{\Delta t}{2\epsilon} \sum_i \left[\frac{1}{2} \left(\frac{\Delta x_i}{\Delta t} \right)^2 + \frac{1}{2} F(x_i)^2 - \frac{\Delta x_i}{\Delta t} F(x_i) \right]$$

The continuous-time probability measure (double-ended)

$$-\ln \mathbb{P}_{\text{Informal}} = \frac{1}{2\,\epsilon} \int_0^T dt \left[\frac{1}{2} \left(\frac{dx}{dt} \right)^2 + \frac{1}{2} \Big| F(x) \Big|^2 - \epsilon \nabla^2 V(x) \right]$$

Regularizing this with the free-particle measure ($\mathbb{Q}$)

Ito-Girsanov probability measure

$$-\ln \mathbb{P}_{IG} = -ln \frac{d\mathbb{P}_{\text{Informal}}}{d\mathbb{Q}} = \frac{1}{2\epsilon} \int_0^T dt \left(\frac{1}{2} \Big| F(x) \Big|^2 - \epsilon \nabla^2 V(x) \right)$$

Sampling Paths with the *Thermodynamic* Action

Think of this functional as an action in the classical sense.

$$S = \frac{1}{2\,\epsilon} \int_0^T dt \left[\frac{1}{2} \left(\frac{dx}{dt} \right)^2 + \frac{1}{2} \Big| F(x) \Big|^2 - \epsilon \nabla^2 V(x) \right]$$

Impose *boundary conditions* at each end to form a path of time $T = N\;\Delta t$

- Short paths may lead to unphysical results
- Long paths should reproduce the Boltzmann distribution
- Path potential: $G(x) = \frac{1}{2} \Big| F(x) \Big|^2 - \epsilon \nabla^2 V(x)$

Sampling Paths with the *Thermodynamic* Action

Think of this functional as an action in the classical sense.

$$S = \frac{1}{2\,\epsilon}\int_0^T dt\,\left[\frac{1}{2}\left(\frac{dx}{dt}\right)^2 + \frac{1}{2}\Big|F(x)\Big|^2 - \epsilon\nabla^2 V(x)\right]$$

Impose *boundary conditions* at each end to form a path of time $T = N\,\Delta t$

- Short paths may lead to unphysical results
- Long paths should reproduce the Boltzmann distribution
- Path potential: $G(x) = \frac{1}{2}\Big|F(x)\Big|^2 - \epsilon\nabla^2 V(x)$

Construct an effective Hamiltonian from the Thermodynamic action

$$-\ln\mathbb{P} \propto S \Rightarrow H_{eff} \Rightarrow \Lambda$$

UNIVERSITY OF Cincinnati

Constructing the effective Hamiltonian Λ

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

1. Integration by parts
2. $\mathbf{L} = d^2/dt^2$
3. Add in auxiliary variables
4. Choose mass matrix to be $\mathbf{M} = -\mathbf{L}$

Constructing the effective Hamiltonian Λ

$$\Lambda = \int_0^T dt \left[-\frac{1}{2} x \frac{d^2}{dt^2} x + G(x) \right]$$

1. Integration by parts
2. $\mathbf{L} = d^2/dt^2$
3. Add in auxiliary variables
4. Choose mass matrix to be $\mathbf{M} = -\mathbf{L}$

Constructing the effective Hamiltonian Λ

$$\Lambda = \int_0^T dt \left[-\frac{1}{2} x \cdot \mathbf{L} \cdot x + G(x) \right]$$

1. Integration by parts
2. $\mathbf{L} = d^2/dt^2$
3. Add in auxiliary variables
4. Choose mass matrix to be $\mathbf{M} = -\mathbf{L}$

Constructing the effective Hamiltonian Λ

$$\Lambda = \int_0^T dt \left[-\frac{1}{2} x \cdot \mathbf{L} \cdot x + G(x) \right] + \int_0^T dt \frac{1}{2} p \cdot \mathbf{M}^{-1} \cdot p$$

1. Integration by parts
2. $\mathbf{L} = d^2/dt^2$
3. Add in auxiliary variables
4. Choose mass matrix to be $\mathbf{M} = -\mathbf{L}$

Constructing the effective Hamiltonian Λ

$$\Lambda = \int_0^T dt \left[-\frac{1}{2} x \cdot \mathbf{L} \cdot x + G(x) \right] - \int_0^T dt \frac{1}{2} p \cdot \mathbf{L}^{-1} \cdot p$$

1. Integration by parts
2. $\mathbf{L} = d^2/dt^2$
3. Add in auxiliary variables
4. Choose mass matrix to be $\mathbf{M} = -\mathbf{L}$

Constructing the effective Hamiltonian Λ

$$\Lambda = \int_0^T dt \left[-\frac{1}{2} p \cdot \mathbf{L}^{-1} \cdot p - \frac{1}{2} x \cdot \mathbf{L} \cdot x + G(x) \right]$$

1. Integration by parts
2. $\mathbf{L} = d^2/dt^2$
3. Add in auxiliary variables
4. Choose mass matrix to be $\mathbf{M} = -\mathbf{L}$

Constructing the effective Hamiltonian Λ

$$\Lambda = \int_0^T dt \left[-\frac{1}{2} p \cdot \mathbf{L}^{-1} \cdot p - \frac{1}{2} x \cdot \mathbf{L} \cdot x + G(x) \right]$$

1. Integration by parts
2. $\mathbf{L} = d^2/dt^2$
3. Add in auxiliary variables
4. Choose mass matrix to be $\mathbf{M} = -\mathbf{L}$

Use Hamilton's equations to find the equation of motion

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

Integrating Hamilton's Equations Λ

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

Splitting allows exact integration of high frequency modes

- First half step of the velocities
- Full step on positions - Exact
- Second half step of the velocities

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

$$\frac{\partial^2 x}{\partial \tau^2} = -x$$

UNIVERSITY OF Cincinnati

Integrating Hamilton's Equations Λ

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

Splitting allows exact integration of high frequency modes

- First half step of the velocities
- Full step on positions - Exact
- Second half step of the velocities

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

$$\frac{\partial^2 x}{\partial \tau^2} = -x \qquad \Leftarrow \textit{Harmonic Osc}$$

Sampling Paths with Hybrid Monte-Carlo

A. Beskos, F. Pinski, J. Sanz-Serna, and A. Stuart, Stochastic Processes and their Applications 121, 2201 (2011)

$$\Lambda = \int_0^T dt \left[-\frac{1}{2} p \cdot \mathbf{L}^{-1} \cdot p - \frac{1}{2} x \cdot \mathbf{L} \cdot x + G(x) \right]$$

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

1. Pick auxiliary variables with known distribution
2. Perform many deterministic integrations
3. Accept/Reject based on Metropolis-Hastings criteria

Sampling Paths with Hybrid Monte-Carlo

A. Beskos, F. Pinski, J. Sanz-Serna, and A. Stuart, Stochastic Processes and their Applications 121, 2201 (2011)

$$\Lambda = \int_0^T dt \left[-\frac{1}{2} p \cdot \mathbf{L}^{-1} \cdot p - \frac{1}{2} x \cdot \mathbf{L} \cdot x + G(x) \right]$$

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

1. Pick auxiliary variables with known distribution
2. Perform many deterministic integrations
3. Accept/Reject based on Metropolis-Hastings criteria

Sampling Paths with Hybrid Monte-Carlo

A. Beskos, F. Pinski, J. Sanz-Serna, and A. Stuart, Stochastic Processes and their Applications 121, 2201 (2011)

$$\Lambda = \int_0^T dt \left[-\frac{1}{2} p \cdot \mathbf{L}^{-1} \cdot p - \frac{1}{2} x \cdot \mathbf{L} \cdot x + G(x) \right]$$

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

1. Pick auxiliary variables with known distribution
2. Perform many deterministic integrations
3. Accept/Reject based on Metropolis-Hastings criteria

Sampling Paths with Hybrid Monte-Carlo

A. Beskos, F. Pinski, J. Sanz-Serna, and A. Stuart, Stochastic Processes and their Applications 121, 2201 (2011)

$$\Lambda = \int_0^T dt \left[-\frac{1}{2} p \cdot \mathbf{L}^{-1} \cdot p - \frac{1}{2} x \cdot \mathbf{L} \cdot x + G(x) \right]$$

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

1. Pick auxiliary variables with known distribution
2. Perform many deterministic integrations
3. Accept/Reject based on Metropolis-Hastings criteria

UNIVERSITY OF Cincinnati

Sampling Paths with Hybrid Monte-Carlo

A. Beskos, F. Pinski, J. Sanz-Serna, and A. Stuart, Stochastic Processes and their Applications 121, 2201 (2011)

$$\Lambda = \int_0^T dt \left[-\frac{1}{2} p \cdot \mathbf{L}^{-1} \cdot p - \frac{1}{2} x \cdot \mathbf{L} \cdot x + G(x) \right]$$

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

1. Pick auxiliary variables with known distribution
2. Perform many deterministic integrations
3. Accept/Reject based on Metropolis-Hastings criteria

Λ is almost conserved in this procedure.
The effective energy flows between different modes in step 2.

This introduces correlation in the path!

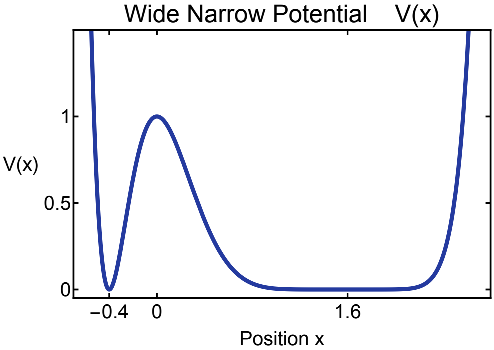
UNIVERSITY OF Cincinnati

Part 2: Results and Implications

Three Perspectives

One Dimensional Potential

I explore a simple one dimensional potential



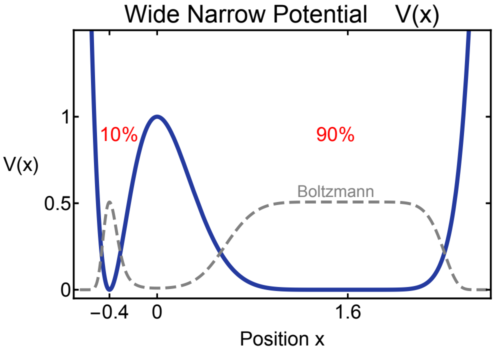
$$V(x) = \frac{(8-5x)^8(2+5x)^2}{2^{26}}$$

- Minima: $-\frac{2}{5}$ and $\frac{8}{5}$
- Barrier Height: $V_{\mathrm{bar}} = 1$

UNIVERSITY OF Cincinnati

One Dimensional Potential

I explore a simple one dimensional potential



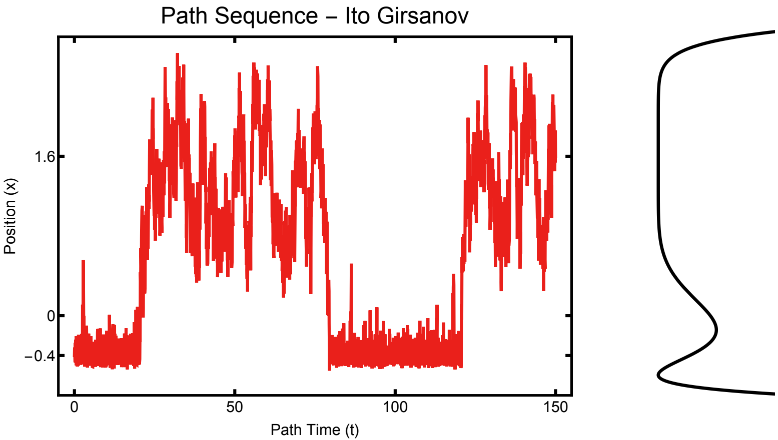
$$V(x) = \frac{(8-5x)^8(2+5x)^2}{2^{26}}$$

- Minima: $-\frac{2}{5}$ and $\frac{8}{5}$
- Barrier Height: $V_{\text{bar}} = 1$
- $\varepsilon = 0.25$

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

Example 1: Path Sampling Results

Use the Path-Space HMC machinery to sample double ended paths:



UNIVERSITY OF Cincinnati

The Garden Path

Time to think about the regularized Ito-Girsanov probability

$$-\ln \mathbb{P}_{IG} = -ln\frac{d\mathbb{P}_{\text{Informal}}}{d\mathbb{Q}} = \frac{1}{2\epsilon}\int_0^T dt\left(\frac{1}{2}\Big|F(x)\Big|^2 - \epsilon\nabla^2 V(x)\right)$$

This expression points to the idea of a Most Probable Path (MPP).

UNIVERSITY OF Cincinnati

The Garden Path

Time to think about the regularized Ito-Girsanov probability

$$-\ln \mathbb{P}_{IG} = -ln\frac{d\mathbb{P}_{\text{Informal}}}{d\mathbb{Q}} = \frac{1}{2\epsilon}\int_0^T dt\left(\frac{1}{2}\Big|F(x)\Big|^2 - \epsilon\nabla^2 V(x)\right)$$

This expression points to the idea of a Most Probable Path (MPP).

The original Onsager-Machlup probability

$$-\ln \mathbb{P}_{\text{OM}} \propto \frac{\Delta t}{2\epsilon}\sum_i \frac{1}{2}\left(\frac{\Delta x}{\Delta t} - F(x_i)\right)^2 = \sum_i \frac{\xi_i^2}{2}$$

The path probability is only dependent on the noise history.
All paths of equal length should have equal probability!

The Garden Path

Time to think about the regularized Ito-Girsanov probability

$$-\ln \mathbb{P}_{IG} = -ln \frac{d\mathbb{P}_{\text{Informal}}}{d\mathbb{Q}} = \frac{1}{2\epsilon} \int_0^T dt \left(\frac{1}{2} \left| F(x) \right|^2 - \epsilon \nabla^2 V(x) \right)$$

This expression points to the idea of a Most Probable Path (MPP).

The original Onsager-Machlup probability

$$-\ln \mathbb{P}_{\text{OM}} \propto \frac{\Delta t}{2\epsilon} \sum_i \frac{1}{2} \left(\frac{\Delta x}{\Delta t} - F(x_i) \right)^2 = \sum_i \frac{\xi_i^2}{2}$$

The path probability is only dependent on the noise history.
All paths of equal length should have equal probability!

UNIVERSITY OF Cincinnati

Example 2: Dispel the idea of the MPP

Perform the following experiment:

0. Create a set of N Gaussian random numbers.
   - 2 million GRN's
1. Generate a trajectory using the discrete SDE
   - Small time step: $\Delta t = 0.0005$
   - Large path time: $T = 1000 \Rightarrow \mathcal{O}(10)$ transitions
2. Scramble the random numbers
3. Repeat from step 1

The end points of these trajectories are approximately distributed according to the Boltzmann distribution.

Example 2: Dispel the idea of the MPP

Perform the following experiment:

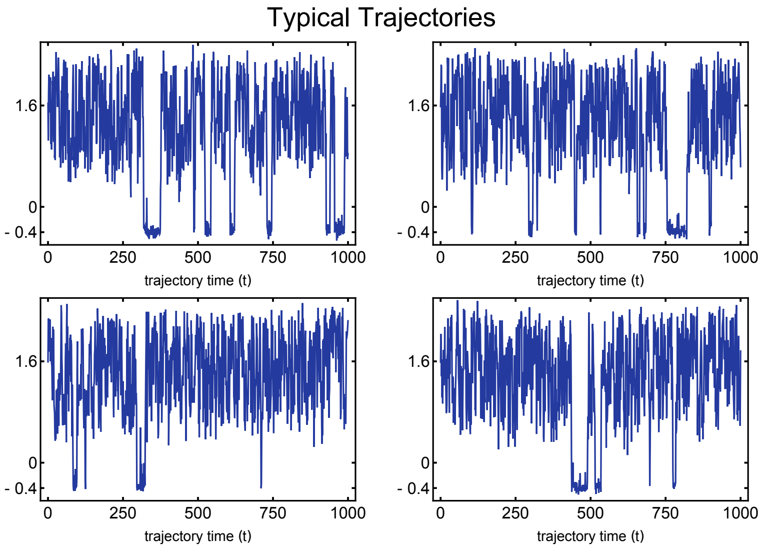
0. Create a set of N Gaussian random numbers.
   - 2 million GRN's
1. Generate a trajectory using the discrete SDE
   - Small time step: $\Delta t = 0.0005$
   - Large path time: $T = 1000 \Rightarrow \mathcal{O}(10)$ transitions
2. Scramble the random numbers
3. Repeat from step 1

The end points of these trajectories are approximately distributed according to the Boltzmann distribution.

No path is more probable than any other!

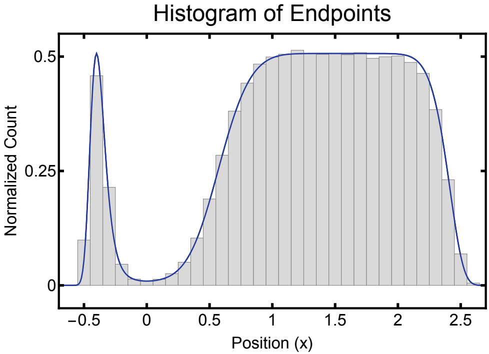
- The Most Probable Path is a misidentification
- The density of paths leads to the Boltzmann distribution

Example Paths of Equal Probability



Distribution of Endpoints

Generating many trajectories with this one set of random numbers



- Histogram shown above uses 472,640 trajectories
- More than $10^{11,733,474}$ possible trajectories
- All paths are **equally probable**

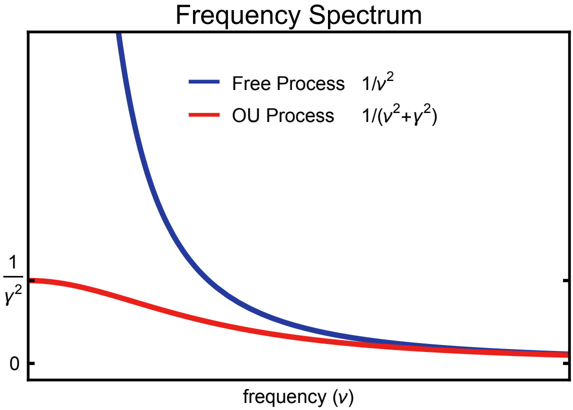
UNIVERSITY OF Cincinnati

Example 3: Frequency spectrum of OU process

Recall: *Regularization* of Ito-Girsanov measure by free measure $\mathbb{Q}$

$$-\ln \mathbb{P}_{IG} = -ln \frac{d\mathbb{P}_{\text{Informal}}}{d\mathbb{Q}}$$

Diffusion with a linear force known as the OU process.



Mixture of the free (Brownian) bridges $\neq$ solution to OU SDE

UNIVERSITY OF Cincinnati

What has gone wrong?

The previous arguments provide evidence, but what is the cause?

$$-\ln \mathbb{P}_{IG} = -ln\frac{d\mathbb{P}_{\text{Informal}}}{d\mathbb{Q}} = \frac{1}{2\epsilon}\int_0^T dt\left(\frac{1}{2}\Big|F(x)\Big|^2 - \epsilon\nabla^2 V(x)\right)$$

The noise embodied in $\mathbb{Q}$ is white noise $\rightarrow$ uncorrelated noise

The question now becomes

$$\left\langle \Delta x^2\ \frac{\Delta F}{\Delta x}\right\rangle \stackrel{?}{=} \left\langle \Delta x^2\right\rangle \left\langle\frac{\Delta F}{\Delta x}\right\rangle$$

By using this to construct the measure, robust sampling methods will allow correlation to build along the path.

Messages to take home

- The Ito-Girsanov expression has been misidentified as a path probability distribution function.
- This should not be surprising because the Ito-Girsanov expression was a change of measure from the Brownian Bridge measure and thus only has meaning in this context
- Robust sampling introduces correlation that is incompatible with the underlying dynamics (SDE)
- A full sampling of the Ito-Girsanov expression produces unphysical results

Part 3:
The Metropolis Lens

Use a well understood
algorithm to illuminate errors

Deriving OM-like functionals using Metropolis

Metropolis algorithm was designed to perform the correct thermodynamic sampling. The source of errors are transparent, unlike when integrating the Brownian SDE.

Deriving OM-like functionals using Metropolis

Metropolis algorithm was designed to perform the correct thermodynamic sampling. The source of errors are transparent, unlike when integrating the Brownian SDE.

One step HMC with Leapfrog is identical to Euler integration of Brownian SDE (no rejections)

- pick a Gaussian random distributed velocity $\sqrt{\epsilon}\,\xi_0$
- propagate Hamiltonian forward one time step
  $x_1 = x_0 + h\sqrt{\epsilon}\,\xi_0 + \frac{h^2}{2}F(x_0) \qquad h = \sqrt{2\,\Delta t}$
- repeat

Deriving OM-like functionals using Metropolis

Metropolis algorithm was designed to perform the correct thermodynamic sampling. The source of errors are transparent, unlike when integrating the Brownian SDE.

One step HMC with Leapfrog is identical to Euler integration of Brownian SDE (no rejections)

- pick a Gaussian random distributed velocity $\sqrt{\epsilon}\,\xi_0$
- propagate Hamiltonian forward one time step
  $x_1 = x_0 + h\sqrt{\epsilon}\,\xi_0 + \frac{h^2}{2}F(x_0) \qquad h = \sqrt{2\,\Delta t}$
- repeat

Derive an OM-like functional by solving for $\xi_0^2/2$
Improve the integrator to generate a more accurate functional

UNIVERSITY OF Cincinnati

Errors along the Path (time t)

$$
\begin{aligned}
\delta e &= \Delta PE + \Delta KE \\
&= V(x_1) - V(x_0) + \frac{1}{2}(v_1 - v_0)(v_1 + v_0)
\end{aligned}
$$

Note: No rejections along the path, have to live with these errors.

Errors along the Path (time t)

$$\begin{aligned}\delta e &= \Delta PE + \Delta KE \\ &= V(x_1) - V(x_0) + \frac{1}{2}(v_1 - v_0)(v_1 + v_0)\end{aligned}$$

Note: No rejections along the path, have to live with these errors.
Leapfrog:

$$\delta e = \Delta PE + \frac{1}{2}(x_1 - x_0)\big(F(x_1) + F(x_0)\big) + \frac{h^2}{8}\big(F(x_1)^2 - F(x_0)^2\big)$$

Mid-Point:

$$\delta e = \Delta PE + (x_1 - x_0)F(\bar{x})$$

Using Mid-Point integration improves the fidelity of the path.

Sampling paths with Mid-Point

An analogous routine to the Path Space HMC presented earlier is used

$$-\ln \mathbb{P}_{MP} = \frac{\Delta t}{2\epsilon}\sum_n \left[\frac{1}{2}\left(\frac{\Delta x_n}{\Delta t} - F\left(\bar{x_n}\right)\right)^2 - \frac{2\epsilon}{\Delta t}\ln\left(1 - \frac{\Delta t}{2}F'(\bar{x}_n)\right)\right]$$

The cross term is handled explicitly in this probability meausre

$$H_{eff} = \sum_n \frac{1}{2}p_n^2 + \frac{1}{2}\left|\frac{x_{n+1} - x_n}{\Delta t}\right|^2 + \frac{1}{2}F(\bar{x}_n)^2 - \frac{2\epsilon}{\Delta t}\ln\left(1 - \frac{\Delta t}{2}F'(\bar{x}_n)\right)$$

Sampling paths with Mid-Point

An analogous routine to the Path Space HMC presented earlier is used

$$-\ln \mathbb{P}_{MP} = \frac{\Delta t}{2\epsilon} \sum_n \left[\frac{1}{2} \left(\frac{\Delta x_n}{\Delta t} - F\left(\bar{x_n}\right) \right)^2 - \frac{2\epsilon}{\Delta t} \ln \left(1 - \frac{\Delta t}{2} F'(\bar{x}_n) \right) \right]$$

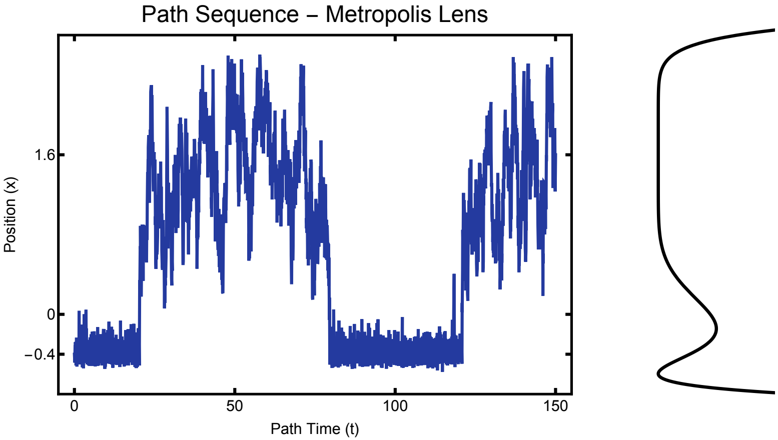
The cross term is handled explicitly in this probability meausre

$$H_{eff} = \sum_n \frac{1}{2} p_n^2 + \frac{1}{2} \left| \frac{x_{n+1} - x_n}{\Delta t} \right|^2 + \frac{1}{2} F(\bar{x}_n)^2 - \frac{2\epsilon}{\Delta t} \ln \left(1 - \frac{\Delta t}{2} F'(\bar{x}_n) \right) - \epsilon \nabla^2 V(\bar{x})$$

UNIVERSITY OF Cincinnati

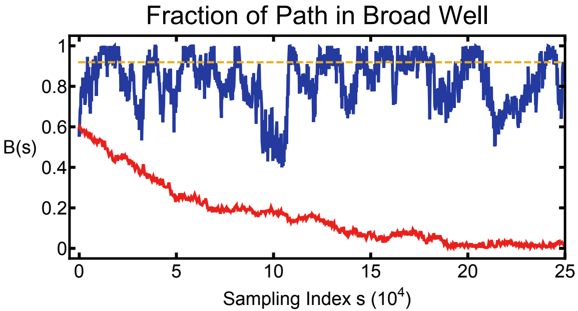
Results of the new HMC algorithm

Use this novel HMC machinery to sample double ended paths



UNIVERSITY OF Cincinnati

Conclusions



- Ito-Girsanov was misidentified as a path probability distribution functional.
- Improved OM-like functions were derived by looking at diffusion through the lens of the Metropolis algorithm.

Messages to take home

- The Ito-Girsanov expression has been misidentified as a path probability distribution function.
- This should not be surprising because the Ito-Girsanov expression was a change of measure from the Brownian Bridge measure and thus only has meaning in this context
- Robust sampling introduces correlation that is incompatible with the underlying dynamics (SDE)
- A full sampling of the Ito-Girsanov expression produces unphysical results
- Used Metropolis lens to generate novel OM-like functionals
- This novel method, using the Mid-Point integrator, correctly handled the entropic effects in the one-dimensional system