

## ME346C Advanced Techniques for Molecular Simulations -- Syllabus

Summer quarter, 2011-2012, 3 units, Stanford

MW(F) 11:00-12:15pm, 200-305

Instructor: Wei Cai, [caiwei@stanford.edu](mailto:caiwei@stanford.edu), <http://www.stanford.edu/~caiwei>

Office hour: *see coursework homepage*

TA: William (Billy) Cash, [wcash@stanford.edu](mailto:wcash@stanford.edu)

Office hour: *see coursework homepage*

Course website:

<http://coursework.stanford.edu> (official)

(last modified 6/24/2012)

### Course objective

The students will learn to perform free energy and free energy barrier calculations by molecular simulations, and understand the statistical mechanics basis of the computational methods. Methods for free energy calculation, such as thermodynamic integration (TI) and adiabatic switching (AS), will be applied to accurately predict the melting point of crystals. Methods for (free) energy barrier calculation, such as nudged-elastic band (NEB) and umbrella sampling (US), will be applied to predict the rate of rare transition events.

Prerequisite: ME346A and B

### Relation with other courses

The statistical mechanics basis has been introduced in ME346A *Introduction to statistical mechanics*. Basic molecular dynamics and Monte Carlo simulation techniques as well as how to use the MD++ program have been introduced in ME346B *Introduction to molecular simulations*. This is the first time ME346C is offered, even though the ME346ABC sequence have been contemplated for many years.

### Grading Policies

The final grade is given based on a final score, which is computed from homework (20%), midterm project (40%) and final project (40%). Homework will be issued biweekly and due on Wednesdays in class unless otherwise announced. To be fair to all students, homeworks that are late by  $n$  days will have their grades multiplied by a factor  $f(n) = (0.8)^n$ . Students are encouraged to discuss with and learn from others when solving the homework problems, but are required to write up the solution (including computer programs) individually. Each projects require a written report and a ppt presentation from each team. Students with a final score higher than 90 will receive an A.

Homework problems involve programming and numerical calculations in Matlab, simulations using MD++, and theoretical problems. Homework reports are submitted in paper. Numerical codes are submitted electronically to *coursework.stanford.edu*.

### **Textbooks and course materials**

Daan Frenkel and Berend Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. Computational science series; v.1, Academic Press, (2002). QD461.F86

Course *lecture notes* will be distributed in class and posted on coursework.

The MD++ program and tutorial materials are available through a link on coursework. Tutorials of MD++ will be given in TA sessions (location to be announced).

Other suggested readings:

Vasily V. Bulatov and Wei Cai, *Computer Simulation of Dislocations*, Oxford University Press. Companion website: <http://micro.stanford.edu/>

### **Course outline**

#### Part I. Free Energy Methods

1. Statistical Mechanics Basis (2 lectures)
2. Thermodynamic Integration (1 lectures)
3. Reversible Scaling (1 lecture)
4. Free Energy of Solids (1 lecture)
5. Free Energy of Fluids (1 lecture)
6. Predicting Melting Temperature (1 lecture)

#### Midterm project presentation

#### Part II. Free Energy Barrier Methods

7. Rate of Rare Events, Nucleation Theory (2 lectures)
8. Energy Barrier Methods, Nudged Elastic Band (2 lectures)
9. Free Energy Barrier Methods, Umbrella Sampling (2 lectures)

#### Final project presentation

## Important dates

Mon	Wed	Fri	Tentative topics	
6/25			Overview	Finding a computer to run MD++ / Matlab
	6/27		Stat Mech of Free Energy	Start looking for team mates
		6/29	Thermodynamic Integration	<b>Make-up lecture</b>
7/2			Reversible Scaling	Pick midterm project topics
	7/4		-----	Independence day, No class
		7/6	-----	Travel, No class
7/9			-----	Travel, No class <b>Final day to add / drop</b>
	7/11		-----	Travel, No class
		7/13	Free Energy of Solids	HW 1 due <b>Make-up lecture</b>
7/16			Free Energy of Fluids	
	7/18		Melting Temperature	
		7/20	-----	
7/23			<i>Student presentation</i>	<b>Midterm project presentation</b>
	7/25		Nucleation Theory	Pick final project topics
		7/27	-----	<b>Final project proposal due</b>
7/30			Energy barrier methods	
	8/1		Energy barrier methods	
		8/3	-----	<b>Course withdraw deadline</b>
8/6			Umbrella Sampling	HW 2 due
	8/8		Umbrella Sampling	
		8/10	-----	
8/13			Applications	
	8/15		Applications	
		8/17	<i>Student presentation</i>	<b>Final project presentation</b>
8/20			-----	<b>Final report due</b>

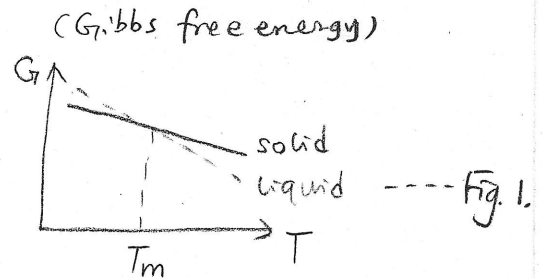
\*Exact dates of lectures may subject to change.

## 1. Importance of Free Energy

system evolves towards the direction that minimizes free energy

examples:

- why crystals melt?
- protein folding
- hydrophobic interaction — entropic origin  
(molecules <sup>that</sup> don't form hydrogen bond disrupts the # of ways water molecules can form hydrogen bonds)



## 2. Definition of Free Energy

(Helmholtz free energy,  $F$   
Gibbs free energy,  $G$ )

Thermodynamics:

- microcanonical ensemble  $(S, N, V, E)$   
 $N, V, E$  constant, system goes to equilibrium  $\Leftrightarrow S \rightarrow \max$

$E(S, V, N)$  (equilibrium) equation of state  
 $dE = TdS - pdV + \mu dN$        $T \equiv \left(\frac{\partial E}{\partial S}\right)_{V, N}$

- Canonical Ensemble

$N, V, T$  constant  $\leftarrow$  Helmholtz free energy  
Legendre transform  $F = E - TS$

$dF = -SdT - pdV + \mu dN$   
 $F(T, V, N)$        $S \equiv -\left(\frac{\partial F}{\partial T}\right)_{V, N}$

- $(N, p, S)$  Ensemble

$\leftarrow$  Enthalpy  
Legendre transform  $H = E + pV$

$dH = TdS + Vdp + \mu dN$   
 $H(S, p, N)$        $T \equiv \left(\frac{\partial H}{\partial S}\right)_{p, N}$

- $(N, p, T)$  Ensemble

Legendre transform  $G = H - TS = F + pV$

$dG = -SdT + Vdp + \mu dN$   
 $G(T, p, N)$        $S \equiv -\left(\frac{\partial G}{\partial T}\right)_{p, N}$  — see Fig. 1

Note  $G = F + pV$

So Gibbs free energy  $G$  equals Helmholtz free energy  $F$  when  $p=0$ .

### 3. Ensemble Average

Many thermodynamic quantities can be expressed as ensemble averages, which can be computed (easily) as time averages.

$$\langle A \rangle \equiv \int d^N q_i d^N p_i \rho(q_i, p_i) A(q_i, p_i) = \bar{A} \equiv \frac{1}{t_{\text{sim}}} \int_0^{t_{\text{sim}}} A(q(t), p(t)) dt$$

ensemble average

distribution function in the

ergodicity assumption

time average

trajectory in simulation

For example = internal energy  $E = \langle H \rangle$

where  $H(q_i, p_i) = \sum \frac{p_i^2}{2m} + V(\{q_i\})$  is the Hamiltonian

stress  $\sigma_{ij} = \left\langle \frac{1}{\Omega} \left( \sum_{\alpha} m v_i^{(\alpha)} v_j^{(\alpha)} - \sum_{\alpha\beta} \frac{\partial V}{\partial r_i^{(\alpha)} - r_i^{(\beta)}} (r_j^{(\alpha)} - r_j^{(\beta)}) \right) \right\rangle$

pressure  $p = -\frac{1}{3} (\sigma_{11} + \sigma_{22} + \sigma_{33})$

microcanonical ensemble

$$p_{mc}(q_i, p_i) = \begin{cases} \frac{1}{\tilde{\Omega}} & E \leq H(q_i, p_i) \leq E + \Delta E \\ 0 & \text{otherwise} \end{cases}$$

where  $\tilde{\Omega} = \int_{E \leq H(q, p) \leq E + \Delta E} d^N q d^N p$ , i.e. volume of const. enp. shell

$$\langle A \rangle_{mc} = \frac{\int_{E \leq H(q, p) \leq E + \Delta E} d^N q d^N p A(q, p)}{\int_{E \leq H(q, p) \leq E + \Delta E} d^N q d^N p}$$

i.e. an average over the const enp. shell

The microcanonical ensemble is sampled by Molecular dynamics simulation without any thermostat or barostat.

### Canonical ensemble

$$P_c(\mathbf{z}_i, \mathbf{p}_i) = \frac{1}{\tilde{Z}} e^{-\beta \mathcal{H}(\mathbf{z}_i, \mathbf{p}_i)}$$

$$\text{where } \tilde{Z} = \int d^{3N} \mathbf{q}_i d^{3N} \mathbf{p}_i e^{-\beta \mathcal{H}(\mathbf{z}_i, \mathbf{p}_i)}$$

$$\langle A \rangle_c = \frac{\int d^{3N} \mathbf{q}_i d^{3N} \mathbf{p}_i e^{-\beta \mathcal{H}(\mathbf{z}_i, \mathbf{p}_i)} A(\mathbf{z}_i, \mathbf{p}_i)}{\int d^{3N} \mathbf{q}_i d^{3N} \mathbf{p}_i e^{-\beta \mathcal{H}(\mathbf{z}_i, \mathbf{p}_i)}}$$

The canonical ensemble is sampled by Molecular dynamics simulation when a thermostat (e.g. Nose-Hoover or Andersen) is used.

It is also sampled by Monte Carlo simulation (e.g. Metropolis algorithm).

### (NPT) Ensemble

$$P_{\text{NPT}}(\mathbf{z}_i, \mathbf{p}_i, V) = \frac{1}{\tilde{\Xi}} e^{-\beta [\mathcal{H}(\mathbf{z}_i, \mathbf{p}_i) + PV]}$$

$$\text{where } \tilde{\Xi} = \int_0^\infty dV \int d^{3N} \mathbf{q}_i d^{3N} \mathbf{p}_i e^{-\beta [\mathcal{H}(\mathbf{z}_i, \mathbf{p}_i) + PV]}$$

$$\langle A \rangle_{\text{NPT}} = \frac{\int_0^\infty dV \int d^{3N} \mathbf{q}_i d^{3N} \mathbf{p}_i e^{-\beta [\mathcal{H}(\mathbf{z}_i, \mathbf{p}_i) + PV]} A(\mathbf{z}_i, \mathbf{p}_i)}{\int_0^\infty dV \int d^{3N} \mathbf{q}_i d^{3N} \mathbf{p}_i e^{-\beta [\mathcal{H}(\mathbf{z}_i, \mathbf{p}_i) + PV]}}$$

The NPT ensemble is sampled by Molecular Dynamics simulation when both a thermostat and a barostat (e.g. Parrinello-Rahman) are used.

#### 4. Free Energy cannot be expressed as Ensemble Average

\* Note that energy  $E = \langle \mathcal{H} \rangle$  can be expressed as ensemble average of Hamiltonian  $\mathcal{H}$ , and that

$F = E - TS$ , so the entropy  $S$  also cannot be expressed as ensemble average.

(otherwise, we can compute  $F$  easily from  $E - TS$ , if both  $E$  and  $S$  can be computed from ensemble/time averages)

#### Microcanonical Ensemble

Boltzmann's formula:  $S = k_B \ln \Omega$

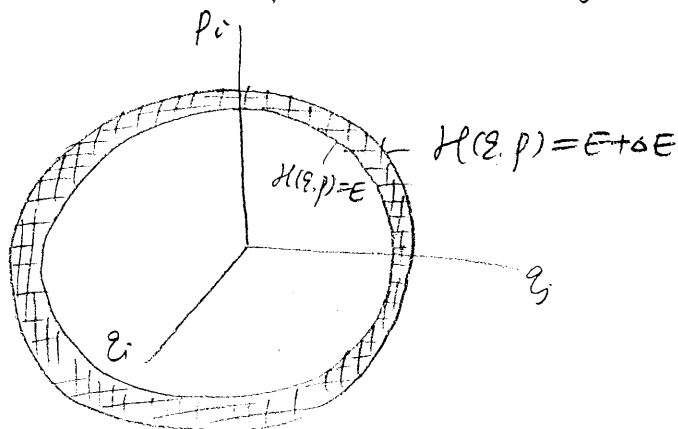
$$\text{where } \Omega = \frac{1}{N! h^{3N}} \tilde{\Omega}$$

↑  
quantum  
corrections

$h$ : Planck's constant

$$S(N, V, E) = k_B \ln \left( \frac{1}{N! h^{3N}} \int_{E \leq \mathcal{H}(\mathbf{q}, \mathbf{p}) \leq E + \Delta E} d^{3N} q_i d^{3N} p_i \right)$$

note that  $S$  is not an average over the const. enj. shell. Instead  $S$  is proportional to the  $\ln$  of the entire phase space volume of the const. enj. shell.



## Canonical Ensemble

$$F = -k_B T \ln Z$$

$$\text{where } Z = \frac{1}{N! h^{3N}} \tilde{Z}$$

↑  
quantum  
corrections

$$F(N, V, T) = -k_B T \ln \left( \frac{1}{N! h^{3N}} \int d^{3N} q_i d^{3N} p_i e^{-\beta \mathcal{H}(q_i, p_i)} \right)$$

Note that  $Z$  (the partition function) is a measure of the entire phase space volume (weighted by the Boltzmann factor  $e^{-\beta \mathcal{H}}$ ).

## (NPT) Ensemble

$$G = -k_B T \ln \Xi$$

$$\text{where } \Xi = \frac{1}{N! h^{3N}} \tilde{\Xi}$$

$$G(N, P, T) = -k_B T \ln \left( \frac{1}{N! h^{3N}} \int_0^\infty dV \int d^{3N} q_i d^{3N} p_i e^{-[\beta \mathcal{H}(q_i, p_i) + PV]} \right)$$

Here we have the statistical mechanics definitions of Helmholtz free energy ( $F$ ) and Gibbs free energy ( $G$ ), but how can we compute them numerically?

How can molecular simulation help?



## 5. Derivative of Free Energy can be written as Ensemble Average

Take canonical ensemble as an example.

Imagine that the Hamiltonian has a parameter  $\lambda$ ,  
i.e.  $H(q_i, p_i, \lambda)$ , then the Helmholtz free energy  
must also depend on  $\lambda$ . (e.g.  $\lambda$  can be some parameter in the potential func.)

$$F(N, V, T, \lambda) = -k_B T \ln \left( \frac{1}{N! h^{3N}} \int d^{3N} q_i d^{3N} p_i e^{-\beta H(q_i, p_i, \lambda)} \right)$$

We now consider the derivative of  $F$  w.r.t.  $\lambda$

$$\frac{\partial F}{\partial \lambda} = (-k_B T) \cdot \frac{\frac{1}{N! h^{3N}} \int d^{3N} q_i d^{3N} p_i e^{-\beta H(q_i, p_i, \lambda)} \cdot (-\beta) \cdot \frac{\partial H(q_i, p_i, \lambda)}{\partial \lambda}}{\frac{1}{N! h^{3N}} \int d^{3N} q_i d^{3N} p_i e^{-\beta H(q_i, p_i, \lambda)}}$$

$$\frac{\partial F}{\partial \lambda} = \frac{\int d^{3N} q_i d^{3N} p_i e^{-\beta H(q_i, p_i, \lambda)} \frac{\partial H(q_i, p_i, \lambda)}{\partial \lambda}}{\int d^{3N} q_i d^{3N} p_i e^{-\beta H(q_i, p_i, \lambda)}} \equiv \left\langle \frac{\partial H(q_i, p_i, \lambda)}{\partial \lambda} \right\rangle_{\lambda}$$

(notice  $\beta = \frac{1}{k_B T}$ )

This means that  $\frac{\partial F}{\partial \lambda}$  can be expressed as an ensemble average, and hence can be computed as a time average.

We need to be careful that the time average must be performed in a simulation under Hamiltonian  $H(q_i, p_i, \lambda)$ .

Define  $f_{\lambda} \equiv -\frac{\partial H(q_i, p_i, \lambda)}{\partial \lambda}$  as the "generalized force"

acting on parameter  $\lambda$ , then

$$-\frac{\partial F}{\partial \lambda} = \left\langle f_{\lambda} \right\rangle_{\lambda} \equiv \frac{\int d^{3N} q_i d^{3N} p_i e^{-\beta H(q_i, p_i, \lambda)} f_{\lambda}(q_i, p_i, \lambda)}{\int d^{3N} q_i d^{3N} p_i e^{-\beta H(q_i, p_i, \lambda)}}$$

i.e.  $-\frac{\partial F}{\partial \lambda}$  is the ensemble average of the 'generalized force'  $f_{\lambda}$ .

↑  
Canonical ensemble when  $H = H(q_i, p_i, \lambda)$

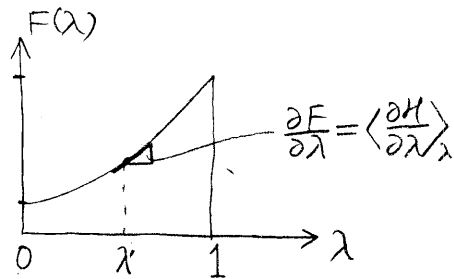
In short,  $-\frac{\partial F}{\partial \lambda}$  is the average force on parameter  $\lambda$ .

We now integrate  $\frac{\partial F}{\partial \lambda}$  from  $\lambda=0$  to  $\lambda=1$ , then

$$\begin{aligned} F(N, V, T, \lambda=1) - F(N, V, T, \lambda=0) \\ = \int_0^1 \frac{\partial F}{\partial \lambda} d\lambda = \int_0^1 \left\langle \frac{\partial H(q_i, p_i; \lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \end{aligned}$$

Similarly,

$$F(N, V, T, \lambda) = F(N, V, T, \lambda=0) + \int_0^{\lambda} \left\langle \frac{\partial H(q_i, p_i; \lambda')}{\partial \lambda'} \right\rangle_{\lambda'} d\lambda'$$



Because  $F(\lambda)$  can be obtained by integrating the average force  $f_{\lambda}$ ,  $F(\lambda)$  is often called the "potential of mean force".

## 6. Thermodynamic Integration (TI)

The TI method allows us to compute the free energy difference between two systems. If the free energy of one system (i.e. the reference system) is known analytically, then we also know the free energy of the other system.

As an example, imagine that we want to compute the Helmholtz free energy of a crystal, described by Hamiltonian

$$H_1(q_i, p_i) = \sum_i \frac{p_i^2}{2m} + V_1(\{q_i\})$$

Because  $V_1(\{q_i\})$  is a complicated interatomic potential, we do not know the free energy of system  $H_1$  analytically.

However, we do know that the free energy of an "Einstein crystal" is available analytically.

The Hamiltonian of an "Einstein crystal" is

$$H_0(q_i, p_i) = \sum_{i=1}^{3N} \frac{p_i^2}{2m} + V_0(\{q_i\}), \quad V_0(\{q_i\}) = -NE_{\text{coh}} + \sum_{i=1}^{3N} \frac{1}{2} k (q_i - q_i^0)^2$$

where  $q_i^0$  is the perfect lattice positions.

↑  
cohesive energy

Since the "Einstein crystal" is equivalent to a set of  $3N$  independent Harmonic Oscillators, its free energy is

$$F_0 \equiv -k_B T \ln \left( \frac{1}{h^{3N}} \int d^{3N} q_i d^{3N} p_i e^{-\beta H_0(q_i, p_i)} \right) = -NE_{\text{coh}} - k_B T \cdot 3N \ln \frac{k_B T}{\hbar \omega}$$

N! term is absent  
because particles are now  
distinguishable from  $q_i^0$

where  $\omega \equiv \sqrt{\frac{k}{m}}$  is the angular frequency of the oscillators  
 $\hbar \equiv \frac{h}{2\pi}$  (see ME346A Lecture Notes)

Now consider a family of Hamiltonians  $H(q_i, p_i, \lambda)$

such that as  $\lambda$  goes from 0 to 1,

$H(q_i, p_i, \lambda)$  goes smoothly from  $H_0(q_i, p_i)$  to  $H_1(q_i, p_i)$ ,

i.e.  $\lambda$  is a "knob" with which we can turn the

system from  $H_0$  (reference system) to  $H_1$  (system of interest).

$H(q_i, p_i, \lambda)$  can be constructed in many ways.

The simplest way is by linear interpolation, i.e.

$$H(q_i, p_i, \lambda) = (1-\lambda) \cdot H_0(q_i, p_i) + \lambda H_1(q_i, p_i)$$

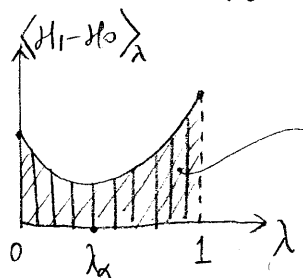
Then

$$F_1 \equiv F(N, V, T, \lambda=1) = F_0 + \int_0^1 \frac{\partial F(N, V, T)}{\partial \lambda} d\lambda$$

$$F_1 = F_0 + \int_0^1 \left\langle \frac{\partial H(q_i, p_i, \lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda$$

If  $H(\lambda) = (1-\lambda)H_0 + \lambda H_1$ , then  $\frac{\partial H}{\partial \lambda} = H_1 - H_0$ , so that

$$F_1 = F_0 + \int_0^1 \langle H_1(q_i, p_i) - H_0(q_i, p_i) \rangle_\lambda d\lambda$$



$F_1 - F_0$  is the total area under the curve and can be computed by numerical quadrature.

The integrand  $\langle H_1(q_i, p_i) - H_0(q_i, p_i) \rangle_\lambda$  is the difference between the two Hamiltonians averaged over the ensemble defined by the "mixed Hamiltonian",  $H(q, p; \lambda) = (1-\lambda)H_0(q, p) + \lambda H_1(q, p)$ .

This integral can be performed numerically, leading to the Thermodynamic Integration (TI) method.

### Algorithm 1 TI with Linear Switching Function

1. Choose number of subdivision points  $N_s$ .  
define  $\lambda_\alpha = \frac{\alpha}{N_s}$ ,  $\alpha = 0, 1, 2, \dots, N_s$ .
2. For each  $\lambda_\alpha$ , perform NVT (MD or MC) simulation using Hamiltonian  $H(q_i, p_i, \lambda_\alpha) = (1-\alpha)H_1(q_i, p_i) + \alpha H_0(q_i, p_i)$ .
3. For each  $\lambda_\alpha$ , compute time average of  $(H_1 - H_0)$  after equilibration has reached.  
The result is called  $\langle H_1 - H_0 \rangle_{\lambda_\alpha}$ .
4.  $F_1 - F_0 = \frac{1}{2} \langle H_1 - H_0 \rangle_0 + \sum_{\alpha=1}^{N_s-1} \langle H_1 - H_0 \rangle_{\lambda_\alpha} + \frac{1}{2} \langle H_1 - H_0 \rangle_1$   
(the trapezoid rule).

Today's goal:

## Physical Interpretations of Parameter $\lambda$

Because of the expression

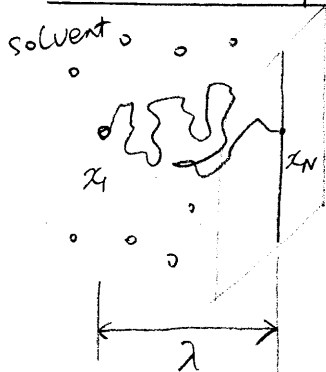
$$F(N, V, T, \lambda) = F(N, V, T, \lambda=0) + \int_0^\lambda \left\langle \frac{\partial H(\{e_i, p_i, \lambda'\})}{\partial \lambda'} \right\rangle_{\lambda'} d\lambda'$$

the Thermodynamic Integration (TI) method of computing free energy is also called  $\lambda$ -integration.

In this expression,  $\lambda$  is very general, so  $\lambda$  can mean many different things.

In practice, we take advantage of this flexibility to make the free energy calculation more efficient, i.e. reach a certain error bar with less computation time.

### 1. $\lambda$ as a physical constraint



Imaging a polymer (or protein) molecule with one end fixed at origin and the other end confined on a plane, such that

$$x_N - x_1 = \lambda$$

there may also be solvent molecules around the polymer that are not subjected to the constraint

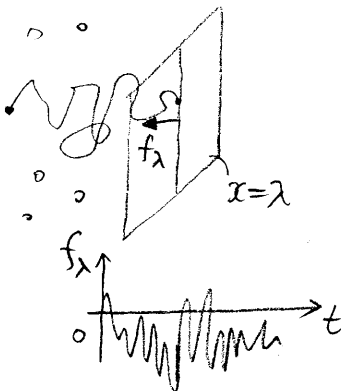
Obviously, the systems equation of motion depends on the value of constraint  $\lambda$ . We will designate this dependence in the Hamiltonian  $H(\{e_i, p_i, \lambda\})$

At temperature  $T$ , the Helmholtz free energy of the system must depend on the value  $\lambda$ , i.e.  $F(N, V, T, \lambda)$

recall

$$-\frac{\partial F(N, V, T, \lambda)}{\partial \lambda} = \langle f_\lambda \rangle_\lambda, \text{ where } f_\lambda \equiv -\frac{\partial \mathcal{H}(\mathcal{Q}_i, \mathcal{P}_i; \lambda)}{\partial \lambda}$$

Here  $f_\lambda$  is the physical force acting on the constraint (i.e. the plane  $x=\lambda$ ), which fluctuates with time.



$\langle f_\lambda \rangle_\lambda$  is the ensemble average of this force for a given value of  $\lambda$

Obviously,  $\langle f_\lambda \rangle_\lambda$  can be computed from a time average in a simulation where  $\lambda$  is fixed.

The equation

$$F(N, V, T, \lambda) - F(N, V, T, \lambda=0) = - \int_0^\lambda \langle f_{\lambda'} \rangle_{\lambda'} d\lambda'$$

can be interpreted as the following:

Imagine  $\lambda$  is gradually changed from 0 to  $\lambda$ , the free energy change of the system equals the work done to the system

$$\Delta W = - \int_0^\lambda \langle f_{\lambda'} \rangle_{\lambda'} d\lambda' \text{ is the work done as } \lambda \text{ goes from } 0 \rightarrow \lambda.$$

$$\delta W = - \langle f_{\lambda'} \rangle_{\lambda'} d\lambda' \text{ is the work done as } \lambda \text{ changes by } d\lambda'$$

The minus sign appears because the external force that does the work must balance the internal force,

$$f_\lambda^{\text{ext}} = - \langle f_\lambda \rangle_\lambda$$

\* Note: Free energy change equals stored work. so free energy is a measure of the system's ability to do work.

This is where the word "free" come from.  $F$  measures the part of energy  $E$  that is "free" to do work. The other part ( $-TS$ ) is locked up as heat and not available to do work.

## 2 $\lambda$ as volume $V$

In this case we will not write  $\lambda$  separately, but consider the derivative  $\frac{\partial F(N, V, T)}{\partial V}$

Since  $dF = -SdT - pdV + \mu dT$

we expect  $\frac{\partial F(N, V, T)}{\partial V} = -p$

If we carry out this derivative from the definition of  $F$ ,

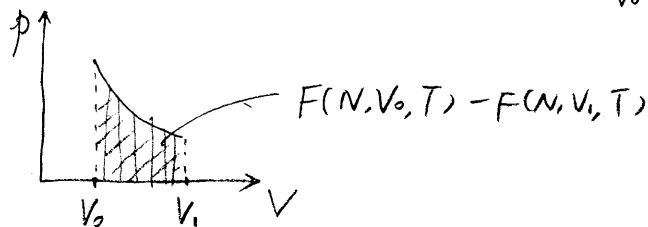
$$F(N, V, T) = -k_B T \ln \left\{ \frac{1}{N! h^{3N}} \int d^3q_i d^3p_i \exp \left[ -\beta \left( \sum \frac{p_i^2}{2m} + U(\{q_i\}) \right) \right] \right\}$$

we will arrive at the virial expression for pressure  $p$ .

(see Lecture 1, p. 2)

(for more discussion see JMP 56, 3242, 2008, p. 3247  
on course work, Extra Reading)

$$F(N, V_1, T) = F(N, V_0, T) - \int_{V_0}^{V_1} p(N, V, T) dV$$



Note that  $p$  at a given  $N, V, T$  can be computed by time average. After  $p(N, V, T)$  is computed for a range of  $V$  from  $V_0$  to  $V_1$ , the integration will tell us  $F(N, V_1, T) - F(N, V_0, T)$ .

### 3. $\lambda$ as inverse temperature $\beta = \frac{1}{k_B T}$

\* First, a short note that letting  $\lambda$  be the temperature  $T$  does not seem to get us anywhere.

$$dF = -SdT - pdV + \mu dN$$

$\therefore \frac{\partial F(N, V, T)}{\partial T} = -S$ , but  $S$  is as difficult to compute as  $F$

Let's rewrite the definition of  $F$  as

$$-\beta F(N, V, T) = \ln \left( \frac{1}{N! h^{3N}} \int d^{3N} q_i d^{3N} p_i e^{-\beta H(q_i, p_i)} \right)$$

take derivative  $\frac{\partial}{\partial \beta}$  on both sides

$$\begin{aligned} \frac{\partial}{\partial \beta} (-\beta F(N, V, T)) &= \frac{\frac{1}{N! h^{3N}} \int d^{3N} q_i d^{3N} p_i e^{-\beta H(q_i, p_i)} [-H(q_i, p_i)]}{\frac{1}{N! h^{3N}} \int d^{3N} q_i d^{3N} p_i e^{-\beta H(q_i, p_i)}} \\ &= - \langle H(q_i, p_i) \rangle \end{aligned}$$

$$\therefore \frac{\partial}{\partial \beta} (\beta F(N, V, T)) = \langle H(q_i, p_i) \rangle \equiv E$$

do we recognize this equation?

expand the L.H.S.  $\frac{\partial}{\partial \beta} (\beta F) = F + \beta \frac{\partial F}{\partial \beta}$

$$\beta = \frac{1}{k_B T}, \quad T = \frac{1}{k_B \beta}, \quad \frac{\partial T}{\partial \beta} = -\frac{1}{k_B \beta^2}, \quad \frac{\partial F}{\partial \beta} = \frac{\partial F}{\partial T} \cdot \frac{\partial T}{\partial \beta} = -\frac{1}{k_B \beta^2} \frac{\partial F}{\partial T}$$

$$\therefore \frac{\partial}{\partial \beta} (\beta F) = F - \frac{1}{k_B \beta} \frac{\partial F}{\partial T} = F - T \frac{\partial F}{\partial T} = F + TS = E$$

So  $\frac{\partial}{\partial \beta} (\beta F) = E$  is equivalent to  $F + TS = E$ .

But we can integrate this equation to compute  $F$ .

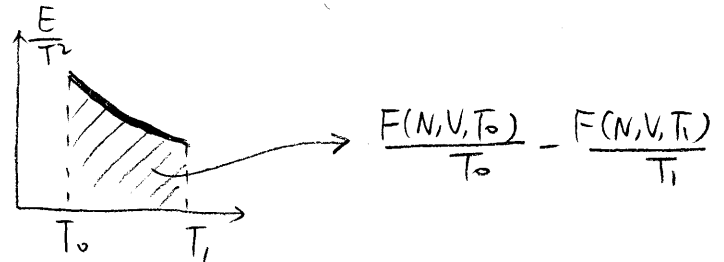
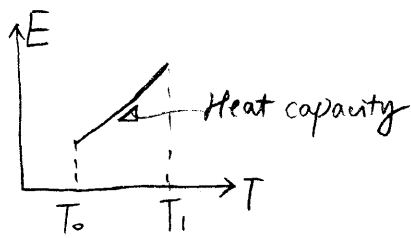
$$\beta_1 F(N, V, T_1) - \beta_0 F(N, V, T_0) = \int_{\beta_0}^{\beta_1} E(N, V, T) d\beta$$



We can write this in more familiar form in terms of  $T$ .

$$\begin{aligned} \frac{F(N, V, T_1)}{k_B T_1} - \frac{F(N, V, T_0)}{k_B T_0} &= \int_{T_0}^{T_1} E(N, V, T) \frac{d\beta}{dT} dT \\ &= \int_{T_0}^{T_1} -\frac{E(N, V, T)}{k_B T^2} dT \end{aligned}$$

$$\frac{F(N, V, T_1)}{T_1} - \frac{F(N, V, T_0)}{T_0} = - \int_{T_0}^{T_1} \frac{E(N, V, T)}{T^2} dT$$



We can compute the internal energy  $E$  in a range of temperature  $[T_0, T_1]$  using time average at each temperature.

The integral of  $\frac{E}{T^2}$  in this range of  $[T_0, T_1]$  tells us about

$$\frac{F(T_1)}{T_1} - \frac{F(T_0)}{T_0}.$$

(This is the basis of the "reversible scaling" method developed by Maurice de Groot to rapidly compute  $F$  as a function of  $T$ .)

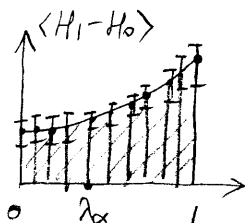
Today's goal:

Adiabatic Switching (AS) method

Ref: Computing in Science & Engineering, 2, 88-96 (2000).

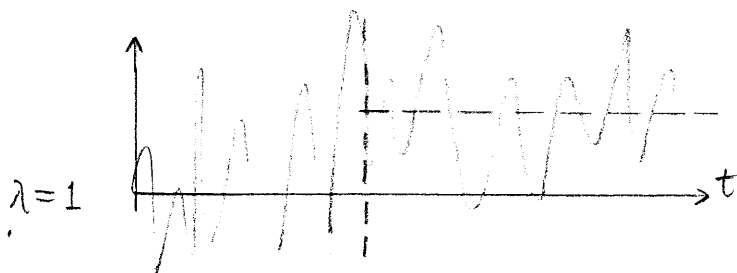
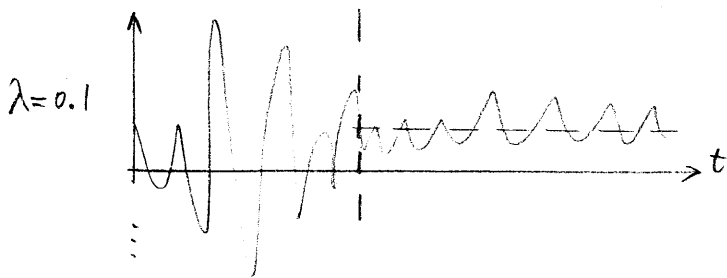
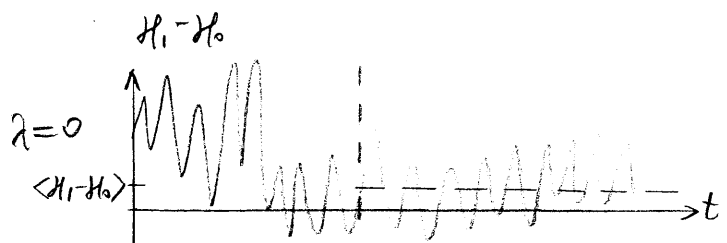
## 1. Problems with the TI Method

The TI method is essentially exact. So the only problem is a matter of efficiency and convenience.



$$H(\lambda) = (1-\lambda)H_0 + \lambda H_1$$

$$F_1 - F_0 = \int_0^1 \langle H_1 - H_0 \rangle_\lambda d\lambda$$



equilibration period

time average is performed only after equilibration has completed

$$\lambda_\alpha = \frac{\alpha}{N_s}, \quad \alpha = 0, 1, 2, \dots, N_s$$

(say,  $N_s = 10$ )

A set of NVT (MD or MC) simulations are performed, each with a different Hamiltonian  $H(\lambda_\alpha)$ .

From each simulation, the average value of  $\langle H_1 - H_0 \rangle_{\lambda_\alpha}$  is computed

Advantage:

All these simulations are independent and can be run in parallel

Disadvantage:

In each simulation, we "waste" computational time on equilibration period.

Other problems with the TI method

- We don't know a priori how many subdivision points ( $N_s$ ) (i.e. how many independent simulations) are sufficient, and how long each simulation should run.
- The fluctuation of  $H_1 - H_0$  is larger for  $\lambda$  closer to 0 or 1. So for  $\lambda$  close to 0 or 1 the simulation time should be longer to reduce the error on the average  $\langle H_1 - H_0 \rangle$ .
- Not clear whether error is dominated by numerical quadrature (insufficient  $N_s$ ) or statistical error (insufficient simulation time to compute  $\langle H_1 - H_0 \rangle$ ).

## 2. Changing $\lambda$ in a Non-equilibrium Simulation

Imagine a long MD simulation in which  $\lambda$  gradually changes from 0 to 1, from the beginning to the end of the simulation

Because  $\lambda$  is now a function of time, i.e.  $\lambda(t)$ , the Hamiltonian is also a function of time, i.e.

$$H(\lambda(t)) = (1 - \lambda(t))H_0 + \lambda(t)H_1$$

The MD simulation is by definition a non-equilibrium simulation.

It does not sample the equilibrium distribution of any equilibrium statistical ensemble.

However, if the changes of  $\lambda$  happens sufficiently slowly, we expect that the system at time  $t$  should be close to the equilibrium distribution for  $H(\hat{\lambda})$ , where  $\hat{\lambda}$  takes the value of  $\lambda(t)$ .

This is exact when the total simulation time  $t_{sim} \rightarrow \infty$ .

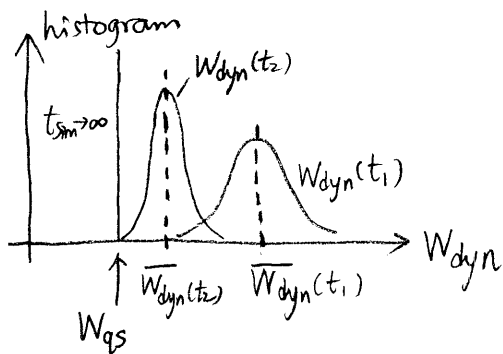
We can define the dynamical work during the non-equilibrium simulation.

$$W_{\text{dyn}} = \int_0^{t_{\text{sim}}} \frac{\partial \mathcal{H}(q_i(t), p_i(t), \lambda(t))}{\partial \lambda} \cdot \frac{d\lambda(t)}{dt} \cdot dt$$

In the limit of an infinitely long simulation, we recover the quasi-static limit, i.e. the system is always at equilibrium,

$$\lim_{t_{\text{sim}} \rightarrow \infty} W_{\text{dyn}} = W_{\text{qs}} \equiv \int_0^1 \left\langle \frac{\partial \mathcal{H}(q_i, p_i, \lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda = F_1 - F_0$$

In reality,  $t_{\text{sim}}$  is finite (although we can try to make it as long as affordable), so  $W_{\text{dyn}}$  is a random number.



You can check this by running the "switching simulation" several times.

Suppose we run the switching simulation for  $t_{\text{sim}} = t_1$  (say 1 ns). We can repeat the simulation many times, each with a different (random) initial velocity condition, and collect a histogram of  $W_{\text{dyn}}(t_1)$ .

There are both statistical error (the spread in the histogram) and systematic error (the average  $\overline{W_{\text{dyn}}(t_1)}$  does not match  $W_{\text{qs}}$ ) in  $\overline{W_{\text{dyn}}(t_1)}$ .

If we increase  $t_{\text{sim}}$  to  $t_{\text{sim}} = t_2$  (say 10 ns), then both the statistical error and the systematic errors reduce. In the limit of  $t_{\text{sim}} \rightarrow \infty$ , both errors vanish.

Q: For a given fixed total simulation time  
(constrained by your computing resource and your patience),  
do you want to run a single very long switching simulation,  
or many shorter switching simulations?

A: I want to run

---

because

---

### 3. Systematic Error as Dissipation

Define  $E_{\text{diss}} \equiv \overline{W_{\text{dyn}}} - W_{\text{qs}}$

↑  
averaged work  
from many switching  
simulations over a  
finite period  $t_{\text{sim}} = t$ .

← work in the quasi-static  
limit,  $t_{\text{sim}} \rightarrow \infty$ .  
 $W_{\text{qs}} = F_1 - F_0$ .

We can show analytically (more difficult)  
and numerically (e.g. if we know  $W_{\text{qs}}$  a priori)  
that  $E_{\text{diss}} > 0$ .

← This is guaranteed  
by the second law of  
thermodynamics.  
Otherwise, we can  
create a perpetual  
engine of the 2nd  
type.

So  $E_{\text{diss}}$  is the energy dissipated during the  
non-equilibrium process of switching  $H(\lambda(t))$  at a finite rate.

(Imagine moving a sphere through water from point A  
to point B and back to point A. The work dissipated  
decreases with decreasing speed of the sphere.)

There is a simple way to estimate  $E_{diss}$  without knowing  $W_{qs}$ .

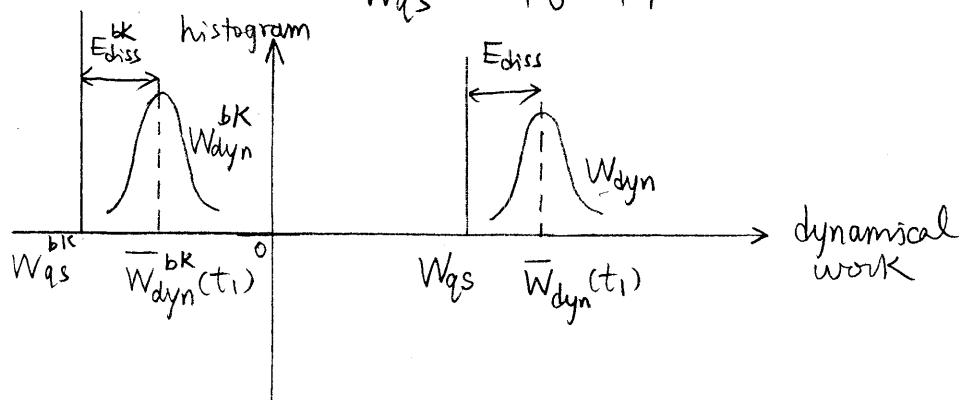
Imagine that we also perform a series of "backward switching" simulations, in which  $\lambda$  changes from 1 to 0, i.e.  $\lambda^{bk}(t)$

The work done in these backward switching simulations is

$$W_{dyn}^{bk} = \int_0^{t_{sim}} \frac{\partial \mathcal{H}(q_i(t), p_i(t), \lambda^{bk}(t))}{\partial \lambda} \cdot \underbrace{\frac{d\lambda^{bk}(t)}{dt}}_{< 0} dt$$

$$\lim_{t_{sim} \rightarrow \infty} W_{dyn}^{bk} = W_{qs}^{bk} \equiv \int_1^0 \left\langle \frac{\partial \mathcal{H}(q_i, p_i, \lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda$$

$$= -W_{qs} = F_0 - F_1$$



$$E_{diss}^{bk} \equiv \overline{W_{dyn}^{bk}} - W_{qs}^{bk} > 0$$

recall  $E_{diss} \equiv \overline{W_{dyn}} - W_{qs} > 0$

(note that the energy dissipation is positive in both forward and reverse switching.)

- Add the two equations, and note  $W_{qs}^{bk} + W_{qs} = 0$

$$\overline{W_{dyn}} + \overline{W_{dyn}^{bk}} = E_{diss} + E_{diss}^{bk} > 0$$

If we expect  $E_{diss} \approx E_{diss}^{bk}$ , then

$$E_{diss} \approx \frac{1}{2} (\overline{W_{dyn}} + \overline{W_{dyn}^{bk}})$$

- Subtract the two equations.

$$\overline{W_{dyn}} - \overline{W_{dyn}^{bk}} = 2W_{qs} + (E_{diss} - E_{diss}^{bk})$$

$$\frac{1}{2}(\overline{W}_{\text{dyn}} - \overline{W}_{\text{dyn}}^{\text{bk}}) = W_{\text{qs}} + \frac{1}{2}(E_{\text{diss}} - E_{\text{diss}}^{\text{bk}})$$

It is reasonable to expect  $E_{\text{diss}} \approx E_{\text{diss}}^{\text{bk}}$ , hence we expect

$\frac{1}{2}(\overline{W}_{\text{dyn}} - \overline{W}_{\text{dyn}}^{\text{bk}})$  to be a better estimate of  $W_{\text{qs}}$  than either  $\overline{W}_{\text{dyn}}$  or  $\overline{W}_{\text{dyn}}^{\text{bk}}$  alone. (cancellation of errors)

This is indeed true, and it is a good practice to always run a reverse switching simulation for each forward switching simulation.

However, in general  $E_{\text{diss}} \neq E_{\text{diss}}^{\text{bk}}$ ,

(alternatively, we can say that we cannot prove  $E_{\text{diss}} = E_{\text{diss}}^{\text{bk}}$ )

so  $\frac{1}{2}(\overline{W}_{\text{dyn}} - \overline{W}_{\text{dyn}}^{\text{bk}}) \neq W_{\text{qs}}$ ,

i.e. there is still a systematic error in using  $\frac{1}{2}(\overline{W}_{\text{dyn}} - \overline{W}_{\text{dyn}}^{\text{bk}})$  to estimate  $W_{\text{qs}} = F_i - F_o$ .

Today's goal:

Computational details of the Adiabatic Switching method.

1.  $\lambda(t)$
2. reference for solid
3. reference for liquid
4. GCT)

### 1. Switching function $\lambda(t)$

In principle, any continuous monotonic function that satisfies the boundary condition:

$$\begin{cases} \lambda(0) = 0 \\ \lambda(t_{\text{sim}}) = 1 \end{cases}$$

will work.

In practice, some switching functions are more efficient (i.e. leading to smaller statistical/systematic errors for the same  $t_{\text{sim}}$ ) than others.

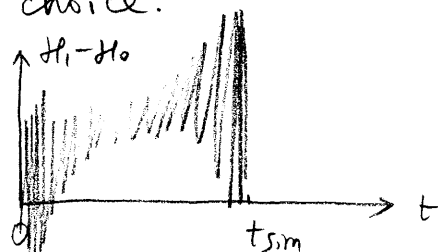
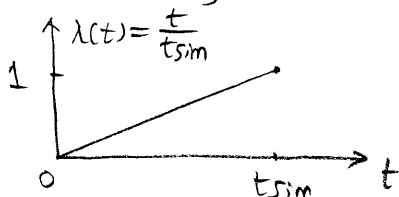
The simplest choice is a linear switching function

$$\lambda(t) = \frac{t}{t_{\text{sim}}} \quad \left( \lambda^{\text{bk}}(t) = 1 - \lambda(t) \text{ always} \right)$$

(For MC Simulation, there is no time  $t$ . But there is step number  $i_{\text{step}} = 0, 1, \dots, N_{\text{step}}$ . So equivalently,

$$\lambda(i_{\text{step}}) = \frac{i_{\text{step}}}{N_{\text{step}}}$$

This is often not an optimal choice.



because  $H_1 - H_0$  usually fluctuate violently for  $\lambda$  close to 0 or 1.

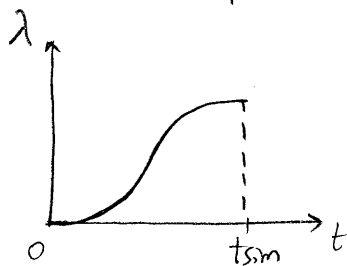
So it makes sense to switch more slowly (i.e. smaller  $\frac{d\lambda}{dt}$ ) for  $\lambda$  close to 0 or 1.



A reasonable choice is a polynomial function

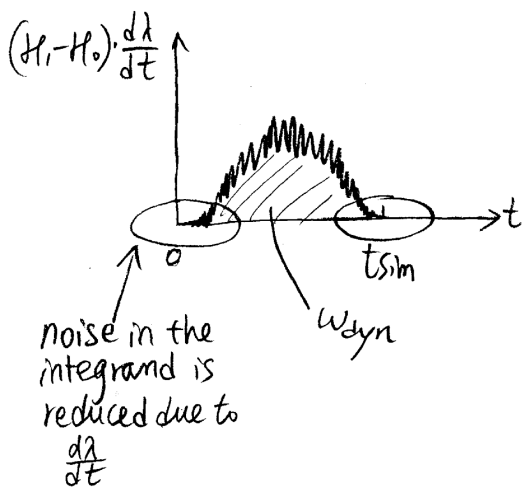
$$\lambda(t) = 70s^9 - 315s^8 + 540s^7 - 420s^6 + 126s^5, \quad s = \frac{t}{t_{\text{sim}}}$$

(used in MD++ , switchfunc = 0)



This function has zero derivative  $\frac{d\lambda}{dt}$  at  $\lambda=0, 1$ .

The time derivative  $\frac{d\lambda}{dt}$  is very small near  $\lambda=0$  or  $1$ , which cancels the large fluctuation in  $\frac{\partial H}{\partial \lambda} = H_1 - H_0$ .



Recall

$$W_{\text{dyn}} = \int_0^{t_{\text{sim}}} \underbrace{\left( \frac{\partial H}{\partial \lambda} \cdot \frac{d\lambda(t)}{dt} \right)}_{\text{integrand}} dt$$

## 2. Choice of Reference State $H_0$ for Solids

There are two criteria in choosing the reference Hamiltonian

- ① Its free energy  $F_0$  is known (either analytically or by other means)
- ② It is as close to  $H_1$  as possible.

The smaller is  $|H_1 - H_0|$ , the smaller are the statistical and systematic error in  $W_{\text{dyn}}$  for a given  $t_{\text{sim}}$ .

In lecture 1, (p.8) we have used the Einstein crystal (i.e. a set of independent Hamiltonian Oscillators) as  $H_0$ .

- We can do better than that.

Consider the Hamiltonian of a crystal (e.g. SW model of Si)

$$H_1(\{q_i, p_i\}) = \sum_i \frac{p_i^2}{2m} + U_1(\{q_i\})$$

Taylor expand  $U_1$  around perfect crystal positions  $q_i^0$

$$U_1(\{q_i\}) = \underbrace{U_1(\{q_i^0\})}_{-NE_{coh}} + \frac{1}{2} \sum_{ij} \underbrace{\frac{\partial^2 U}{\partial q_i \partial q_j} \Big|_{q_i=q_i^0}}_{K_{ij} \text{ Hessian matrix}} (q_i - q_i^0)(q_j - q_j^0) + \dots$$

define  $U_0(\{q_i\}) = -NE_{coh} + \frac{1}{2} \sum_{ij} K_{ij} (q_i - q_i^0)(q_j - q_j^0)$

which is the Taylor expansion of  $U_1$  up to 2nd order.

Let the reference Hamiltonian be

$$H_0(\{q_i, p_i\}) = \sum_i \frac{p_i^2}{2m} + U_0(\{q_i\})$$

The Helmholtz free energy of the reference system is

$$F_0 = -NE_{coh} - k_B T \sum_{i=1}^{3(N-1)} \ln \frac{k_B T}{\hbar \omega_i}$$

where  $\omega_i = \sqrt{\frac{\lambda_i}{m}}$ ,  $\lambda_i$  is the eigenvalue of the  $K_{ij}$  matrix

the three translational modes with  $\lambda_i = 0$ ,  $\omega_i = 0$  are excluded from the sum.

(See ME346A Lecture Note Chapter 10, Section 4)

$F_0$  is called the Harmonic Approximation (HA) of the free energy of interest.

The HA of  $H_1$  is a better reference system than the Einstein crystal.

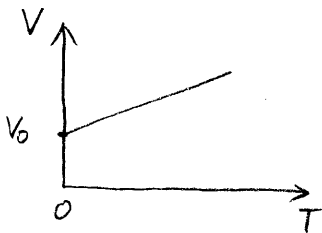
- But we can do even better than that.

We are often interested in the Gibbs free energy of the crystal under zero pressure. Due to the zero pressure, the value of the Gibbs free energy equals that of the Helmholtz free energy,

i.e. 
$$G(N, p=0, T) = F(N, V, T)$$

But we need to make sure  $V$  is the volume of the crystal at zero pressure and temperature  $T$ .

Due to thermal expansion effect,  $V > V_0$ , where  $V_0$  is the equilibrium volume of the crystal at zero pressure and zero  $T$ .



The Harmonic Approximation (HA) introduced earlier, being perfectly harmonic, contains no thermal expansion effect.

The thermal expansion effect can be incorporated by requiring  $\{q_i^0\}$  be the perfect crystal positions stretched to match the equilibrium volume  $V$  of the "real" crystal at temperature  $T$ ,

i.e. 
$$\{q_i^0(V)\}$$
, where  $V = V(T)$ .

$$U_0(\{q_i\}) = U_1(\{q_i^0(V)\}) + \frac{1}{2} \sum_{ij} K_{ij}(V) (q_i - q_i^0(V))(q_j - q_j^0(V))$$

$$F_0(N, V, T) = U_1(\{q_i^0(V)\}) - k_B T \sum_{i=1}^{3(N-1)} \ln \frac{k_B T}{\hbar \omega_i(V)}$$

where  $\omega_i = \sqrt{\frac{\lambda_i(V)}{m}}$ ,  $\lambda_i(V)$  is the eigenvalue of the  $K_{ij}(V)$  matrix.

Note that both the "cohesive energy" and the eigenfrequencies now depends on  $V$ , which implicitly depends on  $T$  due to thermal expansion.

This is called the quasi-Harmonic Approximation (QHA).

The QHA is expected to be a very good approximation,

$$\text{i.e. } \frac{|F_1 - F_0^{\text{QHA}}|}{|F_1|} \ll 1 \quad \text{at } T < T_m/2$$

but significant deviation from QHA is expected at  $T \approx T_m$ .

Using Adiabatic Switching, we can compute  $(F_1 - F_0^{\text{QHA}})$  at any temperature  $T$ , provided we first determine the zero-pressure volume  $V$  at temperature  $T$ .

(Note: In MD++<sub>1</sub>,  $K_{ij}$  can be calculated by the calHessian command. Search calHessian in <http://micro.stanford.edu/MDpp/docs/List>)

### 3. Choice of Reference State $H_0$ for Liquid

A natural reference state for a liquid is the ideal gas.

Consider a liquid at temperature  $T$  and zero pressure, with volume  $V$

$$G_l(N, p=0, T) = F_l(N, V, T)$$

The Helmholtz free energy of an ideal gas of  $N$  atoms is

$$F_{i.g.} = -N k_B T \left[ \ln\left(\frac{V}{N \Lambda^3}\right) + 1 \right]$$

where  $\Lambda = \frac{h}{\sqrt{2\pi m k_B T}}$  is the de Broglie wave length. (see ME346A Lecture Note chapter 11)

\* Note: We need to replace  $N$  by  $(N-1)$  when applying the above formula to MD simulation cells because the center-of-mass velocities are constant (zero).

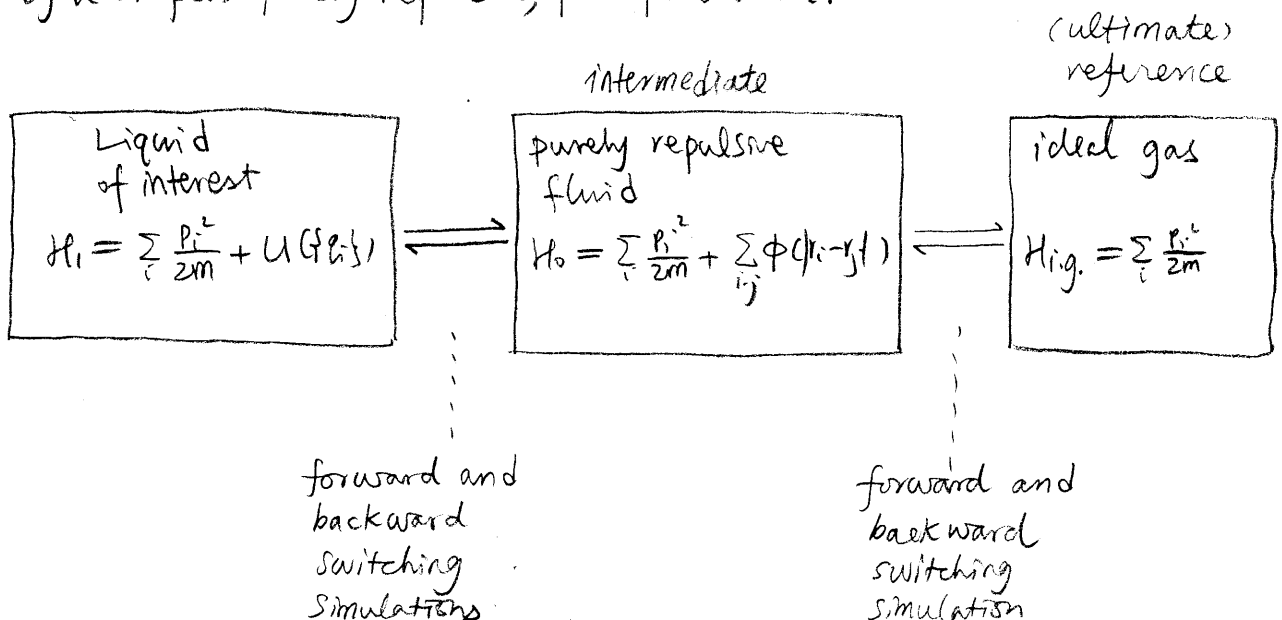
However, there are good reasons not to use the ideal gas as the (immediated) reference system  $H_0$ .

1. Liquid and gas are two different phases. On the journey of switching from a liquid to an ideal gas, the system is likely to cross a first order phase transition (i.e. boiling), which will introduce large systematic & statistical errors to the  $W_{dyn}$ .

In general, we would like to avoid phase transitions if at all possible, during a switching simulation.

2. The interatomic potential  $U_1$  is usually time consuming to evaluate. The interatomic potential for the ideal gas is  $U_0 = 0$ . But during the entire switching simulation, the expensive  $U_1$  has to be evaluated at every step, even though close to the ideal gas limit, the system behaves like a gas, and hence can be well approximated by something much cheaper to evaluate.

A reasonable choice to address these two deficiencies is to introduce an intermediate state, in which the atoms interact by a simple, purely repulsive, pair potential.



A purely repulsive pair potential we have used is

$$\phi(r_i, r_j) = \epsilon \exp\left(-\frac{|r_i - r_j|^2}{2\sigma^2}\right)$$



(Model Simul Mat Sci Eng 16, 085005, 2008)

This pair potential has no singularity and decays exponentially fast for  $r \gg \sigma$ .

If  $\epsilon \gg k_B T$ , then no two atoms will get too close to each other, capturing some of the physics of a real liquid.

The advantage of a purely repulsive model is that it has no liquid-gas transition. There is only one fluid phase.

I. Adiabatic Switching simulations can be performed to compute  $F_1 - F_0$

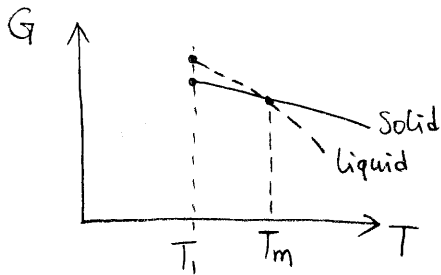
II. Another set of Adiabatic Switching Simulations is performed to compute  $F_0 - F_{i.g.}$

Because the pair potential  $\phi$  is very cheap to evaluate, the switching time from  $H_0$  to  $H_{i.g.}$  can be very long, making the errors very small.

$$\text{III. } F_1 = (F_1 - F_0) + (F_0 - F_{i.g.}) + F_{i.g.}$$

$\uparrow$                        $\uparrow$   
 error analysis need to be  
 performed for both sets of  
 switching simulations.

#### 4. Gibbs Free Energy as a Function of Temperature



For simplicity, assume pressure  $p=0$ .

The previous sections tell us how to compute  $G(N, p=0, T)$  for both solid and liquid phases at some  $T_1$ .

To find  $T_m$ , we need to know  $G$  as a function of  $T$  and see where do the two curves cross.

We can, of course, do that by repeating the above procedure at another set of temperatures,  $T_2, T_3, \dots$

But more efficient approaches exist.

We have seen in lecture 2 (p.5) that the change of free energy with respect to temperature can be obtained by monitoring the internal energy,

$$\frac{F(N, V, T_1)}{T_1} - \frac{F(N, V, T_0)}{T_0} = - \int_{T_0}^{T_1} \frac{E(N, V, T)}{T^2} dT$$

This suggests the possibility of "switching" along the temperature axis. The method is called "reversible scaling".

Ref: de Koning, Antonelli, Yip, Phys. Rev. Lett. 83, 3973 (1999).

In Reversible Scaling method, we consider a family of Hamiltonians

$$H(\{q_i, p_i, \lambda\}) = \sum_i \frac{p_i^2}{2m} + \lambda U_1(\{q_i\})$$

Note that at  $\lambda=1$ ,  $H(\lambda) = H_1$  is the system of interest

For  $0 < \lambda < 1$ ,  $H(\lambda)$  describes a system with "weakened" interatomic interaction.

The free energy of  $H(\lambda)$  at temperature  $T_1$  is connected to the free energy of the original system at temperature  $T_2 = \frac{T_1}{\lambda}$

$$F(N, V, T_1, \lambda) = -k_B T_1 \ln \left\{ \frac{1}{N! h^{3N}} \int d^{3N} q_i d^{3N} p_i \exp \left[ -\frac{1}{k_B T_1} \left( \sum_i \frac{p_i^2}{2m} + \lambda U_1(q_i) \right) \right] \right\}$$

$$= N k_B T_1 \ln(N \Lambda_1^3) - k_B T_1 \underbrace{\ln \left\{ \int d^{3N} q_i \exp \left[ -\frac{\lambda U_1(q_i)}{k_B T_1} \right] \right\}}_{\text{(same)}}$$

where  $\Lambda_1 = \frac{h}{\sqrt{2\pi m k_B T_1}}$

$$F_1(N, V, T_2) = N k_B T_2 \ln(N \Lambda_2^3) - k_B T_2 \underbrace{\ln \left\{ \int d^{3N} q_i \exp \left[ -\frac{U_1(q_i)}{k_B T_2} \right] \right\}}_{\text{(same)}}$$

where  $\Lambda_2 = \frac{h}{\sqrt{2\pi m k_B T_2}}$

$$\therefore \frac{F_1(N, V, T_2)}{T_2} - \frac{F(N, V, T_1, \lambda)}{T_1} = N k_B \ln \left( \frac{\Lambda_2^3}{\Lambda_1^3} \right) = \frac{3}{2} N k_B \ln \left( \frac{T_1}{T_2} \right)$$

$$F_1(N, V, T_2) = \frac{T_2}{T_1} F(N, V, T_1, \lambda) + \frac{3}{2} N k_B T_2 \ln \left( \frac{T_1}{T_2} \right)$$

note  $\lambda = \frac{T_1}{T_2}$

$$F_1 \left( N, V, \frac{T_1}{\lambda} \right) = \frac{1}{\lambda} F(N, V, T_1, \lambda) + \frac{3}{2} N k_B T_1 \frac{1}{\lambda} \ln \lambda$$

This means that, if we obtain  $F(N, V, T_1, \lambda)$  as a function of  $\lambda$  at a single temperature  $T_1$ , we can construct  $F_1$  as a function of temperature.

To compute  $F(N, V, T_1, \lambda)$ , we note

$$F(N, V, T_1, \lambda) = F(N, V, T_1, \lambda=1) + \int_1^\lambda \left\langle \frac{\partial \mathcal{H}(\lambda')}{\partial \lambda'} \right\rangle_{\lambda'} d\lambda'$$

$$= F_1(N, V, T_1) + \int_1^\lambda \left\langle U_1(q_i) \right\rangle_{\lambda'} d\lambda'$$

↑  
already computed

↑  
ensemble average  
of potential energy of unscaled  
Hamiltonian



During the switching simulation, we can also use a barostat to keep the pressure zero. Then the volume will depend on the parameter  $\lambda$ . We can show that

$$G(N, p=0, T_1, \lambda) = G_1(N, p=0, T_1) + \int_1^\lambda \langle U_1(q_i) \rangle_{\lambda'} d\lambda'$$

↑  
already computed

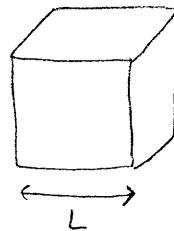
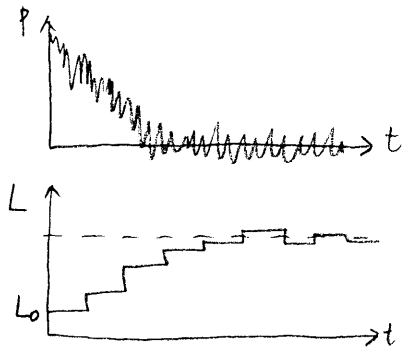
$$G_1(N, p=0, \frac{T_1}{\lambda}) = \frac{1}{\lambda} G_1(N, p=0, T_1) + \frac{3}{2} N k_B T_1 \frac{1}{\lambda} \ln \lambda .$$

Today's goal,

1. discuss HW 1 — free energy calculation of Si crystal
2. introduce rare event problem

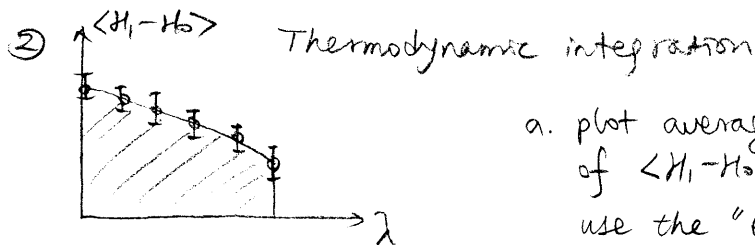
## 1. HW 1 compute free energy of Si crystal using Thermodynamic Integration

① Find equilibrium volume of Si crystal by a series of NVT MD siml.



$$V = L^3$$

- a. check if pressure is reduced to be around zero
- b. reduce fluctuation of box size  $L$
- c. each MD simulation should be long enough to get an accurate average of stress.



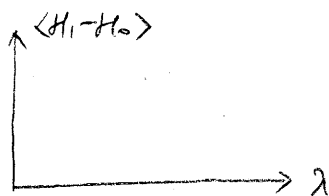
- a. plot average and standard deviation of  $\langle H_1 - H_0 \rangle$  as a function of  $\lambda$  use the "errorbar" command in Matlab.
- b. MD++ does not compute  $H_1 - H_0$  for  $\lambda = 0, 1$  use  $\lambda = 10^{-6}, 1 - 10^{-6}$  instead.

c. why  $\langle H_1 - H_0 \rangle > 0$  ?

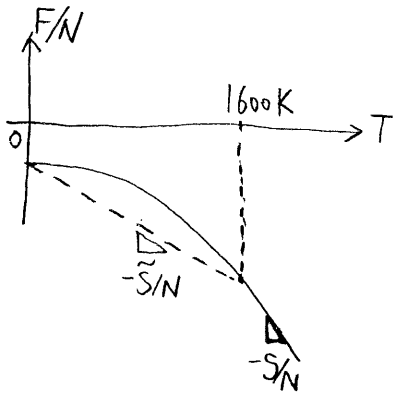
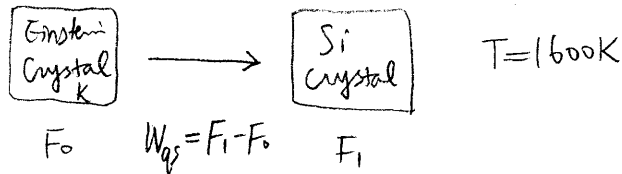
d. why  $\langle H_1 - H_0 \rangle$  decrease with  $\lambda$  ?

e. How would  $\langle H_1 - H_0 \rangle$  vs  $\lambda$  look like for "optimal" choice of  $K$  ?

( $K$  is spring constant in Einstein crystal)



③ How to gain confidence that the computed free energy is correct?  
 i.e. does the result make sense?



$1 k_B = 8.314 \text{ J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$   
 $= 8.617 \times 10^{-5} \text{ eV} \cdot \text{K}^{-1}$

a. repeat calculation for a different  $K$

b.  $S = - \left. \frac{\partial F}{\partial T} \right|_{N,V}$

$\therefore F(T=1600K) < F(T=0K)$

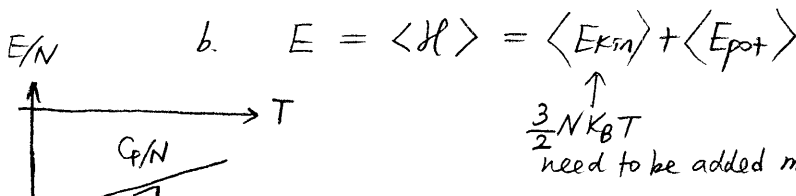
c. a crude estimate of entropy  
 $S = \frac{F(0K) - F(1600K)}{1600K}$

$\frac{S}{N} \sim 6 k_B$

d. compare with Fig 1 & Table 1 in MSMSE 16, 085005 (2008) "Coursework / Extra Reading"

④ Compute  $F(T)$

a. make sure that for each  $T$ , the box size  $L$  is equilibrated to give zero (average) pressure.



c. Check  $C_p = \left. \frac{\partial E}{\partial T} \right|_{p=0}$

$C_p \approx 3 N k_B T$  for Solids

d. verify that  $\frac{F(T_1)}{T_1} - \frac{F(T_0)}{T_0} = - \int_{T_0}^{T_1} \frac{E(T)}{T^2} dT$  for Einstein Crystal

For Einstein crystal:  $F(T_1) = E_0 - 3 N k_B T_1 \ln \frac{k_B T_1}{\hbar \omega}$ ,  $E(T) = E_0 + 3 N k_B T$   
 $F(T_0) = E_0 - 3 N k_B T_0 \ln \frac{k_B T_0}{\hbar \omega}$

$-\int_{T_0}^{T_1} \frac{E(T)}{T^2} dT = -\int_{T_0}^{T_1} \frac{E_0 + 3 N k_B T}{T^2} dT = E_0 \left( \frac{1}{T_1} - \frac{1}{T_0} \right) - 3 N k_B (\ln T_1 - \ln T_0)$   
 $= \frac{F(T_1)}{T_1} - \frac{F(T_0)}{T_0}$

## 2. Rare Event Problem

The timescale limitation is the greatest challenge of molecular simulations today.

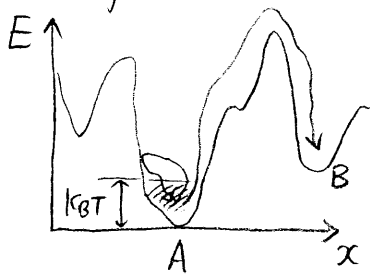
Typical time step of MD:  $\Delta t \approx 1 \text{ fs} = 10^{-15} \text{ s}$

1 million time steps  $\sim 10^6 \Delta t \sim 1 \text{ ns} = 10^{-9} \text{ s}$

	hard to reach:	$1 \mu\text{s} = 10^{-6} \text{ s}$
		$1 \text{ ms} = 10^{-3} \text{ s}$
		$1 \text{ s}$
macroscopic time scale of interest	} ↓ ? ↓	$1 \text{ day} = 86400 \text{ s} \sim 10^5 \text{ s}$
		$1 \text{ year} \sim 10^7 \text{ s}$

The gap in time scale is more than 10 orders of magnitude and is impossible to be bridged by brute force (faster computers) in the foreseeable future.

A typical reason for the appearance of this gap in time scale can be traced to the energy landscape of typical many-body systems, which consists of many local energy minima separated by energy barriers much higher than  $k_B T$ .



Every deep energy minimum corresponds to a meta-stable state.

The system spends a long time fluctuating within one meta-stable state, before a rare transition takes it to a neighboring meta-stable state.

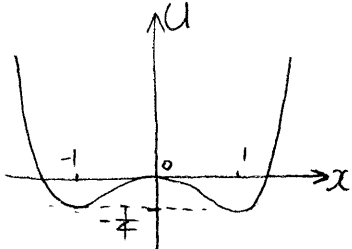
Direct MD or MC simulation is highly inefficient in predicting to long-time evolution of the system, because most of the computation time is "wasted" on sampling the "quasi-equilibrium" fluctuation within one meta-stable state.

Therefore, it is of considerable interest to compute the rate of rare transition events between metastable states.

Examples of rare transition events include:

- nucleation of crystals in supercooled liquid
- nucleation of dislocations in a perfect crystal under high stress
- diffusion of vacancies in crystals
- conformation change of proteins
- 
- 
- 

### A one-dimensional example



Consider a particle in a 1D potential field

$$U(x) = -\frac{1}{2}x^2 + \frac{1}{4}x^4$$

satisfying the following equation of motion:

$$\begin{cases} \dot{x}(t) = v(t) \\ m\dot{v}(t) = -m\gamma v(t) + R(t) + f(x(t)) \end{cases}$$

$\uparrow$  friction coefficient       $\uparrow$  random force       $\uparrow$  external force from potential field

$$f(x) = -\frac{\partial U(x)}{\partial x}$$



(If  $k_B T \ll \frac{1}{4}$ )

$$\langle R(t) \rangle = 0$$

(i.e. white noise)

$$\langle R(t'+t)R(t') \rangle = 2m k_B T \gamma \delta(t)$$

The physical meaning of this equation will be explained in the following.

The Langevin equation is an important model in statistical mechanics. For example, it can be used to describe Brownian motion (pollen in water).

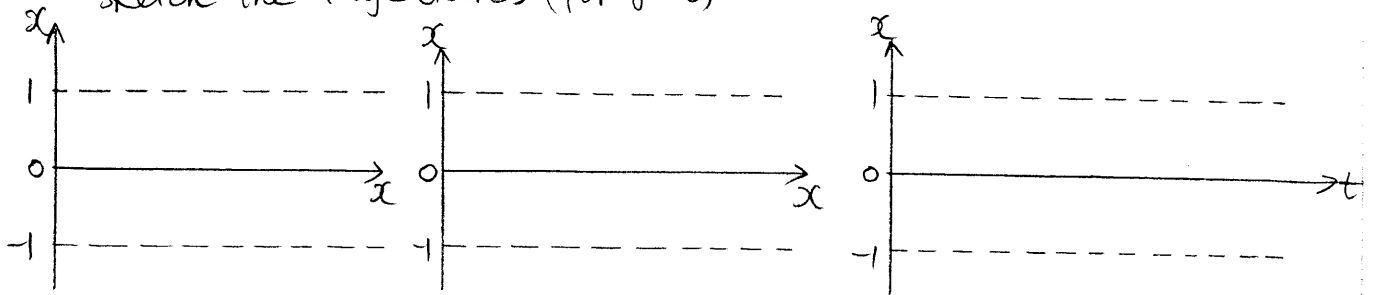
The frictional force  $-m\gamma v(t)$  and random force  $R(t)$  are essential in the Langevin equation.

If they are absent (i.e. set  $\gamma=0$ ), then the equation of motion becomes

$$\begin{cases} \dot{x}(t) = v(t) \\ m\dot{v}(t) = -\frac{\partial U(x)}{\partial x} \end{cases} \quad \text{i.e.} \quad m\ddot{x}(t) = -\frac{\partial U(x)}{\partial x}$$

Given  $U(x) = -\frac{1}{2}x^2 + \frac{1}{4}x^4$ , there are only three different types of trajectories, all being periodic in time.

Sketch the trajectories (for  $\gamma=0$ )



These trajectories are too simple and are not suitable for studying rare transition events.

This is because the system do not reach thermal equilibrium and hence does not have a temperature.

To fix this problem, we can couple the system with a thermostat. A natural thermostat is to embed the particle in a fluid. The fluid molecules (e.g. water molecules) bombard our particle (e.g. a pollen particle) incessantly. We shall refer to the particle as the Brownian particle, due to obvious connection to Brownian motion. From collisions with the water molecule, the Brownian particle acquires the temperature of the water, i.e.  $\langle mV^2 \rangle = k_B T$

Because water molecules ( $\sim 1\text{nm}$ ) are much smaller than the Brownian particle ( $\sim 1\mu\text{m}$ ), we shall consider the collisions with water molecules as random events.

The water molecules are not explicitly considered in the equation of motion (there are too many of them!).

The effects of the water molecules are represented by the two terms in the Langevin equation:

$$\text{frictional force: } f^{\text{fric}} = -m\gamma v(t)$$

$$\text{random force: } f^{\text{rand}} = R(t)$$

$$\langle R(t) \rangle = 0$$

$$\langle R(t+\tau) R(t) \rangle = 2m\gamma T \delta(\tau)$$

notice that the strength of the random force is proportional to the friction coefficient  $\gamma$  (times  $T$ )

This is a manifestation of the fluctuation-dissipation theorem.

The two forces are related because they are both the consequences of the "invisible" water molecules.

Through collisions, the Brownian particle acquires a thermal energy that can be used to jump over the barrier ( $E_b = \frac{1}{4}$ ) between two metastable states.

If  $k_B T \ll E_b$  (say  $E_b = 0.25\text{ eV}$ ,  $T = 300\text{ K}$ ,  $k_B T = 0.0259\text{ eV}$ )

then the transition across the barrier will be a rare event.

Today's goal:

Transition rate of 1D Langevin equation.

(Ref: Kubo Rep. Prog. Phys. 29, 255, 1966; Hänggi et al. Rev. Mod. Phys. 62, 251, 1990)

1. Physical Properties of the Langevin equation. Coursework - Extra Reading

$$\begin{cases} \dot{x}(t) = v(t) \\ m\dot{v}(t) = -m\gamma v(t) + R(t) + f(x(t)) \end{cases} \quad \text{----- (1)}$$

$\uparrow$  friction coefficient       $\uparrow$  random force

$\langle R(t) \rangle = 0$  ensemble average of random force is zero

$$\langle R(t+t') R(t') \rangle = 2m k_B T \gamma \delta(t) \quad \text{----- (2)}$$

$\therefore \langle R(t_1) R(t_2) \rangle = 0$  if  $t_1 \neq t_2$  random forces at two different times are not correlated.

$$\langle R^2(t) \rangle = \infty$$

$$\int_{-\infty}^{+\infty} \langle R(t+t') R(t') \rangle dt = 2m k_B T \gamma$$

consider the special case of  $f(x(t)) = 0$ ,

i.e. Brownian motion with no external potential.

$$m\dot{v}(t) = -m\gamma v(t) + R(t) \quad \text{--- no longer coupled to } x$$

$$\dot{v}(t) = -\gamma v(t) + \frac{1}{m} R(t) \quad \text{----- (3)}$$

This equation has the following analytic solution

$$v(t) = v(0) e^{-\gamma t} + e^{-\gamma t} \int_0^t e^{\gamma s} \frac{1}{m} R(s) ds \quad \text{----- (4)}$$

This can be verified by taking the time derivative,

$$\begin{aligned} \dot{v}(t) &= v(0) (-\gamma) e^{-\gamma t} + (-\gamma) e^{-\gamma t} \int_0^t e^{\gamma s} \frac{1}{m} R(s) ds \\ &\quad + e^{-\gamma t} e^{\gamma s} \frac{1}{m} R(s) \Big|_{s=t} \\ &= -\gamma v(t) + \frac{1}{m} R(t) \end{aligned}$$



(Reference: Hansen & McDonald, Theory of Simple Liquids, §7.3)

Squaring Eq.(4) and take ensemble average

$$\begin{aligned} \langle |v(t)|^2 \rangle &= \left\langle \left| v(0) e^{-\gamma t} + e^{-\gamma t} \int_0^t e^{\gamma s} \frac{1}{m} R(s) ds \right|^2 \right\rangle \\ &= \langle |v(0)|^2 \rangle e^{-2\gamma t} + 2 \int_0^t e^{\gamma s} \langle v(0) R(s) \rangle ds \cdot \frac{1}{m^2} e^{-2\gamma t} \\ &\quad + e^{-2\gamma t} \int_0^t ds \int_0^t ds' e^{\gamma(s+s')} \frac{1}{m^2} \langle R(s) R(s') \rangle ds ds' \end{aligned}$$

$\langle v(0) R(s) \rangle = 0$  : random force is uncorrelated with velocity

$$\langle R(s) R(s') \rangle = 2m k_B T \gamma \delta(s-s')$$

$$\begin{aligned} \therefore \langle |v(t)|^2 \rangle &= \langle |v(0)|^2 \rangle e^{-2\gamma t} + e^{-2\gamma t} \int_0^t ds \int_0^t ds' e^{\gamma(s+s')} \frac{2m k_B T \gamma \delta(s-s')}{m^2} ds ds' \\ &= \langle |v(0)|^2 \rangle e^{-2\gamma t} + e^{-2\gamma t} \int_0^t ds e^{2\gamma s} ds \cdot \frac{2k_B T \gamma}{m} \\ &= \langle |v(0)|^2 \rangle e^{-2\gamma t} + e^{-2\gamma t} \cdot \frac{1}{2\gamma} (e^{2\gamma t} - 1) \cdot \frac{2k_B T \gamma}{m} \\ &= \langle |v(0)|^2 \rangle e^{-2\gamma t} + (1 - e^{-2\gamma t}) \frac{k_B T}{m} \end{aligned}$$

$$\lim_{t \rightarrow \infty} \langle |v(t)|^2 \rangle = \frac{k_B T}{m}, \quad \text{i.e. } \lim_{t \rightarrow \infty} \langle \frac{1}{2} m |v(t)|^2 \rangle = \frac{1}{2} k_B T$$

as it should.

This justifies why the strength of the random force is

$$\langle R(t'+t) R(t') \rangle = 2m k_B T \gamma \delta(t)$$

The strength of the random force is related to the friction force, which is a manifestation of the fluctuation-dissipation theorem.

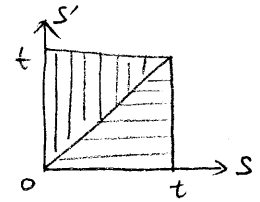
The friction coefficient is also related to the diffusion coefficient.  $D$

$$D = \lim_{t \rightarrow \infty} \frac{1}{2t} \langle |x(t) - x(0)|^2 \rangle$$

$$\text{note: } x(t) - x(0) = \int_0^t v(s) ds$$

$$D = \lim_{t \rightarrow \infty} \frac{1}{2t} \int_0^t ds \int_0^t ds' \langle v(s) v(s') \rangle$$

$$= \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t ds \int_0^{t-s} dt' \langle v(s) v(s+t') \rangle$$

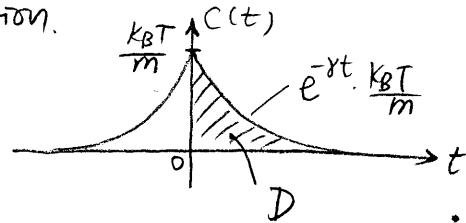


when the system has reached equilibrium  $\langle v(s) v(s+t') \rangle$  is independent of  $s$ .

$$D = \int_0^\infty \langle v(t_0) v(t_0+t') \rangle dt'$$

define  $C(t) = \langle v(t_0) v(t_0+t) \rangle$  as the velocity auto-correlation function.

$$\therefore D = \int_0^\infty C(t') dt'$$



$$C(0) = \langle v(t_0) v(t_0) \rangle = \frac{k_B T}{m}$$

$D$  equals the integral of the velocity auto correlation function.

Using Eq (4), assuming  $t_0$  is very large,

$$\begin{aligned} C(t) &= \langle v(t_0) v(t_0+t) \rangle = \left\langle e^{-\gamma t_0} \int_0^{t_0} e^{\gamma s} \frac{1}{m} R(s) ds \cdot e^{-\gamma(t_0+t)} \int_0^{t_0+t} e^{\gamma s'} \frac{1}{m} R(s') ds' \right\rangle \\ (\text{assume } t > 0) &= e^{-2\gamma t_0} e^{-\gamma t} \int_0^{t_0} ds \int_0^{t_0+t} ds' \frac{1}{m^2} \langle R(s) R(s') \rangle e^{\gamma(s+s')} \\ &= e^{-2\gamma t_0} e^{-\gamma t} \int_0^{t_0} ds \frac{1}{m^2} 2m k_B T \gamma e^{2\gamma s} \\ &= e^{-2\gamma t_0} e^{-\gamma t} \frac{2k_B T \gamma}{m} \frac{1}{2\gamma} (e^{2\gamma t_0} - 1) \\ &= \frac{k_B T}{m} e^{-\gamma t} (1 - e^{-2\gamma t_0}) \\ &= \frac{k_B T}{m} e^{-\gamma t} \quad (t_0 \text{ is very large}) \quad , \quad C(t) = \frac{k_B T}{m} e^{-\gamma |t|} \quad (t > 0) \end{aligned}$$

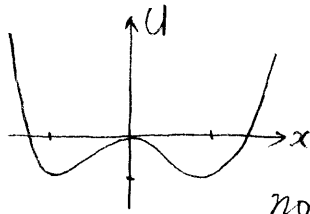
$$D = \int_0^{\infty} \langle c(t') \rangle dt' = \int_0^{\infty} \frac{k_B T}{m} e^{-\gamma t'} dt' = \frac{k_B T}{m\gamma}$$

$$\therefore D = \frac{k_B T}{m\gamma}, \quad m\gamma = \frac{k_B T}{D}$$

This is the Einstein relation, which is also a form of fluctuation-dissipation theorem.

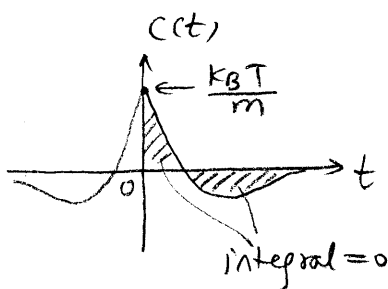
## 2. Langevin equation with confining potential

$$\begin{cases} \dot{x}(t) = v(t) \\ m\dot{v}(t) = -m\gamma v(t) + R(t) - \frac{\partial U(x)}{\partial x} \end{cases}$$



$$U(x) = -\frac{1}{2}x^2 + \frac{1}{4}x^4$$

note: the application of the confining potential will change the velocity auto-correlation function. In particular,  $\langle c(t) \rangle$  must be negative for some  $t$ , so that



$$D = \int_0^{\infty} \langle c(t') \rangle dt' = 0$$

i.e. no long distance diffusion is possible (why?)

Analytic solution is no longer possible, so we will use numerical integration.

Due to the presence of random variable  $R(t)$ , we are integrating a stochastic differential equation (SDE) and special care must be taken.

(Ref: Lawrence C. Evans, "An Introduction to Stochastic Differential Equations", version 1.2.

Department of Mathematics, UC Berkeley.

<http://math.berkeley.edu/~evans/SDE.course.pdf> )

For simplicity we will use the Euler forward integrator.

$$x(t+\Delta t) = x(t) + v(t) \cdot \Delta t$$

We need to be more careful with the rapidly fluctuating random force, i.e. we cannot assume  $R(t)$  remains constant during  $[t, t+\Delta t]$

$$v(t+\Delta t) = v(t) - \gamma v(t) \Delta t + \frac{1}{m} \int_t^{t+\Delta t} R(t') dt' - \frac{1}{m} \frac{\partial U(x(t))}{\partial x} \Delta t$$

$$Q \equiv \int_t^{t+\Delta t} R(t') dt' = \int_0^{\Delta t} R(t') dt' \text{ is a random number}$$

It is the total momentum transferred from the water molecule to the Brownian particle during a time interval  $\Delta t$ .

Because  $Q$  is the sum of a large number of collisions, it has a Gaussian distribution due to the central limit theorem (see ME346A lecture note Chap 3)

We just need to find its expectation value and standard deviation to completely specify the probability distribution of  $Q$ .

$$\langle Q \rangle = \int_0^{\Delta t} \langle R(t') \rangle dt' = 0$$

$$\text{Var}(Q) = \langle Q^2 \rangle - \langle Q \rangle^2 = \langle Q^2 \rangle$$

$$= \int_0^{\Delta t} dt_1 \int_0^{\Delta t} dt_2 \langle R(t_1) R(t_2) \rangle$$

$$= \int_0^{\Delta t} dt_1 \int_0^{\Delta t} dt_2 2m k_B T \gamma \delta(t_1 - t_2)$$

$$= \int_0^{\Delta t} dt_1 2m k_B T \gamma$$

$$= 2m k_B T \gamma \cdot \Delta t$$

$$\sigma_Q = \sqrt{2m k_B T \gamma \cdot \Delta t}$$

Let  $\xi$  be a random variable satisfying the normal distribution

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad \langle \xi \rangle = 0, \quad \langle \xi^2 \rangle = 1$$

Then  $Q = \sqrt{2m k_B T \gamma \cdot \Delta t} \cdot \xi$

In summary, the Euler forward integration looks like the following:

$$\begin{cases} x(t+\Delta t) = x(t) + v(t) \Delta t \\ v(t+\Delta t) = (1 - \gamma \Delta t) v(t) + \sqrt{\frac{2k_B T \gamma}{m}} \cdot \xi \cdot \sqrt{\Delta t} - \frac{1}{m} \frac{\partial U(x(t))}{\partial x} \cdot \Delta t \end{cases}$$

Notice that the term corresponding to the random force contain a random number  $\xi$  (generated at every time step) and is proportional to  $\sqrt{\Delta t}$ .

The  $\sqrt{\Delta t}$  term is a hallmark of SDE.

When  $U(x) = -\frac{1}{2}x^2 + \frac{1}{4}x^4$ , the force is

$$-\frac{\partial U}{\partial x} = x - x^3$$

Numerical integration by Euler Forward method is implemented in `langevin_1d.m` (Coursework/Materials/Homeworks).

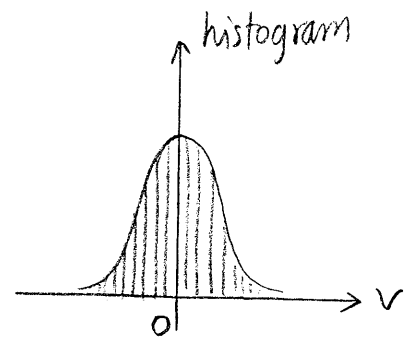
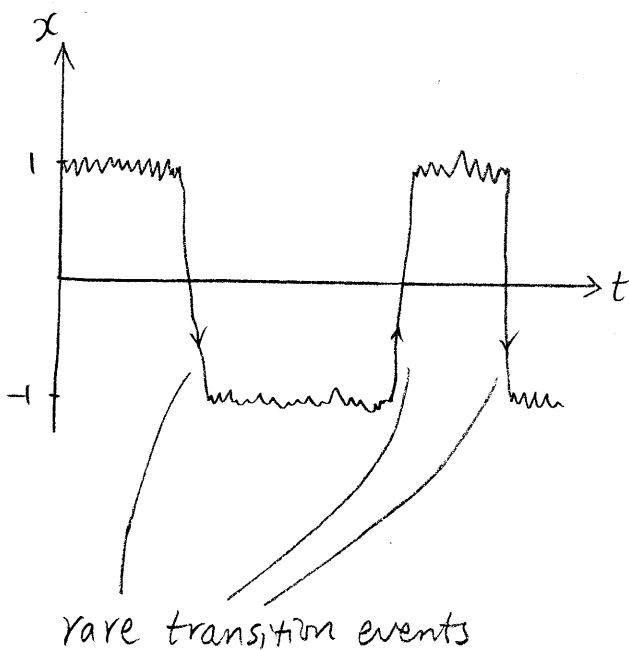
Exercise: run the simulation at several temperatures (in K) and observe the behavior of the system, e.g.

$$T = 1200; \quad \text{langevin\_1d}$$

$$T = 800; \quad \text{langevin\_1d}$$

$$T = 600; \quad \text{langevin\_1d}$$

Typical simulation behavior



velocity satisfies Maxwell Boltzmann distribution

$$f(v) = \frac{1}{\sqrt{2\pi}\sigma_v} e^{-\frac{v^2}{2\sigma_v^2}}$$

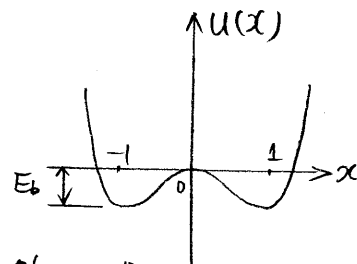
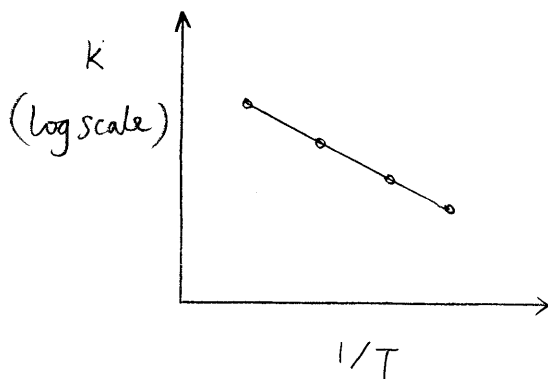
$$\sigma_v = \sqrt{\frac{k_B T}{m}} \quad \text{Standard deviation}$$

transition rate can be estimated as

$$k = \frac{\# \text{ of transition events}}{\text{total simulation time}}$$

← (can be extracted by analyze\_traj\_1d.m)

Arrhenius plot



Empirical observation:

$$k = \nu \exp\left(-\frac{E_b}{k_B T}\right)$$

(HW2)

Q. How to predict frequency prefactor  $\nu$ ?

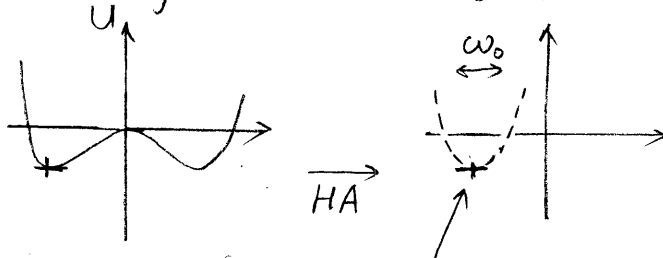
### 3. Transition State Theory (TST)

The transition state theory together with harmonic approximation (HA) predicts that the transition rate is

$$k_{\text{(HA)}}^{\text{TST}} = \frac{\omega_0}{2\pi} \exp\left(-\frac{E_b}{k_B T}\right)$$

where  $\omega_0 = \sqrt{\frac{K}{m}}$ ,  $K = \left. \frac{\partial^2 U}{\partial x^2} \right|_{x=-1}$

is the angular frequency of vibration close to the bottom of the meta-stable state



$$U(x) \approx U(-1) + \frac{1}{2}K(x+1)^2$$

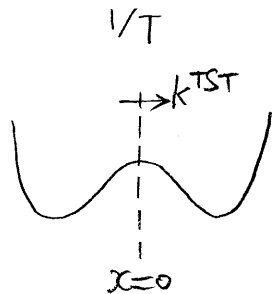
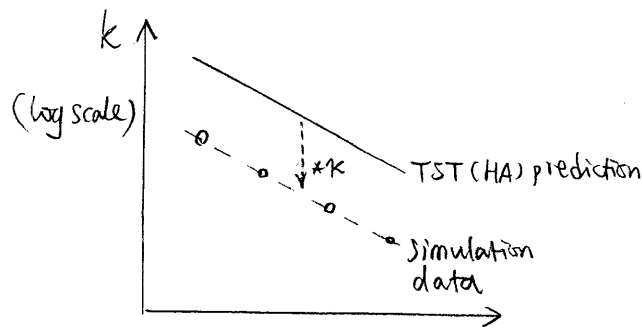
$K$ : effective spring constant

We will derive this result in the next lecture.

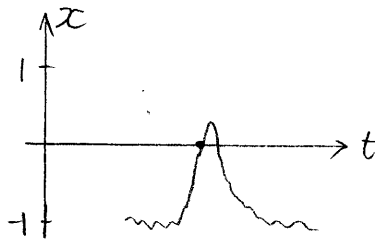
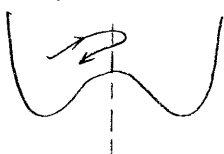
Here we observe that this expression makes physical sense.

- The Boltzmann factor  $\exp\left(-\frac{E_b}{k_B T}\right)$  describes the probability that the system can gain enough thermal energy and reach the neighborhood of  $x=0$  (the dividing surface)
- $\frac{\omega_0}{2\pi}$  describes the frequency at which the system attempts to cross the dividing surface ( $x=0$ )
- The Boltzmann factor  $\exp\left(-\frac{E_b}{k_B T}\right)$  can be regarded as a success probability for each attempt.

## 4. Recrossing Factor $K$



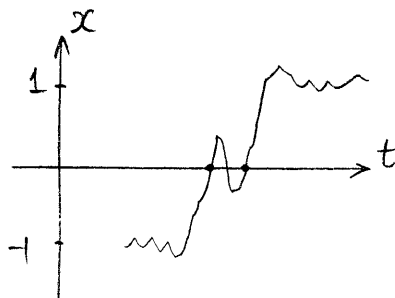
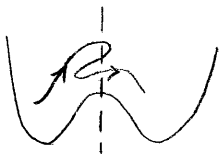
Examples:



# of crossing:  $N_{\text{cross}} = 1$

# of transition:  $N_{\text{trans}} = 0$

(from left to right only)



# of crossing:  $N_{\text{cross}} = 2$

# of transition:  $N_{\text{trans}} = 1$

(from left to right only)

A recrossing correction factor  $K$  can be computed by running a set of short simulations starting from  $x=0$  and  $v > 0$  (satisfying Boltzmann distrib)

Each simulation terminates when  $|x(t)| > 1$

A successful transition is observed if  $x(t_{\text{end}}) > 1$   
(instead of  $x(t_{\text{end}}) < -1$ ).

$$K = \frac{\text{total \# of transitions}}{\text{total \# of crossings}}$$

(see HW 2, langwin-recross\_1d.m)  
on Coursework

TST is found to always over-estimate the transition rate.

The main reason is that TST predicts the flux across the dividing surface ( $x=0$ ).

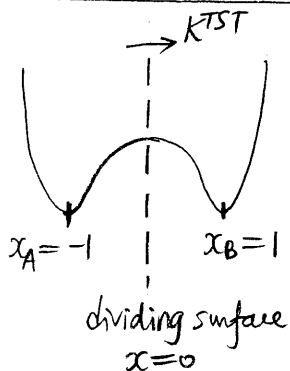
But not all trajectories crossing the dividing surface contribute to a transition event.



Today's goal:

## Transition State Theory in 1D

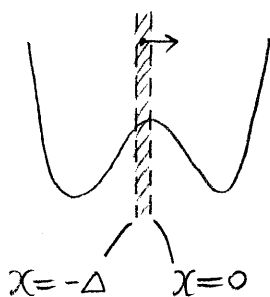
### 1. Flux across dividing surface ( $x=0$ )



Consider a large ensemble of replicas of the system at thermal equilibrium and described by the canonical (NVT) ensemble.

At any instant, there are some systems crossing the dividing surface ( $x=0$ ) from the left (i.e. from  $x=0^-$  to  $x=0^+$ ).

$k^{TST}$  is the probability flux (per unit time) of these crossings.



Consider a small domain to the left of  $x=0$ ,  $[-\Delta, 0)$  ( $\Delta \ll 1$ )

In the canonical ensemble, the probability of finding the system in this domain out of all systems in the left metastable state is

$$p(-\Delta, 0) = \frac{\int_{-\Delta}^0 e^{-\beta U(x)} dx}{\int_{-\infty}^0 e^{-\beta U(x)} dx} \approx \frac{e^{-\beta E_b} \cdot \Delta}{\int_{-\infty}^0 e^{\beta[U(x) - U(-1)]} dx}$$

Assuming that the velocity for these systems still satisfies the Maxwell Boltzmann distribution (i.e. Gaussian distribution with standard deviation  $\sigma_v = \sqrt{\frac{k_B T}{m}}$ ,  $f(v) = \frac{1}{\sqrt{2\pi}\sigma_v} \exp(-\frac{v^2}{2\sigma_v^2})$ )

Half of these systems have  $v > 0$ , and have the potential to cross  $x=0$  a moment later.

Consider a short time period  $\delta t$ . We will compute the fraction of systems that crosses  $x=0$  during this short time interval.

The average velocity moving to the right is

$$\bar{v} = 2 \cdot \int_{-\infty}^0 f(v) \cdot v \, dv = 2 \cdot \frac{1}{\sqrt{2\pi} \sigma_v} \int_{-\infty}^0 v \cdot e^{-\frac{v^2}{2\sigma_v^2}} \, dv = \sqrt{\frac{2}{\pi}} \sigma_v$$

Therefore, on the average, all systems within the domain of  $[-\Delta, 0)$  where  $\Delta = \bar{v} \cdot \delta t = \sqrt{\frac{2}{\pi}} \sigma_v \cdot \delta t$  and with a positive velocity will cross the dividing surface after a period of  $\delta t$ .

The probability of finding such systems is

$$P_{\text{cross}} = P([- \Delta, 0]) \cdot \frac{1}{2} = \frac{e^{-\beta E_b} \cdot \Delta}{\int_{-\infty}^0 e^{-\beta [u(x) - u(x_A)]} \, dx} \cdot \frac{1}{2}$$

$\uparrow$   
 only half of  
 the systems have  
 positive velocity

$x_A = -1$

define  $\tilde{Z}_A = \int_{-\infty}^0 e^{-\beta [u(x) - u(x_A)]} \, dx$ , which is proportional to the partition function of the left metastable state (state A)

$$P_{\text{cross}} = \frac{e^{-\beta E_b}}{\tilde{Z}_A} \cdot \sqrt{\frac{2}{\pi}} \cdot \sqrt{\frac{k_B T}{m}} \cdot \frac{1}{2} \cdot \delta t$$

The flux across the dividing surface is

$$k^{\text{TST}} = \frac{P_{\text{cross}}}{\delta t} = \sqrt{\frac{k_B T}{2\pi m}} \frac{e^{-\beta E_b}}{\tilde{Z}_A}$$

Eyring (1935)

The partition function of state A is

$$\begin{aligned} Z_A &= \frac{1}{h} \int_{-\infty}^{+\infty} d(mv) \int_{-\infty}^0 dx \exp[-\beta(\frac{1}{2}mv^2 + U(x) - U(x_A))] \\ &= \frac{\sqrt{2\pi m k_B T}}{h} \int_{-\infty}^0 dx \exp[-\beta(U(x) - U(x_A))] \\ &= \frac{\sqrt{2\pi m k_B T}}{h} \tilde{Z}_A \end{aligned}$$

$$\begin{aligned} k^{TST} &= \sqrt{\frac{k_B T}{2\pi m}} \frac{e^{-\beta E_b}}{Z_A} \cdot \frac{\sqrt{2\pi m k_B T}}{h} \\ &= \frac{k_B T}{h} \frac{e^{-\beta E_b}}{Z_A} \end{aligned}$$

$h$ : Planck's constant

(note  $k^{TST}$  is purely classical, hence independent of  $h$ .  
The  $h$  factor in the above expression cancels that in the definition of  $Z_A$ .)

Note: both  $e^{-\beta E_b}$  and  $Z_A$  are dimensionless quantities

$\frac{k_B T}{h}$  has the dimension of frequency, so is  $k^{TST}$ .

## 2. Harmonic Approximation in TST

In general  $Z_A$  (or  $\tilde{Z}_A$ ) is not available analytically.

However, if we expand  $U(x)$  around the local minimum  $x_A$ ,

$$U(x) \approx U(x_A) + \frac{1}{2}K(x-x_A)^2 + \underbrace{\mathcal{O}(x-x_A)^3}_{\text{ignored}}, \quad K = \left. \frac{\partial^2 U}{\partial x^2} \right|_{x_A}$$

$$\text{Then } \tilde{Z}_A \approx \int_{-\infty}^{+\infty} dx \exp(-\frac{1}{2}\beta K(x-x_A)^2) = \sqrt{\frac{2\pi k_B T}{K}}$$

$$Z_A = \frac{\sqrt{2\pi m k_B T}}{h} \cdot \sqrt{\frac{2\pi k_B T}{K}} = \frac{2\pi k_B T}{h} \sqrt{\frac{m}{K}} = \frac{2\pi k_B T}{h \omega_0} = \frac{k_B T}{\hbar \omega_0}$$

define  $\omega_0 \equiv \sqrt{\frac{K}{m}}$  angular frequency of Harmonic oscillator

$$\hbar \equiv \frac{h}{2\pi}$$

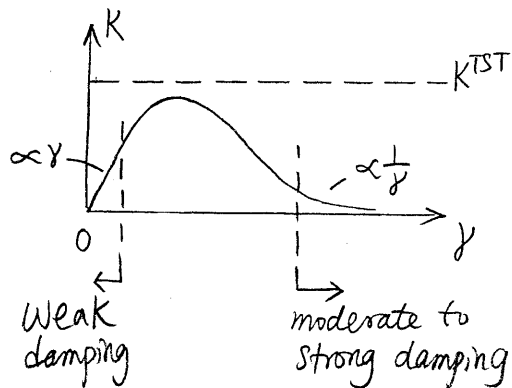
$$k_{(HA)}^{TST} = \frac{k_B T}{h} \cdot \frac{\hbar \omega_0}{2\pi k_B T} e^{-\beta E_b} = \frac{\omega_0}{2\pi} e^{-\beta E_b} \quad (\text{same as in lecture 6, p.8})$$

### 3. Recrossing Factor $\kappa$

We now know that  $k^{TST}$  overestimates the true transition rate  $k$ .

$$k = \kappa \cdot k^{TST}$$

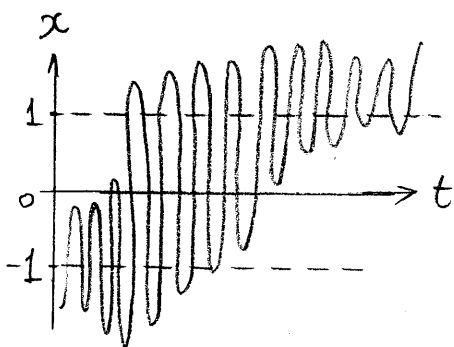
$\kappa < 1$  is introduced to account for the recrossing events.



[Hanggi et al (1990), Fig. 18]

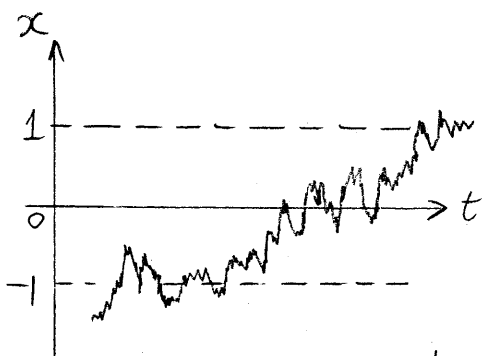
The transition rate of the 1D Langevin equation depends on the friction coefficient  $\gamma$ , while  $k^{TST}$  is independent of  $\gamma$ .

$\kappa \ll k^{TST}$  (i.e.  $\kappa \ll 1$ ) in both the weak damping ( $\gamma \rightarrow 0$ ) and strong damping ( $\gamma \rightarrow \infty$ ) limit.



#### a. Weak damping ( $\gamma \rightarrow 0$ ) limit

Once the system has acquired enough energy to cross the barrier, it will go back and forth between the two states many times before it loses enough energy to settle inside one of the two states.



#### b. Strong damping ( $\gamma \rightarrow \infty$ ) limit

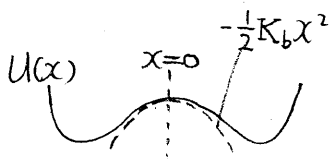
The inertia of the system is not important here and the trajectory resembles that of diffusion.

This is called the Smoluchowski limit.

$$k = \frac{\omega_0 \omega_b}{2\pi \gamma} e^{-\beta E_b}, \quad (\text{i.e. } \kappa = \frac{\omega_b}{\gamma})$$

(\* can be derived from Becker-Döring Theory)

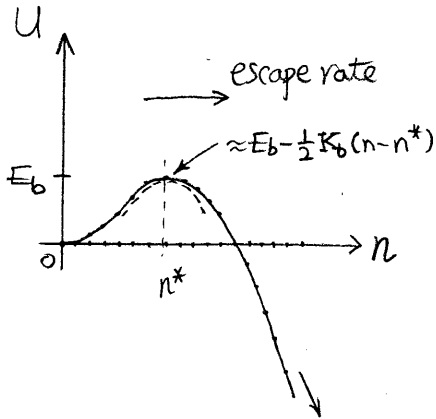
$$\text{where } \omega_b = \sqrt{\frac{K_b}{m}}, \quad K_b = -\left. \frac{\partial^2 U}{\partial x^2} \right|_{x=0}$$



#### 4. Rate in the Spatial Diffusion Limit

commonly known as the Becker-Döring Theory

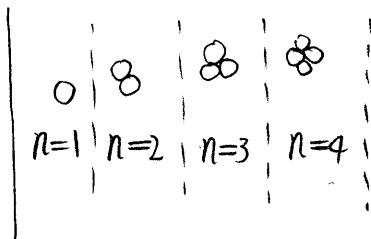
Farkas (1927), Becker-Döring (1935)



For a change, let us now consider a discrete system:  $n=0, 1, 2, \dots$

The system has only one meta-stable state and we are interested in the rate of escape from this meta-stable state.

The discrete system may be considered as an approximation of the Langevin equation in the strong damping limit.



It may also be a good model by itself, e.g.  $x$  can be the number of molecules in a droplet in a nucleation event.

The dynamics along the  $x$ -axis is a random walk, i.e. we have a Markov process.

probability rate of moving to  $(n+1)$  from state  $n$ :  $R_+(n)$

probability rate of moving to  $(n-1)$  from state  $n$ :  $R_-(n)$

detailed balance condition:



$$R_+(n-1) \cdot e^{-\beta U(n-1)} = R_-(n) \cdot e^{-\beta U(n)}$$

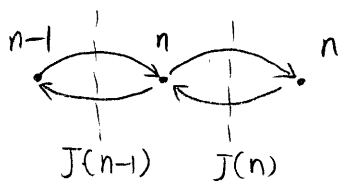
$$\text{i.e. } \frac{R_+(n-1)}{R_-(n)} = e^{-\beta [U(n) - U(n-1)]}$$

$$R_-(n) = R_+(n-1) \cdot e^{\beta [U(n) - U(n-1)]}$$

$$\text{e.g. } R_-(1) = R_+(0) e^{\beta (U_1 - U_0)}, \quad U_1 = U(1), U_0 = U(0)$$

Let  $p(n,t)$  be the probability of finding the system at state  $n$ .  
(corresponding to the density of droplet of size  $n$ )

Let  $J(n,t)$  be the net flux from state  $n$  to state  $n+1$



$$\begin{cases} \frac{\partial}{\partial t} p(n,t) = J(n-1,t) - J(n,t) \\ J(n,t) = p(n,t) \cdot R_+(n) - p(n+1,t) \cdot R_-(n+1) \end{cases}$$

Seek steady state solution (no time dependence) where  $J(n) = J^{st}$  for all  $n$ .

$$J^{st} = J(n) = p(n) R_+(n) - p(n+1) R_-(n+1)$$

$$\begin{aligned} n=0: \quad J^{st} &= p(0) R_+(0) - p(1) R_-(1) = R_+(0) [p(0) - p(1) e^{\beta(u_1 - u_0)}] \\ &= e^{-\beta u_0} \cdot R_+(0) \cdot [p(0) e^{\beta u_0} - p(1) e^{\beta u_1}] \end{aligned}$$

$$\begin{aligned} n=1: \quad J^{st} &= p(1) R_+(1) - p(2) R_-(2) \\ &= e^{-\beta u_1} \cdot R_+(1) \cdot [p(1) e^{\beta u_1} - p(2) e^{\beta u_2}] \end{aligned}$$

$$n=N-1: \quad J^{st} = e^{-\beta u_{N-1}} R_+(N-1) \cdot [p(N-1) e^{\beta u_{N-1}} - p(N) e^{\beta u_N}]$$

$$\frac{J^{st}}{e^{-\beta u_0} \cdot R_+(0)} = p(0) e^{\beta u_0} - p(1) e^{\beta u_1}$$

$$\frac{J^{st}}{e^{-\beta u_1} \cdot R_+(1)} = p(1) e^{\beta u_1} - p(2) e^{\beta u_2}$$

⋮

$$\frac{J^{st}}{e^{-\beta u_{N-1}} \cdot R_+(N-1)} = p(N-1) e^{\beta u_{N-1}} - p(N) e^{\beta u_N}$$

add all together

$$\left( \sum_{n=0}^{N-1} \frac{1}{e^{-\beta u_n} \cdot R_+(n)} \right) \cdot J^{st} = p(0) e^{\beta u_0} - p(N) e^{\beta u_N}$$

Boundary condition:  $p(0) = \frac{e^{-\beta U_0}}{Z}$ ,  $\hat{Z} = \sum_{n=0}^{n^*} e^{-\beta U_n}$  (assume local equilibrium)

$$\lim_{N \rightarrow \infty} p(N) = 0 \quad (n^* \text{ is a maximum of } U(n))$$

$$J^{st} = \frac{1}{Z} \cdot \left( \sum_{n=0}^{\infty} \frac{e^{\beta U_n}}{R_+(n)} \right)^{-1} = \left( \sum_{n=0}^{\infty} \frac{1}{\frac{1}{Z} e^{-\beta U_n} R_+(n)} \right)^{-1}$$

This is the famous Becker-Döring result.

It can be simplified if  $U$  varies slowly with  $n$ .

The dominant contribution to the sum comes from the terms near  $n = n^*$ .

Taylor expand  $U(n)$  around  $n^*$ , and assume  $R_+(n) \approx R_+(n^*)$

$$U(n) = E_b - \frac{1}{2} K_b (n - n^*)^2 + \dots$$

$$\frac{1}{J^{st}} \approx \int_{-\infty}^{+\infty} dn \frac{1}{\frac{1}{Z} e^{-\beta E_b} \cdot e^{\frac{1}{2} \beta K_b (n - n^*)^2} R_+(n^*)}$$

$$= \frac{1}{\frac{1}{Z} e^{-\beta E_b} R_+(n^*)} \int_{-\infty}^{+\infty} dn e^{-\frac{1}{2} \beta K_b (n - n^*)^2}$$

$$= \frac{1}{\frac{1}{Z} e^{-\beta E_b} R_+(n^*)} \sqrt{\frac{2\pi K_b T}{K_b}}$$

$$J^{st} = \frac{1}{Z} e^{-\beta E_b} R_+(n^*) \cdot \frac{K_b}{\sqrt{2\pi K_b T}} \quad \left( \Gamma \equiv \frac{K_b}{\sqrt{2\pi K_b T}} \text{ is called the Zeldovich factor} \right)$$

If  $U(n)$  near  $n=0$  is also approximated by Harmonic Approx.

then  $\hat{Z} = \sqrt{\frac{2\pi K_b T}{K}}$ , where  $K = \frac{\partial^2 U}{\partial n^2} \Big|_{n=0}$  (see p. 3)

$$J^{st} = \sqrt{\frac{K}{2\pi K_b T}} \cdot \sqrt{\frac{K_b}{2\pi K_b T}} \cdot e^{-\beta E_b} R_+(n^*)$$

$$= \frac{\sqrt{K \cdot K_b}}{2\pi K_b T} \cdot e^{-\beta E_b} R_+(n^*)$$

To make connection with the TST result for Langevin equation,  
we note that  $R_+(n^*)$  is related to diffusivity along the  $n$ -axis

$$\left( D = \frac{a^2}{2\tau} \text{ for random walk, see ME346A Lecture Note, Chap 1} \right)$$

$$D = \frac{1}{2} [R_+(n^*) + R_-(n^*)] = R_+(n^*) \quad (\text{here } a=1)$$

At the same time, Einstein's relation requires

$$D = \frac{k_B T}{m \gamma}$$

$$\therefore R_+(n^*) = \frac{k_B T}{m \gamma}$$

$$j_{st} = \frac{\sqrt{K \cdot K_b}}{2\pi k_B T} e^{-\beta E_b} \cdot \frac{k_B T}{m \gamma}$$

$$\text{define } \omega_0 = \sqrt{\frac{K}{m}}, \quad \omega_b = \sqrt{\frac{K_b}{m}}$$

$$j_{st} = \frac{\omega_0 \omega_b}{2\pi \gamma} e^{-\beta E_b}$$

(identical to the result on  
p.4 based on TST with  $K$  correction)



Today's goal:

## Transition State Theory in 2D, 3N-D

### 1. Langevin equation in 2D

Consider a 2D potential energy function (landscape)

$$U(x, y) = \frac{1}{6} \left\{ 4(1-x^2-y^2)^2 + 2(x^2-2)^2 + [(x+y)^2-1]^2 + [(x-y)^2-1]^2 - 2 \right\}$$

It has two local minima:  $\underline{r}_A = (-\frac{\sqrt{5}}{2}, 0)$ ,  $\underline{r}_B = (\frac{\sqrt{5}}{2}, 0)$

118

#### FINDING TRANSITION PATHWAYS

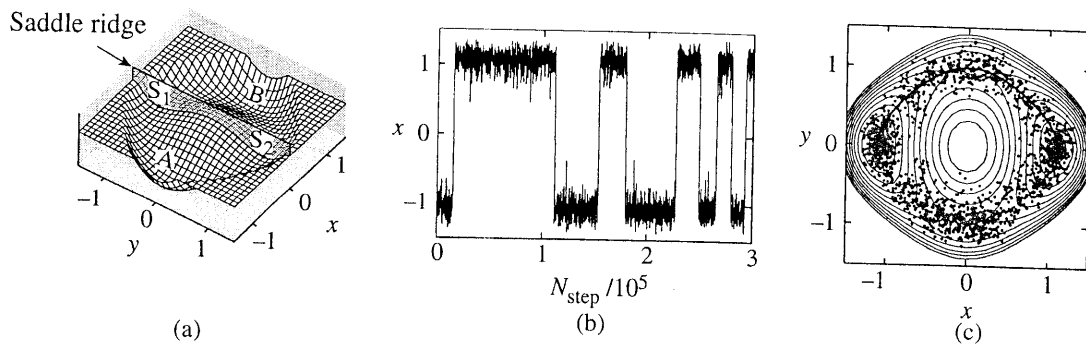


FIG. 7.1. (a) A two-dimensional potential energy function  $U_1(x, y)$ . The two metastable states, A for  $x < 0$  and B for  $x > 0$ , are separated by a ridge at  $x = 0$ .  $S_1$  and  $S_2$  mark the locations of two saddles on the ridge. (b) Instantaneous values of coordinate  $x$  in an MC simulation at  $k_B T = 0.15$ . (c) The microstates (dots) visited during transitions between A and B plotted on top of the contour lines of  $U(x, y)$ . The solid (dashed) line connecting two energy minima is the minimum energy path obtained by steepest descent from saddle  $S_1$  ( $S_2$ ).

The dividing surface (or saddle ridge),  $x=0$ , divides the 2D space into two regions: A and B.

- A: ( $x < 0$ ) starting from any point here and following a steepest descent path will always lead to point  $\underline{r}_A$
- B: ( $x > 0$ ) starting from any point here and following a steepest descent path will always lead to point  $\underline{r}_B$

$$U(\underline{r}_A) = U(\underline{r}_B) = -\frac{1}{12}$$

Assume the system follows the Langevin equation:

$$\begin{cases} \dot{x}(t) = v_x(t) \\ \dot{y}(t) = v_y(t) \\ m\dot{v}_x(t) = -m\gamma v_x(t) + R_1(t) - \frac{\partial U(x,y)}{\partial x} \\ m\dot{v}_y(t) = -m\gamma v_y(t) + R_2(t) - \frac{\partial U(x,y)}{\partial y} \end{cases}$$

$$\langle R_i(t) \rangle = 0$$

$$\langle R_i(t) R_j(t) \rangle = 2m k_B T \gamma \delta(t) \delta_{ij} \quad i, j = 1 \text{ or } 2$$

At sufficiently low  $T$ , the system spends most of the time around  $x_A$  or  $x_B$  (i.e. in state A or state B), with rare transitions between the two states (energy basins).

Numerical integration by Euler Forward method is implemented in `Langevin_2d.m` (Coursework/Materials/Homeworks)

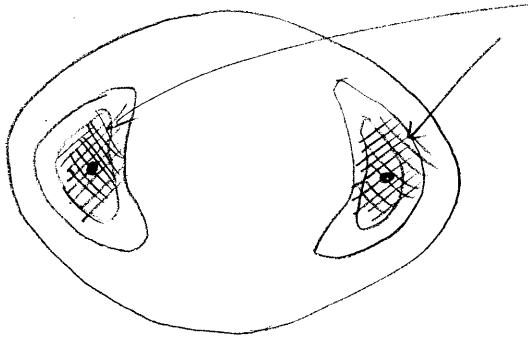
Exercise: run the 2d simulation at different temperatures (in K) and observe the behavior of the system, e.g. •

$$T = 2500; \quad \text{Langevin\_2d}$$

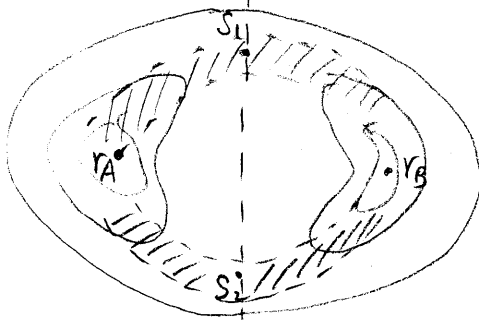
$$T = 2000; \quad \text{Langevin\_2d}$$

$$T = 1500; \quad \text{Langevin\_2d}$$

2. Minimum Energy Path (MEP)

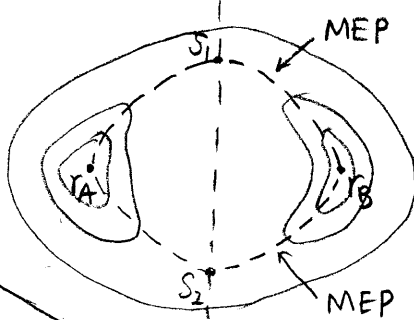


System spends most of time in the energy basin around  $\bar{x}_A, \bar{x}_B$



When rare transition occurs, the trajectories path through a "tubular region" that includes special point  $S_1$  or  $S_2$ .

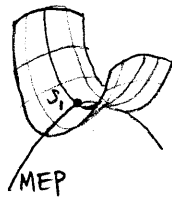
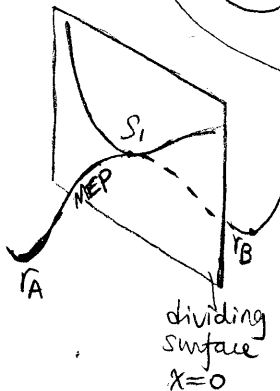
$x=0$   
dividing surface



$S_1$  and  $S_2$  are saddle points, which are potential energy minima within the dividing surface ( $x=0$ ).

$S_1 = (0, 1), S_2 = (0, -1), U(S_1) = U(S_2) = 1$

Starting from  $S_1$  (or  $S_2$ ) plus a small perturbation and following the steepest descent trajectory will lead us to an energy minimum ( $\bar{x}_A$  or  $\bar{x}_B$ ).



A minimum energy path (MEP) connecting  $\bar{x}_A$  and  $\bar{x}_B$  can be constructed from the two steepest descent trajectories from each saddle point ( $S_1$  or  $S_2$ ).

Saddle points are energy maxima along MEPs, but are energy minima within the dividing surface. The shape of the energy landscape around  $S_1$  or  $S_2$  looks like a saddle.

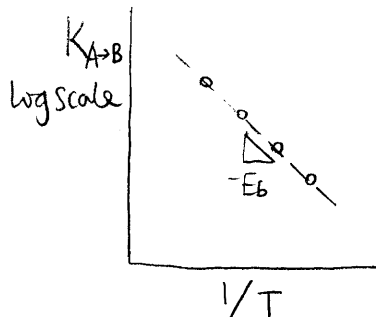
In the limit of  $T \rightarrow 0$ , every transition path must closely follow an MEP.

3. Transition Rate in 2D

Suppose we have already identified the saddle points ( $S_1, S_2$ ) connecting two energy minima ( $r_A, r_B$ )

(which is not an easy task, especially in high dimensions)

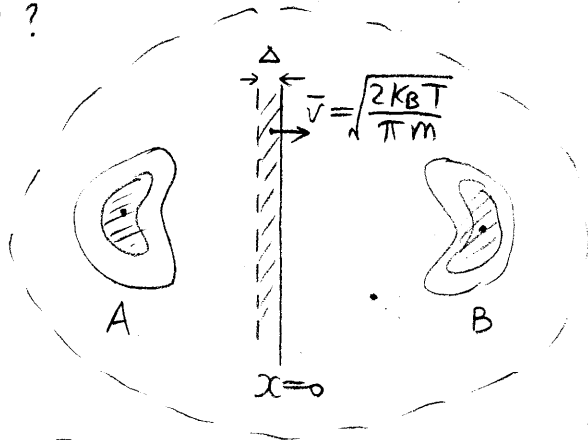
how to predict the transition rate (e.g. from A to B)?



$$K_{A \rightarrow B} = \nu \exp\left(-\frac{E_b}{k_B T}\right)$$

$$E_b = U(S_1) - U(r_A) = 1 + \frac{1}{12} = \frac{13}{12}$$

Q: How to predict frequency prefactor  $\nu$ ?



$$K = \kappa K^{TST}$$

↑  
recrossing factor  
 $0 < \kappa < 1$

↑  
transition state theory prediction

(see lecture note 7, p.2)

$$K^{TST} = \frac{\int_{-\infty}^{+\infty} dy e^{-\beta[U(x=0, y) - U(r_A)]}}{\int_{-\infty}^0 dx \int_{-\infty}^{+\infty} dy e^{-\beta[U(x, y) - U(r_A)]}} \cdot \bar{v} \cdot \frac{1}{2}$$

probability density at dividing surface  $x=0$

↑  
average velocity to the right

↑  
only half of the systems in the ensemble have positive velocity

define partition function of state A:

$$Z_A = \frac{1}{h^2} \int_{-\infty}^{+\infty} dp_x \int_{-\infty}^{+\infty} dp_y \int_{-\infty}^0 dx \int_{-\infty}^{+\infty} dy \exp\left[-\beta\left(\frac{p_x^2}{2m} + \frac{p_y^2}{2m} + U(x, y) - U(r_A)\right)\right]$$

$$= \left(\frac{\sqrt{2\pi m k_B T}}{h}\right)^2 \int_{-\infty}^0 dx \int_{-\infty}^{+\infty} dy e^{-\beta[U(x, y) - U(r_A)]}$$

define partition function of dividing surface ( $x=0$ )

$$\begin{aligned} Z^* &= \frac{1}{h} \int_{-\infty}^{+\infty} dp_y \int_{-\infty}^{+\infty} dy \exp \left[ -\beta \left( \frac{p_y^2}{2m} + U(x=0, y) - U(\underline{r}_A) \right) \right] \\ &= \frac{\sqrt{2\pi m k_B T}}{h} \int_{-\infty}^{+\infty} dy e^{-\beta [U(x=0, y) - U(\underline{r}_A)]} \end{aligned}$$

$$k^{TST} = \frac{\sqrt{2\pi m k_B T}}{h} \cdot \frac{Z^*}{Z_A} \cdot \sqrt{\frac{2k_B T}{\pi m}} \cdot \frac{1}{2}$$

$$\boxed{k^{TST} = \frac{k_B T}{h} \cdot \frac{Z^*}{Z_A}}$$

compare with the 1D result:  $k^{TST} = \frac{k_B T}{h} \cdot \frac{e^{-\beta E_b}}{Z_A}$

$e^{-\beta E_b}$  is replaced by  $Z^*$ , because each transition trajectory does not have to exactly pass through the saddle point for  $T > 0$ .

#### 4. Harmonic Approximation in TST.

Taylor expand  $U(x, y)$  around  $\underline{r}_A = (x_A, y_A)$ :  $x_A = -\frac{\sqrt{5}}{2}$ ,  $y_A = 0$

$$U(x, y) = U(\underline{r}_A) + \frac{1}{2} K_1 (x - x_A)^2 + \frac{1}{2} K_2 (y - y_A)^2 + \dots$$

ignored

$$Z_A = \left( \frac{\sqrt{2\pi m k_B T}}{h} \right)^2 \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy e^{-\beta \left( \frac{1}{2} K_1 (x - x_A)^2 + \frac{1}{2} K_2 (y - y_A)^2 \right)}$$

$$= \left( \frac{\sqrt{2\pi m k_B T}}{h} \right)^2 \cdot \sqrt{\frac{2\pi k_B T}{K_1}} \cdot \sqrt{\frac{2\pi k_B T}{K_2}}$$

$$= \left( \frac{2\pi k_B T}{h} \right)^2 \cdot \sqrt{\frac{m}{K_1}} \cdot \sqrt{\frac{m}{K_2}}$$

define  $\omega_1 = \sqrt{\frac{K_1}{m}}$ ,  $\omega_2 = \sqrt{\frac{K_2}{m}}$ ,  $\hbar \equiv \frac{h}{2\pi}$

$$Z_A = \frac{k_B T}{\hbar \omega_1} \cdot \frac{k_B T}{\hbar \omega_2}$$

(Partition function of two independent harmonic oscillators)

Taylor expand  $U(x, y)$  around  $S_1 = (x_b, y_b) : x_b = 0, y_b = 1$

$$U(x, y) = U(S_1) - \frac{1}{2} K_b (x - x_b)^2 + \frac{1}{2} K_2^* (y - y_b)^2 + \underbrace{\dots}_{\text{ignored}}$$

$$Z^* = \frac{\sqrt{2\pi m k_B T}}{h} \int_{-\infty}^{+\infty} e^{-\beta(\frac{1}{2} K_2^* (y - y_b)^2 + E_b)} dy$$

$E_b \equiv U(S_1) - U(A)$

$$= \frac{\sqrt{2\pi m k_B T}}{h} \cdot \sqrt{\frac{2\pi k_B T}{K_2^*}} \cdot e^{-\beta E_b} \quad \text{define } \omega_2^* = \sqrt{\frac{K_2^*}{m}}$$

$$= \frac{k_B T}{h \omega_2^*} \cdot e^{-\beta E_b}$$

$$K_{\text{(HA)}}^{\text{TST}} = \frac{k_B T}{h} \cdot \frac{\frac{k_B T}{h \omega_2^*}}{\frac{k_B T}{h \omega_1} \cdot \frac{k_B T}{h \omega_2}} \cdot e^{-\beta E_b}$$

$$\boxed{K_{\text{(HA)}}^{\text{TST}} = \frac{\omega_1 \omega_2}{2\pi \omega_2^*} \cdot e^{-\beta E_b}}$$

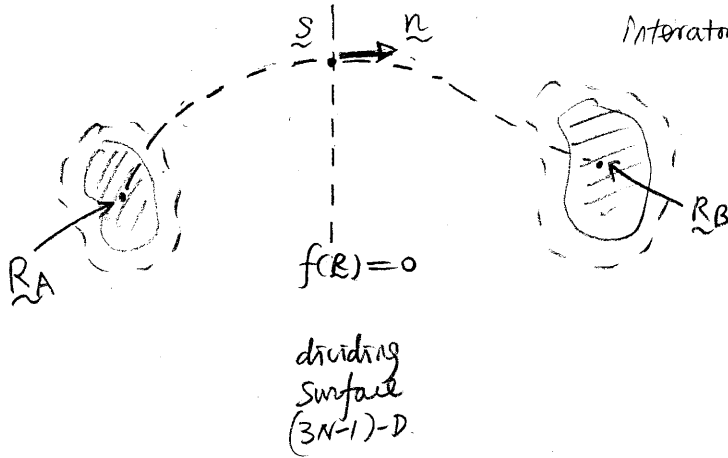
compare with 1D result:  $K_{\text{(HA)}}^{\text{TST}} = \frac{\omega_b}{2\pi} e^{-\beta E_b}$

$$K = \kappa \cdot K^{\text{TST}}$$

In the strong friction limit,  $\kappa = \frac{\omega_b}{\gamma}$ ,  $\omega_b = \sqrt{\frac{m}{K_b}}$

then  $K_{\text{(HA)}} = \frac{\omega_1 \omega_2}{2\pi \omega_2^*} \cdot \frac{\omega_b}{\gamma} \cdot e^{-\beta E_b}$

5. Generalization to 3N-D



Interatomic potential function

$$U(\{q_i\}) \rightarrow U(\underline{R})$$

↑  
3N-D vector

$\underline{R}_A, \underline{R}_B$  local energy minima.  
 $\underline{S}$ : saddle point.

dividing surface defined as

$$f(q_1, q_2, \dots, q_{3N}) = 0$$

State A:  $f(\underline{R}) < 0$   
State B:  $f(\underline{R}) > 0$

It is usually approximated by a plane

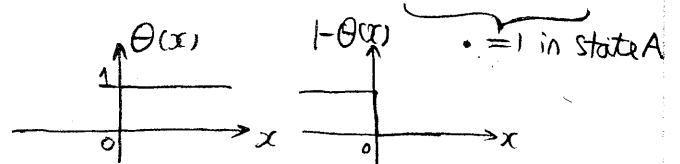
$$(\underline{R} - \underline{S}) \cdot \underline{n} = 0 \quad \underline{n} \text{ is a unit vector}$$

i.e.  $f(q_1, q_2, \dots, q_{3N}) = \sum_i (q_i - n_i) - (\underline{S} \cdot \underline{n})$

$$K^{TST} = \frac{k_B T}{h} \frac{Z^*}{Z_A}$$

$$d^{3N} \underline{R} = dq_1 dq_2 \dots dq_{3N}$$

where  $Z_A = \left( \frac{\sqrt{2\pi m k_B T}}{h} \right)^{3N} \int d^{3N} \underline{R} e^{-\beta[U(\underline{R}) - U(\underline{R}_A)]} \cdot (1 - \Theta(f(\underline{R})))$



$$Z^* = \left( \frac{\sqrt{2\pi m k_B T}}{h} \right)^{3N-1} \int d^{3N} \underline{R} e^{-\beta[U(\underline{R}) - U(\underline{R}_A)]} \cdot \delta(f(\underline{R}))$$

Harmonic approximation:

$$Z_A = \frac{k_B T}{\hbar \omega_1} \cdot \frac{k_B T}{\hbar \omega_2} \dots \frac{k_B T}{\hbar \omega_{3N}}$$

$$Z^* = \frac{k_B T}{\hbar \omega_2^*} \dots \frac{k_B T}{\hbar \omega_{3N}^*} \cdot e^{-\beta E_b}$$

$$E_b = U(\underline{S}) - U(\underline{R}_A)$$

$$K^{TST}_{(HA)} = \frac{\omega_1 \cdot \omega_2 \dots \omega_{3N}}{2\pi \cdot \omega_2^* \dots \omega_{3N}^*} e^{-\beta E_b}$$

$\omega_1, \omega_2, \dots, \omega_{3N}$  are eigenfrequencies near  $\underline{R}_A$ .

$\omega_2^*, \dots, \omega_{3N}^*$  are eigenfrequencies in the dir. surf. near  $\underline{S}$ .

$K = \kappa \cdot K^{TST}$ , In the strong friction limit,  $\kappa = \frac{\omega_b}{\gamma}$ ,  $\omega_b = \sqrt{\frac{m}{Bz_b}}$

$$K_{(HA)} = \frac{\omega_1 \cdot \omega_2 \dots \omega_{3N}}{2\pi \cdot \omega_2^* \dots \omega_{3N}^*} \cdot \frac{\omega_b}{\gamma} \cdot e^{-\beta E_b}$$

Suppose we want to express the transition rate in the following form:

$$k = \nu_0 \exp\left(-\frac{F_b}{k_B T}\right), \text{ where } \nu_0 \text{ is a constant, say } \nu_0 = 10^{13} \text{ s}^{-1}$$

where  $F_b = E_b - T S_b$  is the free energy barrier  
or activation free energy,

$E_b$  is the activation energy

$S_b$  is the activation entropy

$$k = \nu_0 \exp\left(-\frac{E_b - T S_b}{k_B T}\right) = \nu_0 \exp\left(\frac{S_b}{k_B}\right) e^{-\beta E_b}$$

Compare this with the TST predictions. we can see that

the Harmonic Approximation of TST is equivalent to assuming

$$\nu_0 \exp\left(\frac{S_b}{k_B}\right) = \frac{\omega_1 \omega_2 \dots \omega_{3N}}{2\pi \omega_2^* \dots \omega_{3N}^*}$$

$$\text{i.e. } S_b = k_B \ln\left(\frac{\omega_1 \omega_2 \dots \omega_{3N}}{2\pi \nu_0 \omega_2^* \dots \omega_{3N}^*}\right)$$

within HA, activation entropy measures the change of local eigen frequencies from  $R_A$  to  $S$ .

If we use  $\kappa = \frac{\omega_b}{\gamma}$  in the strong friction limit, we have

$$\nu_0 \exp\left(\frac{S_b}{k_B}\right) = \frac{\omega_1 \omega_2 \dots \omega_{3N}}{2\pi \omega_2^* \dots \omega_{3N}^*} \cdot \frac{\omega_b}{\gamma}$$

$$\text{i.e. } S_b = k_B \ln\left(\frac{\omega_1 \omega_2 \dots \omega_{3N}}{2\pi \nu_0 \omega_2^* \dots \omega_{3N}^*} \cdot \frac{\omega_b}{\gamma}\right)$$



Today's goal:

Finding Minimum Energy Path (MEP) in 2D, 3ND.

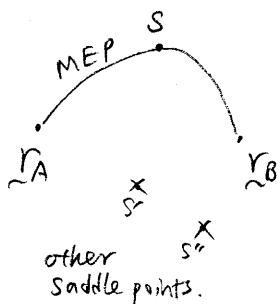
(Ref: Computer Simulations of Dislocations, Bulatov & Cai, 2006, §7)

We have seen that MEP and saddle point plays an important role in determining the transition rate

In fact, the MEP is more important than the saddle point

The maximum energy point on MEP must be a saddle point, as long as the potential energy is a smooth function.

However, a Saddle point may or may not connect the two local energy minima of interest through steepest descent paths.



In high dimension, there are a lot of "irrelevant" saddle points.

In this lecture, we will introduce several methods of searching MEPS.

The energy barrier of a path is defined as the maximum potential energy along the path minus the potential energy at the beginning point of the path.

Searching for MEP means finding a path  $\underline{\Gamma}(s)$  that minimizes the energy barrier.

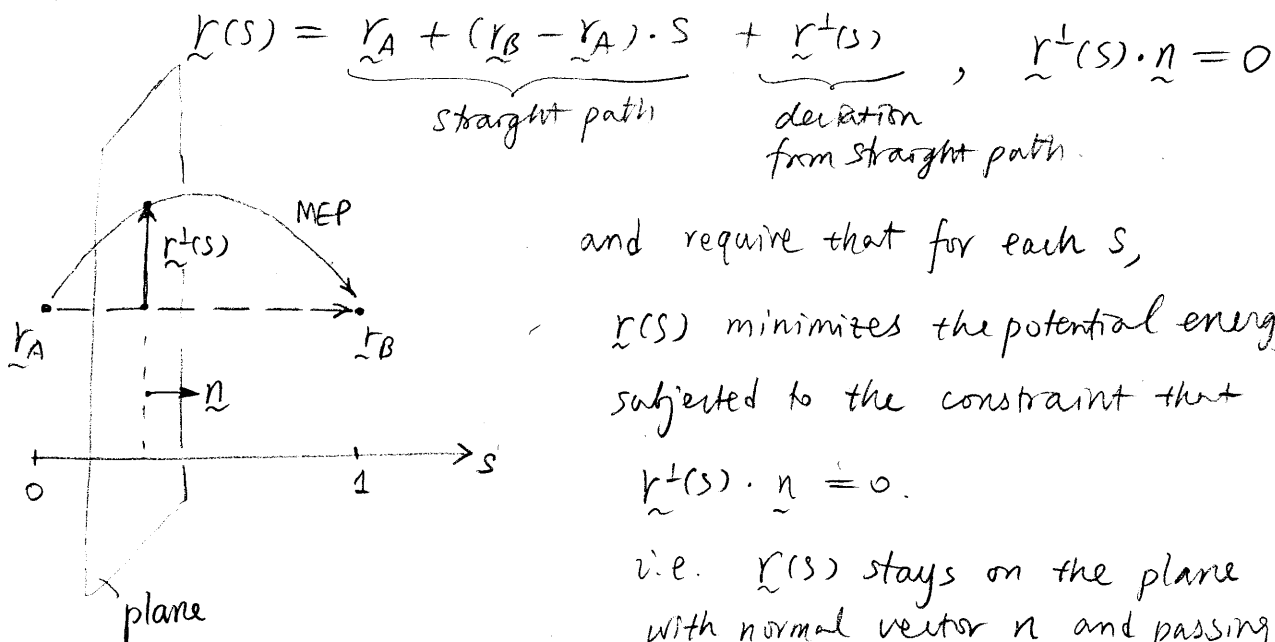
Most likely we will obtain a local minimum in the space of all paths. It is usually impossible to prove that the search result is a global minimum, although there are ways to improve the probability that we find the global minimum.

## 1. Constrained Minimization

A straight path connecting  $\underline{r}_A$  and  $\underline{r}_B$  can be written as

$$\underline{r}(s) = \underline{r}_A + (\underline{r}_B - \underline{r}_A) \cdot s$$

If we assume the MEP is not so different from the straight path, we can try



and require that for each  $s$ ,

$\underline{r}(s)$  minimizes the potential energy subjected to the constraint that

$$\underline{r}^\perp(s) \cdot \underline{n} = 0.$$

i.e.  $\underline{r}(s)$  stays on the plane with normal vector  $\underline{n}$  and passing through point  $\underline{r}_A + (\underline{r}_B - \underline{r}_A) \cdot s$ .

$$(\underline{r} - \underline{r}_A) \cdot \underline{n} = (\underline{r}_B - \underline{r}_A) \cdot \underline{n} \cdot s$$

$$\underline{n} = \frac{\underline{r}_B - \underline{r}_A}{|\underline{r}_B - \underline{r}_A|}$$

Starting from  $s=0$ ,  $\underline{r}(s=0) = \underline{r}_A$ ,  $s$  can be increased

in small steps  $\Delta s$ . The configuration from the previous step

is advanced to the new plane by adding

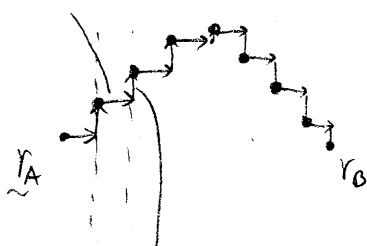
$$(\underline{r}_B - \underline{r}_A) \cdot \Delta s,$$

and is then relaxed to a local energy minimum within the plane.

Repeating the process will lead to a

path that may approximate MEP.

$$(\underline{r}_B - \underline{r}_A) \cdot \Delta s$$



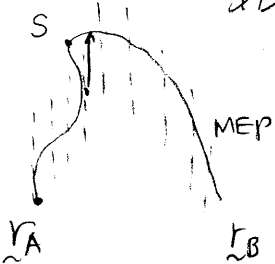
Constrained  
Minimization

Advantage of constrained minimization:

- Easy to implement
- computationally cheap.

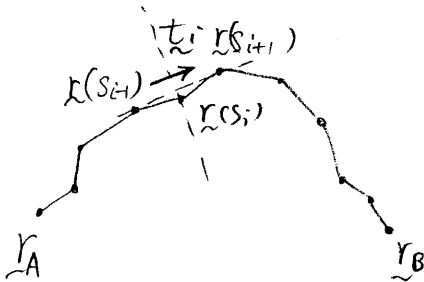
Disadvantages:

- can fail if true MEP is very different from straight path.
- Path may not be continuous, may not reach  $\underline{r}_B$ !



## 2. Chain of State Method

Commonly under the name of Nudged-Elastic band method or string method.



Initialize a chain of configurations  
(i.e. many copies of the system)

$$\underline{r}(S_i), \quad i=0, 1, \dots, N, \quad S_i = \frac{i}{N}$$

$$\underline{r}(0) = \underline{r}_A; \quad \underline{r}(1) = \underline{r}_B.$$

A linear path is commonly used as an initial guess

$$\underline{r}^{\text{init}}(S_i) = \underline{r}_A + (\underline{r}_B - \underline{r}_A) \cdot S_i$$

But it is not necessary, if a better guess is available.

define local tangent vector

$$\underline{t}_i = \frac{\underline{r}(S_{i+1}) - \underline{r}(S_{i-1})}{|\underline{r}(S_{i+1}) - \underline{r}(S_{i-1})|}$$

We now search for path such that  $\underline{r}(S_i)$  is a minimum within the plane of normal vector  $\underline{t}_i$  for every  $i$ .

This can be achieved by moving  $\mathbf{x}(s_i)$  in the direction that reduces the potential energy but orthogonal to  $\underline{t}_i$  at every step, and iterate until convergence.

The points along the path needs to be "resampled" periodically to ensure good quality in the numerical representation i.e. points uniformly distributed along the path.



Advantage of chain of state method:

- More robust than constrained minimization

Disadvantage:

- More time consuming, need parallel computing for large systems.
- Step size need to be adjusted for stability.
- Initial path need to be guessed.

Result may be strongly influenced by initial path.

— i.e. local minimum versus global minimum.

### 3. Global Path Optimization

We can use Monte Carlo / Simulated Annealing simulation in the space of paths to increase the probability of finding the global minimum

For each path (discretized as  $\underline{r}(s_i)$ ), define

$$E_{\max} = \max_i E(\underline{r}(s_i))$$

At each MC Step, randomly perturb one of the states along the chain, subjected to the constrain that

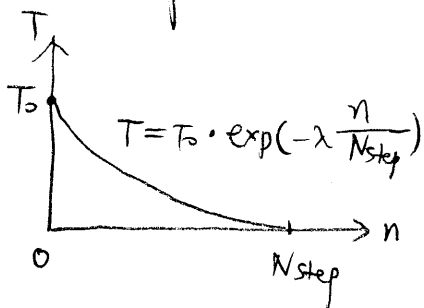
$$\begin{aligned} |r^{\text{new}}(s_i) - r^{\text{new}}(s_{i-1})| &< r_c \\ |r^{\text{new}}(s_i) - r^{\text{new}}(s_{i+1})| &< r_c \end{aligned} \quad (\text{to ensure continuity})$$

and accept the trial if it decreases  $E_{\max}$ .

If the trial more increases  $E_{\max}$ , accept with probability

$$P_{\text{acc}} = \exp\left(-\frac{E_{\max}^{\text{new}} - E_{\max}}{k_B T}\right)$$

Temperature  $T$  gradually decreases during simulation



## Adatom Diffusion on Au Surface

Wei Cai

8/1/2012

Department of Mechanical Engineering, Stanford University, Stanford 94305-4040

### 1. Prepare Au (111) Surface

The atomistic calculations below will be based on the EAM Foiles potential for Au. The tcl file can be download from ME346C Coursework web site.

```
bin/eam_mc2 scripts/work/au/au_surface_atom.tcl 0
```

A perfect FCC crystal of Au is created with dimensions:  $4[1 -2 1] \times 4[1 0 -1] \times 3[1 1 1]$  (1152 atoms).

A 1.4% bi-axial compressive strain is applied in the x and y directions. (Otherwise the surface is under a large tension stress and adatoms will tend to be incorporated into the top surface instead of staying on top of it.)

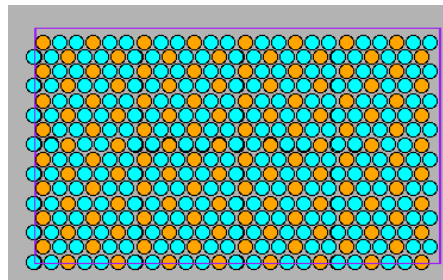


Figure 1. Au (111) surface. The top surface layer atoms are in orange.

Click on the atoms to get the (scaled) coordinates of various sites on the (111) surface. These coordinates will be used in step 2.

### 2. Add One Adatom to Au (111) Surface

```
bin/eam_mc2 scripts/work/au/au_surface_atom.tcl 1
```

One more atom is added on top of the (111) surface. The total number of atoms is 1153.

There are two different types of sites to put the adatom: type A and type B. We create four different structures: A, Ap, B, Bp. A and Ap are of the same type. B and Bp are of the same type. The structures are saved in folder runs/au-surf-1-0-0.

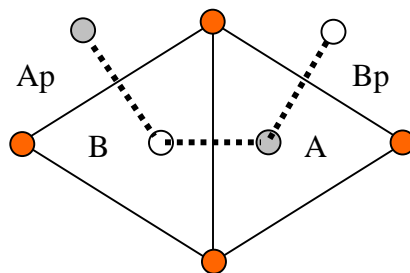


Figure 2. Orange atoms are the top layer of Au (111) surface. There are two types of sites (type A and type B) for adatom, both avoiding the orange atoms. A, Ap are equivalent sites. B, Bp are equivalent sites.

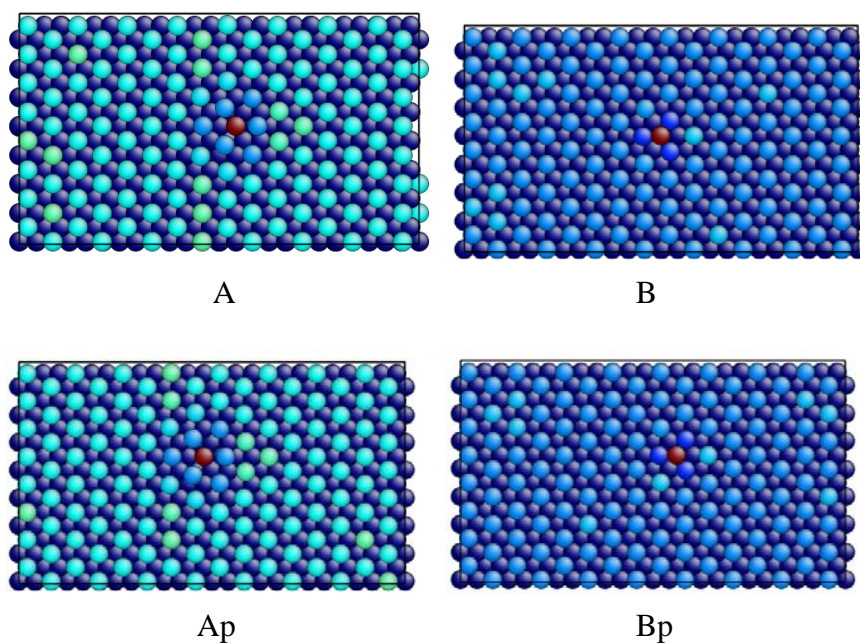


Figure 3. For different adatom structures created by MD++.

### 3. Find Minimum Energy Path

The string method is used to search for the minimum energy path between state A (or Ap) and state B (or Bp).

```
bin/eam_mc2 scripts/work/au/au_surface_atom.tcl 2 A B
or
bin/eam_mc2 scripts/work/au/au_surface_atom.tcl 2 A Bp
```

Figure 4 shows the energy profile for minimum energy path (after 1000 step of stringrelax) between different states.

Plot energy profile:

```
cd runs/au-surf-2-A-B
```

```
octave
```

```
data=load('stringeng.out'); plot(data(end-10:end,3:5:end)', '*-');  
xlabel('replica ID'); ylabel('Energy (eV)');
```

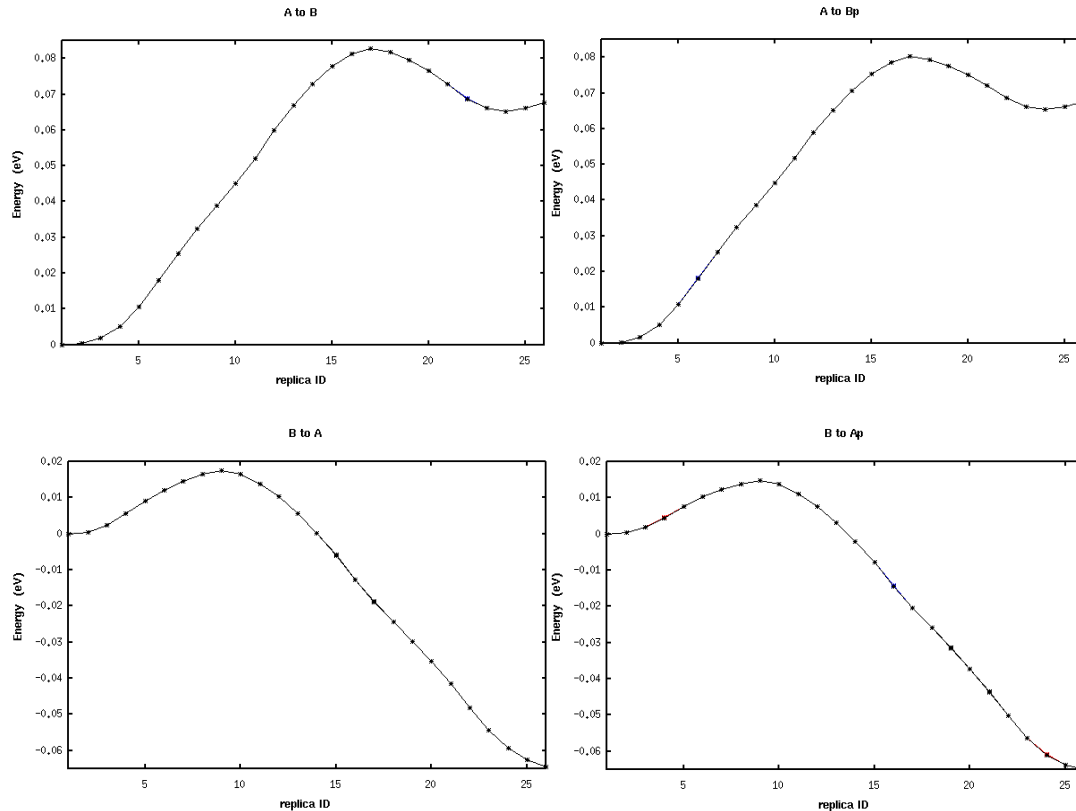


Figure 4. The energy profile for adatom migration on Au (111) surface.

From these calculations, we see that the energy difference between the two adatom sites is:

$$E_B - E_A = 0.065 \text{ eV}$$

The energy barriers for the jump are (approximately):

$$E^* - E_A = 0.083 \text{ eV}$$

$$E^* - E_B = 0.018 \text{ eV}$$

The calculation can be repeated with more relaxation steps to make sure the results have converged. It is also of interest to find out whether there are multiple transition pathways between meta-stable states (e.g. between A and B).

Further Reading:

R. C. Jaklevic and L. Elie, Scanning-Tunneling-Microscope Observation of Surface Diffusion on an Atomic Scale: Au on Au(111), Phys. Rev. Lett. 60, 120–123 (1988)

[http://prl.aps.org/pdf/PRL/v60/i2/p120\\_1](http://prl.aps.org/pdf/PRL/v60/i2/p120_1)



```

# -*-shell-script-*-
# compute energy barrier of surface atom diffusion in FCC Au
#
# make eam build=R SYS=mc2

source "scripts/Examples/Tcl/startup.tcl"

*****
# Definition of procedures
*****
proc initmd { status n m } {
#MD++ setnolog
MD++ setoverwrite
MD++ dirname = runs/au-surf- $\$status-\$n-\$m$ 
MD++ NNM = 200
}

#-----
proc readeam { material } {
#Read in MEAM potential (MD++ format)
MD++ potfile = "~/Codes/MD++/potentials/EAMDATA/eamdata. $\{material\}$ Foiles"
MD++ eamgrid = 500 readeam NNM=300
MD++ element0 =  $\$material$ 
}

proc make_perfect_crystal { nx ny nz } {
MD++ crystalstructure = face-centered-cubic latticeconst = 4.08
MD++ latticesize = \[ 1 -2 1  $\$nx$  1 0 -1  $\$ny$  1 1 1  $\$nz$  \]
MD++ makecrystal #finalcnfile = perf.cn writecn #eval
}

proc make_dislocation_loop { } {
# create structure B'
set bz [expr -0.5/ [MD++_Get latticesize(11)] ]
MD++ input = \[ 3 1 0 0  $\$bz$  0.02 -0.25 0.25 0.3 -10 10 -10 10 1 1 -0.25 0.25 0 \]
MD++ makedipole
}

#-----
proc relax_fixbox { } { MD++ {
# Conjugate-Gradient relaxation
conj_ftol = 1e-4 conj_itmax = 1000 conj_fevalmax = 1000
conj_fixbox = 1
relax
} }
#end of proc relax_fixbox

#-----
proc relax_freebox { } { MD++ {
# Conjugate-Gradient relaxation
conj_ftol = 1e-4 conj_itmax = 1000 conj_fevalmax = 1000
conj_fixbox = 0
conj_fixboxvec = [ 0 0 1
                  1 0 1
                  0 0 0 ]

relax
} }
#end of proc relax_freebox

proc setup_window { } { MD++ {
#-----
#colors for Central symmetry view
color00 = "red" color01 = "blue" color02 = "green"
color03 = "magenta" color04 = "cyan" color05 = "purple"
color06 = "gray80" color07 = "white" color08 = "orange"
#-----
# Plot Configuration
#
atomradius = [0.7 0.78] bondradius = 0.3 bondlength = 0 #2.8285 #for Si
win_width=600 win_height=600
#atomradius = 0.9 bondradius = 0.3 bondlength = 0 #2.8285 #for Si
atomcolor = cyan highlightcolor = purple bondcolor = red
} }

```

```

fixatomcolor = yellow backgroundcolor = gray70
#atomcolor = lightgrey highlightcolor = purple bondcolor = darkgrey
plot_color_axis = 2 NCS = 12
plot_color_windows = [ 5
                        0.6 10 1
                        10 20 6
                        20 80 8
                        80 5000 5
                        0 0.6 4
                      ]
#plot_limits = [ 1 -10 10 -10 0.35 -10 10 ]
#plot_color_windows = 0
plot_atom_info = 1 # reduced coordinates of atoms
#plot_atom_info = 2 # real coordinates of atoms
#plot_atom_info = 3 # energy of atoms
#plot_highlight = [ 0 0 1 2 3 4 5 6 7 8 9 ]
plotfreq = 10
#
#rotateangles = [ 0 0 0 1.7 ]
#rotateangles = [ 0 -90 0 1.7 ]
#openwin alloccolors rotate saverot plot
#plot_color_axis = 0 input = [ -8 -3 10 ] GnuPlotHistogram
#plot_color_axis = 2 input = [ 0.6 50 50 ] GnuPlotHistogram
} }

proc openwindow { } {
  setup_window
  MD++ openwin alloccolors rotate saverot eval plot
}

#-----
proc exitmd { } { MD++ quit }
#end of proc exitmd
#-----

#-----
proc setup_md { } { MD++ {
  equilsteps = 0 timestep = 0.001 # (ps)
  atommass = 196.96655 # (g/mol)
  DOUBLE T = 1
  totalsteps = 1000
  saveprop = 1 savepropfreq = 100 openpropfile #run
  savecn = 1 savecnfreq = 10000 openintercnfile
  plotfreq = 100 printfreq = 100
  ensemble_type = "NVT" integrator_type = "VVerlet" implementation_type = 0
  vt2 = 1e28 #1e28 2e28 5e28
  wallmass = 2e3 # atommass * NP = 14380
  boxdamp = 1e-3 # optimal damping for 216 atoms and wallmass 1e-3
  saveH # Use current H as reference (H0), needed for specifying stress
  fixboxvec = [ 0 0 1
                1 0 1
                0 0 0 ]
  output_fmt = "curstep EPOT KATOM Tinst HELM HELMP TSTRESS_xx TSTRESS_yy TSTRESS_zz H_11 H_22 H_33"
} }
#end of proc setup_md

#*****
# Main program starts here
#*****
# status 0:
# 1:
# 2:
#
# read in status from command line argument
if { $argc == 0 } {
  set status 0
} elseif { $argc > 0 } {
  set status [lindex $argv 0]
}
puts "status = $status"

```

```
if { $argc <= 1 } {
  set n 0
} elseif { $argc > 1 } {
  set n [lindex $argv 1]
}
puts "n = $n"

if { $argc <= 2 } {
  set m 0
} elseif { $argc > 2 } {
  set m [lindex $argv 2]
}
puts "m = $m"

if { $status == 0 } {
  # create initial structure with (111) surface
  MD++ setnolog
  initmd $status $n $m
  readeam Au

  make_perfect_crystal 4 4 3

  setup_window
  openwindow

  relax_freebox
  MD++ eval
  MD++ finalcnfile = "au-perf.cn" writencn

  # make surface
  MD++ input = \[ 3 3 0.5 \] changeH_keepR

  # apply bi-axial compression
  MD++ input = \[ 1 1 -0.014 \] changeH_keepS
  MD++ input = \[ 2 2 -0.014 \] changeH_keepS

  relax_fixbox
  MD++ eval plot

  #setup_md
  #MD++ T OBJ = 30 initvelocity totalsteps = 1000 run
  #relax_fixbox

  MD++ finalcnfile = "au-surf.cn" writencn

  # click on atoms to get (sx,sy) coordinates of sites A, B, C

  MD++ sleep
  exitmd
} elseif { $status == 1 } {
  # put adatom on surface
  # prepare and relax dislocation structure
  MD++ setnolog
  initmd $status $n $m
  readeam Au

  # allocate more memory than needed because we will add atom later
  MD++ allocmultiple = 2
  MD++ incnfile = "../au-surf-0-0-0/au-surf.cn" readcn

  set n [MD++_Get "NP"]
  set indx [expr $n*3 ]
  set indy [expr $n*3+1]
  set indz [expr $n*3+2]

  # add one more atom on surface (site A)
  MD++ NP = [expr $n + 1]
```

```
setup_md

# put new atom on site A
MD++ SR($indx) = 0.0416
MD++ SR($indy) = 0.0
MD++ SR($indz) = 0.35

setup_window
openwindow

relax_fixbox
MD++ T_OBJ = 30 initvelocity totalsteps = 1000 run
relax_fixbox
MD++ finalcnfile = "au-surf-A.cn" writecn
MD++ finalcnfile = "au-surf-A.cfg" writeatomeyecfg

# put new atom on site B
MD++ incnfile = "../au-surf-0-0-0/au-surf.cn" readcn
MD++ NP = [expr $n + 1]
MD++ SR($indx) = 0.0
MD++ SR($indy) = 0.0
MD++ SR($indz) = 0.35
relax_fixbox
MD++ T_OBJ = 30 initvelocity totalsteps = 1000 run
relax_fixbox
MD++ finalcnfile = "au-surf-B.cn" writecn
MD++ finalcnfile = "au-surf-B.cfg" writeatomeyecfg

# put new atom on site Ap
MD++ incnfile = "../au-surf-0-0-0/au-surf.cn" readcn
MD++ NP = [expr $n + 1]
MD++ SR($indx) = -0.0208
MD++ SR($indy) = 0.0625
MD++ SR($indz) = 0.35
relax_fixbox
MD++ T_OBJ = 30 initvelocity totalsteps = 1000 run
relax_fixbox
MD++ finalcnfile = "au-surf-Ap.cn" writecn
MD++ finalcnfile = "au-surf-Ap.cfg" writeatomeyecfg

# put new atom on site Bp
MD++ incnfile = "../au-surf-0-0-0/au-surf.cn" readcn
MD++ NP = [expr $n + 1]
MD++ SR($indx) = 0.0625
MD++ SR($indy) = 0.0625
MD++ SR($indz) = 0.35
relax_fixbox
MD++ T_OBJ = 30 initvelocity totalsteps = 1000 run
relax_fixbox
MD++ finalcnfile = "au-surf-Bp.cn" writecn
MD++ finalcnfile = "au-surf-Bp.cfg" writeatomeyecfg

MD++ sleep
exitmd

} elseif { $status == 2 } {
# find minimum energy path for dislocation loop nucleation
MD++ setnolog
initmd $status $n $m
readeam Au

MD++ incnfile = "../au-surf-1-0-0/au-surf-$n.cn" readcn saveH
MD++ incnfile = "../au-surf-1-0-0/au-surf-$m.cn" readcn restoreH SHToR setconfig2
MD++ incnfile = "../au-surf-1-0-0/au-surf-$n.cn" readcn readcn setconfig1

setup_window
#openwindow

MD++ eval

MD++ fixallatoms constrain_fixedatoms freeallatoms
MD++ chainlength = 25 totalsteps = 1000
```

```
MD++ timestep = 0.01 printfreq = 2
MD++ allocchain initRchain

# nebspec = [ relax_surrounding_atoms reparam_freq fix_left_end fix_right_end climb_image]
MD++ { nebspec = [ 0 1 1 1 1 ] stringrelax }
MD++ finalcnfile = stringrelax.chain.cn writeRchain

#MD++ sleep
exitmd

} elseif { $status == 3 } {

    exitmd

} elseif { $status == 4 } {

    exitmd

} elseif { $status == 10 } {
# visualization
MD++ setnolog
initmd "view" $n $m
readpot

MD++ incnfile = "../au-surf-1-0-0/au-surf-$n.cn" readcn saveH
MD++ incnfile = "../au-surf-1-0-0/au-surf-$m.cn" readcn restoreH SHToR setconfig2
MD++ incnfile = "../au-surf-1-0-0/au-surf-$n.cn" readcn readcn setconfig1

MD++ zipfiles = 0
for { set iter 0 } { $iter <= 25 } { incr iter 1 } {
    MD++ incnfile = "../w-disl-nuc-${strain_yz}/stringrelax.chain.cn" readRchain

    # set chain number
    set chain_no $iter
    set total_no 25

    MD++ input = $iter copyRchaintoCN eval

    if { $iter == 0 } {
        openwindow
    }
    MD++ eval plot
    # write cn or cfg file
    MD++ finalcnfile = "chain_no_${chain_no}.cn" writecn
    MD++ finalcnfile = "chain_no_${chain_no}.cfg" writeatomeyecfg
    MD++ sleepseconds = 1 sleep

}
MD++ sleepseconds = 100 sleep
exitmd

} else {

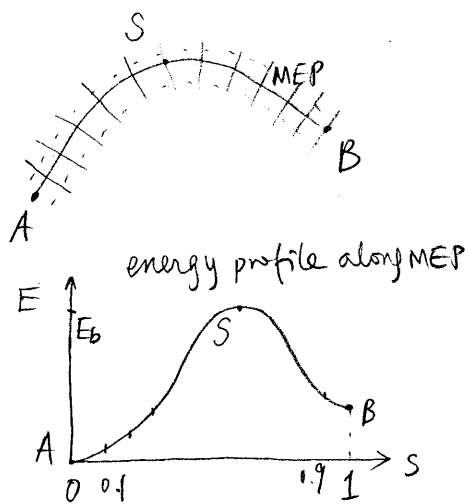
puts "unknown status = $status"
exitmd

}
```

Today's goal:

## Umbrella Sampling

### 1. Reaction Coordinate



S (upper case) is the saddle point.

Imagine an MEP connecting metastable states A and B.

We can divide MEP into many small sections with "ticks" marking 0.1, 0.2, ...

They measure the progress the system has made in transitioning from A to B

Define  $s$  as the arc length along MEP from A to another point on MEP, divided by the arc length of the entire MEP.

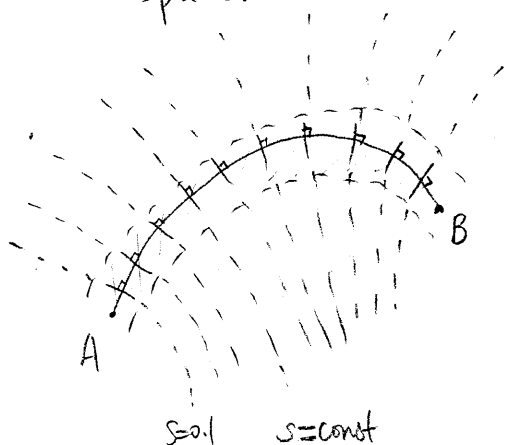
Hence,  $s=0 \Leftrightarrow A$

$s=1 \Leftrightarrow B$

$0 < s < 1 \Leftrightarrow$  somewhere between A & B.

$s$  (lowercase) is a reaction coordinate, measuring the progress the system has made in transitioning from A to B.

Here  $s$  is only defined for points on MEP. We would like a reaction coordinate defined for every point in the configuration space.



For this purpose, imagine a function  $s(\{q_i\})$  and plot the isosurfaces

$$s(\{q_i\}) = 0.1,$$

$$s(\{q_i\}) = 0.2,$$

in configuration space.

$s(\{q_i\})$  is a suitable reaction coordinate if its isosurfaces are locally perpendicular to MEP.

At low enough temperature, the transition paths are in the neighborhood of MEP. So the reaction coordinate need to be well chosen only in the "tubular" region around MEP. Hence, there are some arbitrariness in the selection of reaction coordinates.

\* An optimal reaction coordinate exists: it is a function of the commitor probability  $s(\{q_i\}) = f(P_{\text{comm}}(\{q_i\}))$ .  $P_{\text{comm}}(\{q_i\})$  measures the probability for a system, starting at  $\{q_i\}$ , to reach state B before reaching state A, and is called the commitor (probability).

$s(\{q_i\}) = f(P_{\text{comm}}(\{q_i\}))$  means that all points on the isosurface of  $s(\{q_i\})$  have the same commitor probability.

The dividing surface between state A and state B is the isosurface  $s(\{q_i\}) = 0.5$

Q: Why do we need the reaction coordinate?

From transition state theory

$$K = \kappa K^{\text{TST}}, \quad K^{\text{TST}} = \frac{k_B T}{h} \frac{Z^*}{Z_A}$$

in the strong friction limit  $\kappa = \frac{\omega_b}{\gamma}$  (see lecture 8, p 6)

$$Z_A = \left( \frac{\sqrt{2\pi m k_B T}}{h} \right)^{3N} \int \prod_{\underline{R}} d^3 \underline{R} e^{-\beta[U(\underline{R}) - U(\underline{R}_A)]} \Theta(s(\underline{R}) - s^*)$$

$$Z^* = \left( \frac{\sqrt{2\pi m k_B T}}{h} \right)^{3N-1} \int \prod_{\underline{R}} d^3 \underline{R} e^{-\beta[U(\underline{R}) - U(\underline{R}_A)]} \delta(s(\underline{R}) - s^*)$$

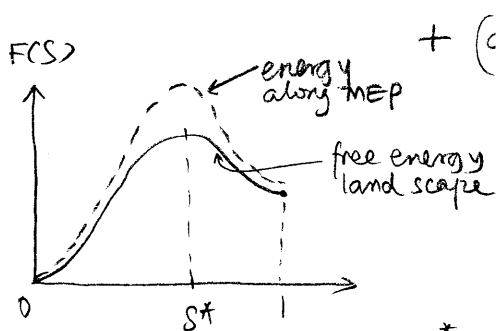
↑  
reaction coordinate at S

define  $F_A = -k_B T \ln Z_A$  ----- free energy of state A

$F^* = -k_B T \ln Z^*$  ----- free energy of dividing surface.

$$K = \kappa \cdot \frac{k_B T}{h} \cdot e^{-\frac{F^* - F_A}{k_B T}}$$

$$\text{define } F(\vec{s}) = \left( \frac{\sqrt{2\pi m k_B T}}{h} \right)^{3N-1} \int d^{3N} \underline{R} e^{-\beta [U(\underline{R}) - U(\underline{R}_A)]} \delta(s(\underline{R}) - \vec{s})$$



+ (arbitrary constant)

to make  $F(s=0) = 0$

The same (arbitrary constant) must be added to  $F_A$ .

$F(s^*)$  is the activation free energy of the transition  
 $F(s) \rightarrow E(s)$  along MEP only in the zero temperature limit.

$$\begin{cases} F^* \equiv F(s^*) \\ F_A = -k_B T \ln \left[ \frac{\sqrt{2\pi m k_B T}}{h} \int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds \right] \\ K = \kappa \cdot \frac{k_B T}{h} \cdot e^{-\frac{F^* - F_A}{k_B T}} \end{cases} \quad (\text{self-consistent})$$

Note that both  $F_A$  and  $\left(\frac{k_B T}{h}\right)$  contains the Planck constant  $h$ , which must cancel.

$$e^{-\beta F_A} = \frac{\sqrt{2\pi m k_B T}}{h} \int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds$$

$$K = \kappa \cdot \frac{k_B T}{h} \cdot \frac{e^{-\frac{F(s^*)}{k_B T}}}{\frac{\sqrt{2\pi m k_B T}}{h} \int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds} = \kappa \cdot \frac{k_B T}{\sqrt{2\pi m}} \cdot \frac{e^{-\frac{F(s^*)}{k_B T}}}{\int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds}$$

It is convenient to define

$$p(s) \equiv \frac{e^{-\frac{F(s)}{k_B T}}}{\int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds}$$

the probability density of finding the system with reaction coordinate  $s$

$$e^{-\frac{\tilde{F}_A}{k_B T}} \equiv \int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds = \frac{h}{\sqrt{2\pi m k_B T}} e^{-\frac{F_A}{k_B T}}$$

$$\text{equivalently, } \tilde{F}_A = F_A - k_B T \ln \frac{h}{\sqrt{2\pi m k_B T}}$$

Then

$$\begin{cases} K = \kappa \cdot \frac{k_B T}{\sqrt{2\pi m}} \cdot e^{-\frac{F(s^*) - \tilde{F}_A}{k_B T}} = \kappa \cdot \frac{k_B T}{\sqrt{2\pi m}} \cdot p(s^*) \\ \tilde{F}_A \equiv -k_B T \ln \left[ \int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds \right] \end{cases} \quad (\text{This is the rate predicted by TST.})$$



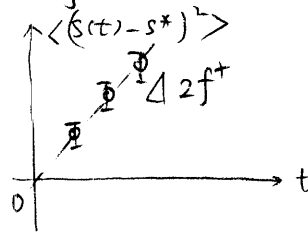
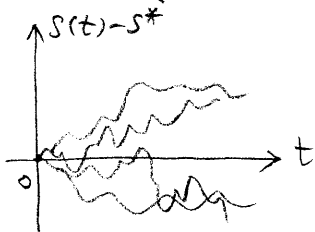
Alternatively, we can use the Becker-Döring theory  
(valid in the strong friction limit, see lecture 7, p. 7)

$$k^{BD} = f^+ \cdot \Gamma \cdot e^{-\frac{F(s^*) - \tilde{F}_A}{k_B T}} = f^+ \cdot \Gamma \cdot p(s^*)$$

note  $p(s) = \frac{e^{-\frac{F(s)}{k_B T}}}{\int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds}$ ,  $\tilde{F}_A = -k_B T \ln \left[ \int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds \right]$

$f^+$  is the "molecular attachment" rate. It measures the "diffusivity" of the system along the  $s$ -axis near  $s^*$ ,

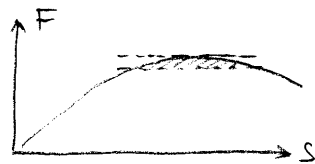
i.e.  $\langle (s(t) - s^*)^2 \rangle = 2 f^+ t$



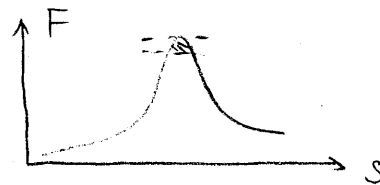
run a series of simulations starting from dividing surface  $s = s^*$

$\Gamma = \sqrt{\frac{K_b}{2\pi k_B T}}$  is the Zeldovich factor,  $K_b = -\frac{\partial^2 F(s)}{\partial s^2} \Big|_{s^*}$

The Zeldovich factor accounts for the recrossing events



smaller  $K_b$   
smaller  $\Gamma$   
smaller rate



larger  $K_b$   
larger  $\Gamma$   
larger rate

In the strong friction limit,  $k^{BD} = k \cdot k^{TST}$  (see lecture 7, p. 8)

The Becker-Döring theory predictions have been numerically verified

in the cases of Ising model: Phys. Rev. E 81, 030601(R), 2010  
82, 011603, 2010

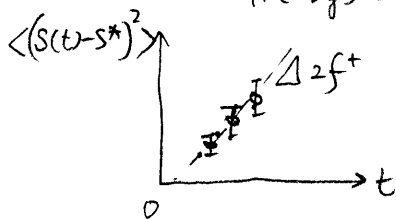
and dislocation nucleation in crystals

Proc. Natl. Acad. Sci. 108, 5174, 2011.  
J. Mat. Res. 26, 2335, 2011.

To use Becker-Döring theory, we need the free energy landscape or free energy profile,  $F(s)$ , from which we obtain

$$\rightarrow \begin{cases} F(s^*) \\ \Gamma = \sqrt{\left(-\frac{\partial^2 F}{\partial s^2}\right)_{s^*} \cdot \frac{1}{2\pi k_B T}} \end{cases}$$

$f^+$  can be computed by running a series of simulations starting from the dividing surface  $S(\mathbf{r}; s) = s^*$  and observe how fast the system diffuse away from it.



## 2. Umbrella Sampling

Q: How do we compute  $F(s)$ ?

- Options =
1. Integration to find  $Z(\hat{s}) = \left(\frac{N^2 T m k_B T}{h}\right)^{3N-1} \int d\mathbf{R} e^{-\beta U} \delta(S(\mathbf{R}) - \hat{s})$   
and  $F(\hat{s}) = -k_B T \ln Z(\hat{s})$
  2. Adiabatic switching  
need to constrain system on isourface  $S(\mathbf{R}) = \hat{s}$
  3. Umbrella Sampling.

Q: Remind ourselves, why do we need FCS?

The probability density of finding system at  $s$  is

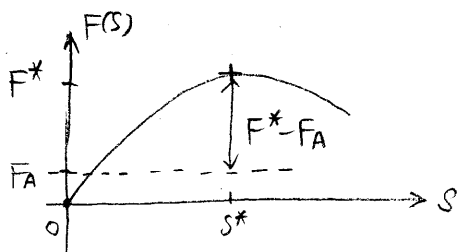
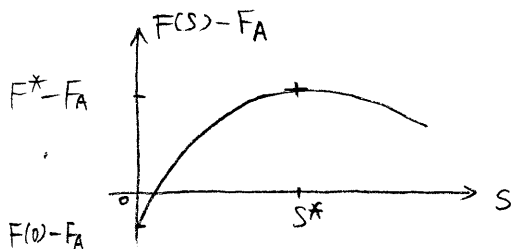
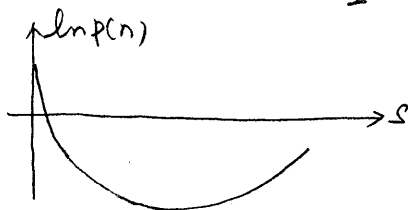
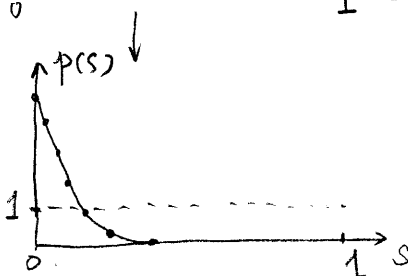
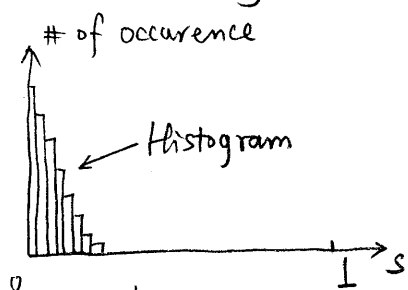
$$p(s) = e^{-\frac{F(s) - \tilde{F}_A}{k_B T}}$$

So if we can compute  $p(s)$ , then

$$F(s) - \tilde{F}_A = -k_B T \ln p(s)$$

Wait a minute...

We can compute  $p(s)$  by running our simulation and building a histogram.



Sounds too easy...

What will be the problem of this approach?

$$p(s) = \frac{e^{-\frac{F(s)}{k_B T}}}{\int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds}$$

$$\ln p(s) = -\frac{F(s)}{k_B T} - \ln \left[ \int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds \right]$$

$$F(s) = -k_B T \ln p(s) - \underbrace{k_B T \ln \left[ \int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds \right]}_{\text{a constant at fixed } T}$$

For convenience, we redefine  $F(s)$  as

$$F(s) = -k_B T \ln p(s) + \text{const.}$$

where the const. is chosen such that

$$F(s=0) = 0$$

This does not change any physics because we can always find  $\tilde{F}_A$  and  $p(s)$  from

$$e^{-\beta \tilde{F}_A} = \int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds$$

$$\tilde{F}_A = -k_B T \ln \left[ \int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds \right]$$

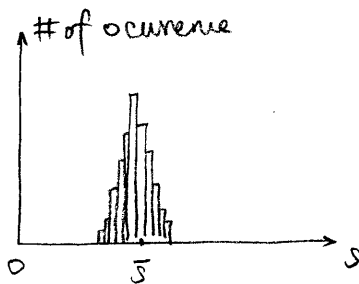
$$p(s) = e^{-\frac{F(s) - \tilde{F}_A}{k_B T}}$$

To gain statistics in arbitrary range of  $s$ , apply a bias potential

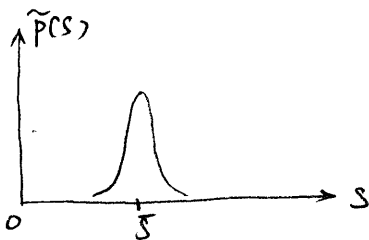
$$U_{\text{bias}}(R) = \frac{1}{2} K_{\text{bias}} (s(R) - \bar{s})^2$$

The new system will have the following free energy landscape

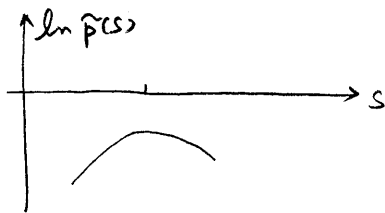
$$\tilde{F}(s) = F(s) + \frac{1}{2} K_{\text{bias}} (s - \bar{s})^2$$



$\bar{s}$  is the center of the window

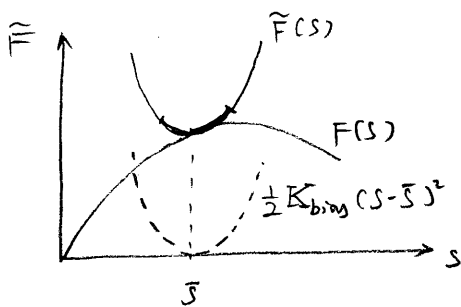


$$\tilde{p}(s) = \frac{e^{-\frac{\tilde{F}(s)}{k_B T}}}{\int_{-\infty}^{+\infty} e^{-\frac{\tilde{F}(s)}{k_B T}} ds}$$



$$\ln \tilde{p}(s) = -\frac{\tilde{F}(s)}{k_B T} - \ln \left[ \int_{-\infty}^{+\infty} e^{-\frac{\tilde{F}(s)}{k_B T}} ds \right]$$

$$\tilde{F}(s) = -k_B T \ln \tilde{p}(s) - \underbrace{k_B T \ln \left[ \int_{-\infty}^{s^*} e^{-\frac{F(s)}{k_B T}} ds \right]}_{\text{a constant at fixed } T}$$



This means that given  $\tilde{p}(s)$  we can determine  $\tilde{F}(s)$  up to a constant.

i.e. shifting  $\tilde{F}(s)$  by any constant will give the same  $\tilde{p}(s)$  through

$$\tilde{p}(s) = \frac{e^{-\frac{\tilde{F}(s)}{k_B T}}}{\int_{-\infty}^{+\infty} e^{-\frac{\tilde{F}(s)}{k_B T}} ds}$$

Hence, we can compute  $F(s)$  up to a constant

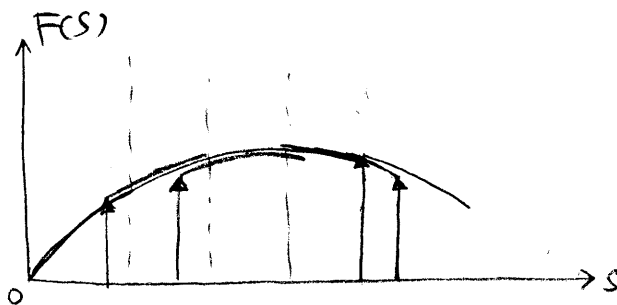
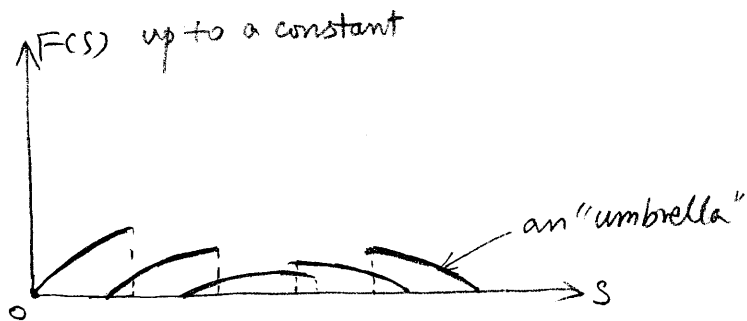
$$F(s) = \tilde{F}(s) - \frac{1}{2} K_{\text{bias}} (s - \bar{s})^2 + (\text{arbitrary constant})$$

$$= -k_B T \ln \tilde{p}(s) - \frac{1}{2} K_{\text{bias}} (s - \bar{s})^2 + (\text{arbitrary constant})$$

## Overlapping Windows.

Imagine we perform a series of simulations, each with a different bias potential, centered at different  $\bar{n}$ .

Each simulation will inform us the shape of  $F(s)$  in the neighborhood of  $\bar{n}$ , up to a constant.



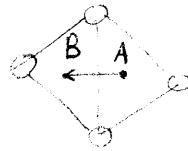
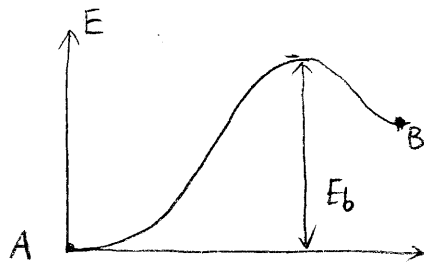
arrows indicates the amount of shift for each curve to piece together a continuous function  $F(s)$ .

Different sections of the  $F(s)$  need to be shifted (up or down) by different amount to obtain a single  $F(s)$  curve.

Today's goal

Umbrella Sampling case study.

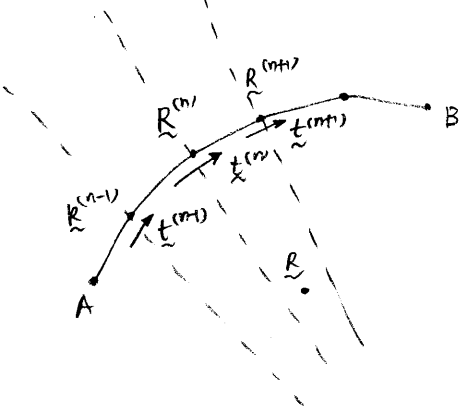
1. An adatom diffusion on (111) surface, reaction coordinate



We have computed MEP using string method  $E_b = 0.083 \text{ eV}$

To apply umbrella sampling method, first we need to define a reaction coordinate  $s$

For convenience, we let  $s(A) = 0$ ,  $s(B) = 100$  and will build histograms for  $s = 0, 1, 2, \dots, 100$ .



The reaction coordinates for chain of states  $\underline{R}^{(n)}$   $n=0, 1, \dots, N_c-1$ , in the relaxed MEP are

$$s(\underline{R}^{(n)}) = \frac{n}{N_c-1} \cdot 100$$

We now need to define  $s$  for arbitrary  $\underline{R}$

Define tangent vector  $\underline{t}^{(n)} = \frac{\underline{R}^{(n+1)} - \underline{R}^{(n)}}{|\underline{R}^{(n+1)} - \underline{R}^{(n)}|}$

for  $n=1, 2, \dots, N_c-1$

compute  $d^{(n)} = (\underline{R} - \underline{R}^{(n)}) \cdot \underline{t}^{(n)}$

$d^{(n)} > 0$  means  $\underline{R}$  is to the "right" of plane  $\underline{t}^{(n)}$

$d^{(n)} < 0$  means  $\underline{R}$  is to the "left" of plane  $\underline{t}^{(n)}$

We will use the following algorithm to compute  $S(\underline{R})$

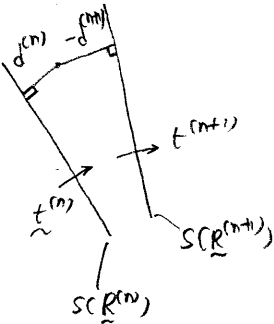
① find the smallest  $n$  in  $1, 2, \dots, N_c - 2$  such that

$$d^{(1)} > 0, d^{(2)} > 0, \dots, d^{(n)} > 0, d^{(n+1)} \leq 0$$

if found, then

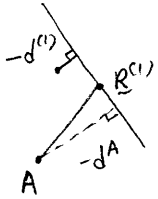
$S(\underline{R})$  is the linear interpolation of  $S(\underline{R}^{(n)})$  and  $S(\underline{R}^{(n+1)})$

$$S(\underline{R}) = \frac{d^{(n)} S(\underline{R}^{(n+1)}) - d^{(n+1)} S(\underline{R}^{(n)})}{d^{(n)} - d^{(n+1)}}$$



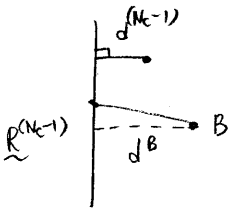
② If  $d^{(1)} \leq 0$ , then compute  $d^A = (\underline{R}^A - \underline{R}^{(1)}) \cdot \underline{t}^{(1)}$

$$S(\underline{R}) = \frac{d^{(1)} - d^A}{-d^A} S(\underline{R}^{(1)}) \quad \text{note: } S(\underline{R}^A) = 0$$

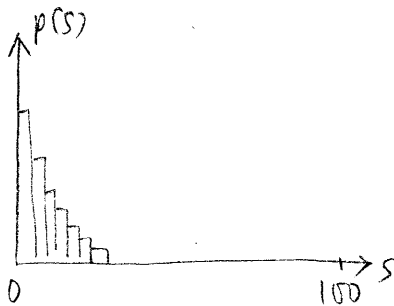


③ If  $d^{(N_c-1)} > 0$ , then compute  $d^B = (\underline{R}^B - \underline{R}^{(N_c-1)}) \cdot \underline{t}^{(N_c-1)}$

$$S(\underline{R}) = S(\underline{R}^B) - (S(\underline{R}^B) - S(\underline{R}^{(N_c-1)})) \frac{d^{(N_c-1)}}{d^B}$$



## 2. Umbrella Sampling



First, build histogram for probability

$p(s)$  using MC simulation with no bias potential.

MD++ react\_coord\_type = 3 ← using MEP to define reaction coordinate.

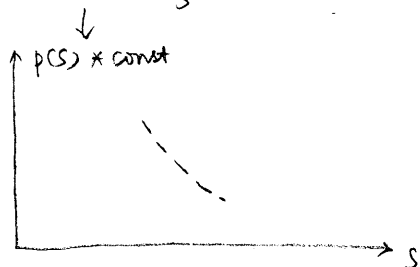
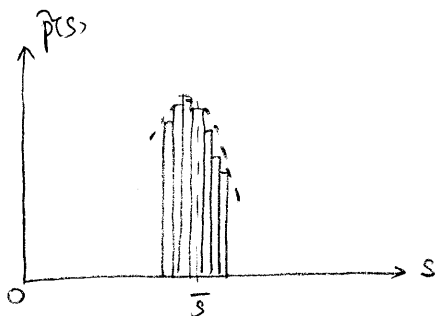
MD++ YES\_UMB = 1 run MC

data saved in Freq.txt, Prob.txt, Narray.txt + Nsample.txt.

Note  $p(s)$  is the absolute probability

$$p(s) = \frac{e^{-\frac{F(s)}{k_B T}}}{\int e^{-\frac{F(s)}{k_B T}} ds}$$

but we only get statistics in a small domain, e.g.  $s \leq 30$ .



Second, perform MC simulation with bias potential  $U_{bias} = \frac{1}{2} K_{bias} (s - \bar{s})^2$

MD++ UMB\_K =  $3e-4$  ←

MD++ n\_center = 50 ←

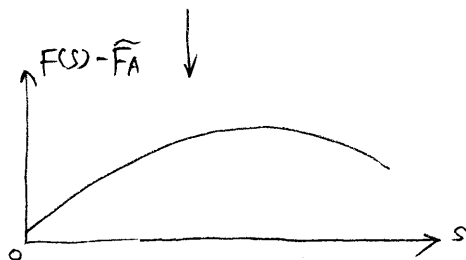
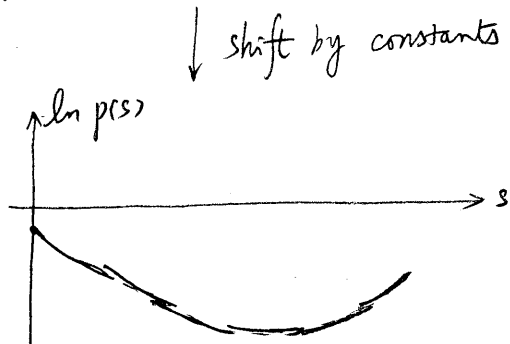
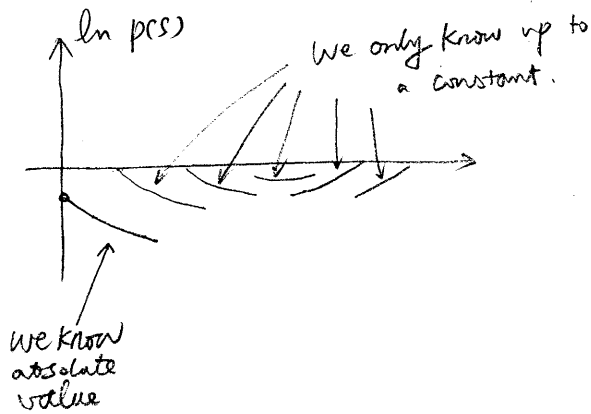
$$\tilde{F}(s) = F(s) + \frac{1}{2} K_{bias} (s - \bar{s})^2$$

Histogram tells us

$$\tilde{p}(s) = \frac{e^{-\frac{\tilde{F}(s)}{k_B T}}}{\int e^{-\frac{\tilde{F}(s)}{k_B T}} ds} = \frac{e^{-\frac{F(s)}{k_B T}} \cdot e^{-\frac{K_{bias}(s-\bar{s})^2}{2k_B T}}}{\int e^{-\frac{F(s)}{k_B T}} ds}$$

$$p(s) = \tilde{p}(s) \cdot e^{\frac{K_{bias}(s-\bar{s})^2}{2k_B T}} \cdot \underbrace{\left[ \frac{\int e^{-\frac{F(s)}{k_B T}} ds}{\int e^{-\frac{F(s)}{k_B T}} ds} \right]}_{\text{unknown constant}}$$





Note that in the unbiased sampling, we know the absolute value of  $p(s)$ .

In the sampling with bias potential, we know  $p(s)$  up to a multiplicative constant. i.e. we know  $\ln p(s)$  up to an additive constant.

If the sampling windows overlap, we can shift them by appropriate constants to form a continuous curve  $\ln p(s)$ , and then

$$F(s) - \hat{F}_A = -k_B T \ln p(s)$$

$$\text{where } e^{-\beta \hat{F}_A} = \int e^{-\beta F(s)} ds$$

The "Self-Consistent Histogram" method is used to determine the constants that each window needs to shift to form a continuous curve.

Today's goal

Self-consistent Histogram Method

(Ref: Frenkel & Smit, Understanding Molecular Simulations,  
2nd ed. p.179-196)

Q: How to piece together the relative probabilities estimated from different sampling windows to get a continuous function of  $p(s)$ ?

### 1. Error in Histograms

For the unperturbed system,  
(i.e. without bias potential)

define

$$p_0(\hat{s}) = \frac{\int d^3N_{\underline{R}} e^{-\beta[U(\underline{R}) - U(\underline{R}_A)]} \delta(s(\underline{R}) - \hat{s})}{\int d^3N_{\underline{R}} e^{-\beta[U(\underline{R}) - U(\underline{R}_A)]}}$$

This is the same as  $p(\hat{s})$  defined in lecture 11.  
Here we added subscript 0 to indicate that there is no bias potential.

define

$$Z_0 \equiv \int d^3N_{\underline{R}} e^{-\beta[U(\underline{R}) - U(\underline{R}_A)]}$$

The Landau Free energy (of the original system) is defined as

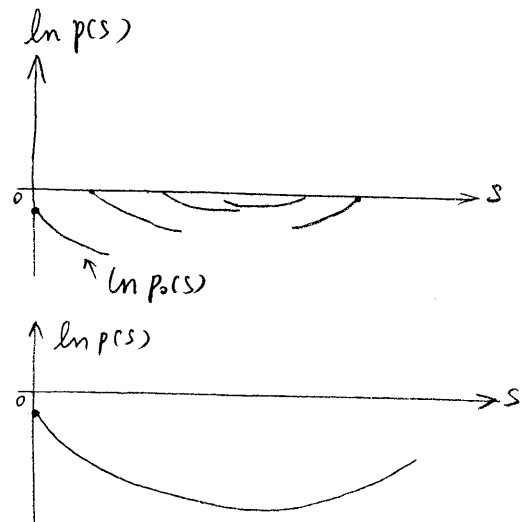
$$F_{\text{Landau}}(s) \equiv -k_B T \ln p_0(s) = F(s) - \tilde{F}_A$$

$p_0(s)$  can be estimated by building a histogram from a simulation of the original system

$$p_0(s) \cdot \Delta s = \frac{\langle H_0(s) \rangle}{M_0}$$

$M_0$  is the total number of points in the histogram

$H_0(s)$  is the number of points falling inside  $[s - \frac{\Delta s}{2}, s + \frac{\Delta s}{2})$



Q: What is the error bar of the estimate  $p_0(s) \Delta S$  by  $\frac{H_0(s)}{M_0}$ ?

Intuitively, we expect that the error will become smaller with increasing  $M_0$ .

Assume  $H_0(s)$  satisfies the Poisson distribution

\* The Poisson distribution expresses the probability of a given number of events occurring in a fixed interval of time, if these events occur with a known average rate and independent of the time since the last event. [http://en.wikipedia.org/wiki/Poisson\\_distribution](http://en.wikipedia.org/wiki/Poisson_distribution)

A discrete stochastic variable  $X$  is said to have a Poisson distribution with parameter  $\lambda > 0$ , for  $k=0, 1, 2, \dots$  if

$$P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Note that  $\sum_{k=0}^{\infty} P(X=k) = e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot e^{\lambda} = 1$  normalization condition

expectation value,  $\langle X \rangle = \sum_{k=0}^{\infty} \frac{k \cdot \lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \cdot \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}$

$$= e^{-\lambda} \cdot \lambda \cdot \frac{\partial}{\partial \lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot \lambda \cdot \frac{\partial}{\partial \lambda} e^{\lambda} = \lambda$$

Variance  $V(X) = \langle X^2 \rangle - \langle X \rangle^2 = \lambda$

standard deviation  $\sigma(X) = \sqrt{\lambda}$

Hence, we expect the variance of estimator  $\frac{H_0(s)}{M_0}$  is

$$V\left(\frac{H_0(s)}{M_0}\right) = \frac{\langle H_0^2(s) \rangle - \langle H_0(s) \rangle^2}{M_0^2} = \frac{\langle H_0(s) \rangle}{M_0^2} = \frac{p_0(s) \Delta S}{M_0}$$

relative error:  $\frac{\sigma\left(\frac{H_0(s)}{M_0}\right)}{\frac{\langle H_0(s) \rangle}{M_0}} = \frac{\sqrt{\frac{p_0(s) \Delta S}{M_0}}}{\frac{p_0(s) \Delta S}{M_0}} = \frac{1}{\sqrt{p_0(s) \Delta S \cdot M_0}} = \frac{1}{\sqrt{\langle H_0(s) \rangle}}$

i.e. the more count there is for bin  $s$ , the more accurate is the estimate of  $p_0(s)$

## 2. Sampling with the help of bias potential

Since the relative error for estimating  $p_0(s)$  is  $\frac{1}{\sqrt{\langle H_0(s) \rangle}}$ , we need to collect sufficient "hits",  $H_0(s)$ , in the bin  $[s - \frac{\Delta s}{2}, s + \frac{\Delta s}{2})$

$\langle H_0(s) \rangle$	relative error
4	50%
100	10%
1000	3%

If  $H_0(s) = 0$  for a given bin  $[s - \frac{\Delta s}{2}, s + \frac{\Delta s}{2})$ , i.e. no hit!, it does not tell much about  $p_0(s)$  (except for an upper bound).

There are regions in which  $p_0(s)$  is very low, i.e.  $p_0(s) \ll (M_0 \Delta s)^{-1}$  then  $\langle H_0(s) \rangle = p_0(s) \Delta s \cdot M_0 \ll 1$ , then most likely we will get  $H_0(s) = 0$  in these regions, in which the histogram does not tell us much about  $p_0(s)$ .

Let us now consider a series of simulations, each with a different bias potential, e.g.

$$W_i(s) = \frac{1}{2} K_{\text{bias}} (s - \bar{s}_i)^2 \quad \bar{s}_i \text{ is the center of each sampling window}$$

From each simulation, we build a histogram  $H_i(s)$  which estimates

$$p_i(s) \Delta s = \frac{\langle H_i(s) \rangle}{M_i}, \quad M_i \text{ is the total number of points in Histogram } i.$$

where

$$p_i(\hat{s}) = \frac{\int d^N \underline{R} e^{-\beta [U(\underline{R}) - U(\underline{R}_A) + W_i(s(\underline{R}))]} \delta(s(\underline{R}) - \hat{s})}{\int d^N \underline{R} e^{-\beta [U(\underline{R}) - U(\underline{R}_A) + W_i(s(\underline{R}))]}}$$

define  $Z_i \equiv \int d^N \underline{R} e^{-\beta [U(\underline{R}) - U(\underline{R}_A) + W_i(s(\underline{R}))]}$

Hence

$$p_i(\hat{s}) = \frac{\int d^{3N}R e^{-\beta[U(R) - U(R_A) + W_i(\hat{s})]} \delta(S(R) - \hat{s})}{Z_i}$$

$$= \frac{1}{Z_i} \cdot e^{-\beta W_i(\hat{s})} \cdot \int d^{3N}R e^{-\beta[U(R) - U(R_A)]} \delta(S(R) - \hat{s})$$

$$\therefore \frac{p_o(\hat{s})}{p_i(\hat{s})} = \frac{Z_i}{Z_o} e^{\beta W_i(\hat{s})}$$

This means that the probabilities of observing  $\hat{s}$  in the original system and the system with bias potential  $W_i(s)$  are related by a known function  $e^{\beta W_i(\hat{s})}$  times a constant  $\frac{Z_i}{Z_o}$  (which remains unknown).

$$p_o(\hat{s}) = e^{\beta W_i(\hat{s})} \cdot \frac{Z_i}{Z_o} \cdot p_i(\hat{s}) = e^{\beta W_i(\hat{s})} \cdot \frac{Z_i}{Z_o} \cdot \frac{\langle H_i(s) \rangle}{M_i} \cdot \frac{1}{\Delta s}$$

This means we can estimate  $p_o(\hat{s})$  through histogram  $H_i(s)$ , provided the ratio  $\frac{Z_i}{Z_o}$  can be found  $\rightarrow$  what we will do next.

Intuitively, there are regions in which  $\langle H_i(s) \rangle$  provides better estimates for  $p_o(s)$  than  $\langle H_o(s) \rangle$ . These are the regions in which  $H_i(s) \gg 1$  while  $H_o(s) \sim 0$ .

These regions vary with  $i$  as the center of the sampling window changes.

There are regions in which several histograms provide non-zero statistics, i.e.  $H_i(s) > 0$ . Intuitively, all these histograms can be combined to provide the "best" estimate of  $p_o(s)$ , i.e.

$$p_o^{\text{est}}(\hat{s}) = \sum_{i=1}^n w_i(\hat{s}) e^{\beta W_i(\hat{s})} \frac{Z_i}{Z_o} \frac{H_i(\hat{s})}{M_i \cdot \Delta s}$$

where  $w_i(\hat{s})$  is a weight function still to be determined.

$$\sum_{i=1}^n w_i(\hat{s}) = 1 \quad \text{for each } \hat{s}. \quad \text{Note } \frac{Z_i}{Z_o} \text{ are still unknown.}$$

### 3. Finding the optimal weight $w_i(\hat{s})$

The "optimal" weight function  $w_i(\hat{s})$  is chosen to minimize the variance of  $p_0^{\text{est}}(\hat{s})$ .

$$\begin{aligned} V(p_0^{\text{est}}(\hat{s})) &= \langle (p_0^{\text{est}}(\hat{s}))^2 \rangle - \langle p_0^{\text{est}}(\hat{s}) \rangle^2 \\ &= \sum_{i=1}^n w_i^2(\hat{s}) e^{2\beta W_i(\hat{s})} \left(\frac{z_i}{z_0}\right)^2 V\left(\frac{H_i(\hat{s})}{M_i \Delta s}\right) \\ &= \sum_{i=1}^n w_i^2(\hat{s}) e^{2\beta W_i(\hat{s})} \left(\frac{z_i}{z_0}\right)^2 \frac{p_i(\hat{s})}{M_i \Delta s} \\ &= p_0(\hat{s}) \sum_{i=1}^n w_i(\hat{s}) e^{\beta W_i(\hat{s})} \frac{z_i}{z_0} \cdot \frac{1}{M_i \Delta s} \end{aligned}$$

minimizing  $V(p_0^{\text{est}}(\hat{s}))$  subjected to the constraint  $\sum_{i=1}^n w_i(\hat{s}) = 1$

gives 
$$w_i(\hat{s}) = \frac{e^{-\beta W_i(\hat{s})} M_i \frac{z_0}{z_i}}{\sum_{i=1}^n e^{-\beta W_i(\hat{s})} M_i \frac{z_0}{z_i}}$$

Note: Histogram  $H_i$  is given more weight at  $\hat{s}$  if

$M_i \uparrow \iff$  more statistics

$e^{-\beta W_i(\hat{s})} \uparrow \iff W_i(\hat{s}) \downarrow \iff$  bias potential at  $\hat{s}$  is small  
i.e.  $\hat{s}$  close to the center of sampling window  $\bar{s}_i$

Plug this expression for  $w_i(\hat{s})$  into the expression for  $p_0^{\text{est}}(\hat{s})$

$$p_0^{\text{est}}(\hat{s}) = \frac{\sum_{i=1}^n H_i(\hat{s})}{\sum_{i=1}^n e^{-\beta W_i(\hat{s})} M_i \frac{z_0}{z_i}}$$

\* note  $e^{\beta W_i(\hat{s})}$  disappears from Numerator and  $e^{-\beta W_i(\hat{s})}$  (a more benign term!) appears in Denominator.

#### 4. Determine $Z_i/Z_0$

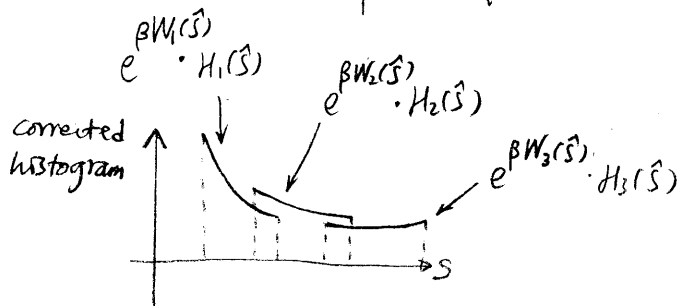
Finally, we return to the definition of  $Z_i$ :

$$\begin{aligned} Z_i &\equiv \int d^N \underline{R} e^{-\beta [U(\underline{R}) - U(\underline{R}_A) + W_i \cdot (S(\underline{R}))]} \\ &= \int d\hat{S} e^{-\beta W_i \cdot (\hat{S})} \cdot \int d^N \underline{R} e^{-\beta [U(\underline{R}) - U(\underline{R}_A)]} \delta(S(\underline{R}) - \hat{S}) \\ &= \int d\hat{S} e^{-\beta W_i \cdot (\hat{S})} \cdot Z_0 \cdot p_0(\hat{S}) \end{aligned}$$

substitute the expression for  $p_0(\hat{S})$

$$Z_i \equiv \int d\hat{S} e^{-\beta W_i \cdot (\hat{S})} \cdot \frac{\sum_{k=1}^n H_k(\hat{S})}{\sum_{k=1}^n e^{-\beta W_k \cdot (\hat{S})} M_k \frac{1}{Z_k}} \quad \begin{array}{l} \text{(note cancellation)} \\ \text{of } Z_0 \end{array}$$

$\{Z_i\}$   $i=1, \dots, n$  must be solved self-consistently from this set of implicit equations. Only their ratio can be found, so we can set (arbitrarily)  $Z_1 = 1$ .



Consider three histograms  $H_1(\hat{S})$ ,  $H_2(\hat{S})$ ,  $H_3(\hat{S})$  with overlaps.

For a fixed  $Z_1$ ,

$Z_2$  can be found by minimizing

objective function

$$f(Z_2, Z_3) = \left( Z_2 - \int d\hat{S} e^{-\beta W_2(\hat{S})} \frac{H_1(\hat{S}) + H_2(\hat{S}) + H_3(\hat{S})}{e^{-\beta W_1(\hat{S})} H_1(\hat{S}) \frac{M_1}{Z_1} + e^{-\beta W_2(\hat{S})} H_2(\hat{S}) \frac{M_2}{Z_2} + e^{-\beta W_3(\hat{S})} H_3(\hat{S}) \frac{M_3}{Z_3}} \right)$$

This can be achieved by calling the `fminsearch` function in Matlab.

See `analyze_us_data.m`, `sclabs.m`

↑ the objective function

using the same routine, we can determine  $Z_3$  by  $\min_{Z_3} f(Z_2, Z_3, Z_4)$ , etc.

In principle, `fminsearch` can be used to find all  $\{Z_i\}$  in one multi-dimensional search.