GEM4 Summer School OpenCourseWare
http://gem4.educommons.net/
http://www.gem4.org/

Molecular Mechanics notes

Please use the following citation format:

> *Notes from* Hwang, Wonmuk and Ju Li. "Parallel tutorial session #7: Molecular Mechanics." Lecture series, GEM4 session at MIT, Cambridge, MA, August 9, 2006. http://gem4.educommons.net/ (accessed MM DD, YYYY). License: Creative Commons Attribution-Noncommercial-Share Alike.

Note: Please use the actual date you accessed this material in your citation.

Parallel tutorial session #7 : MOLECULAR MECHANICS

08 / 09 / 2006

Wonmuk Hwang

Molecular mechanics fundamentals

- <u>macroscopic mechanics</u> (classical mechanics) : position , velocity , acceleration

  describe the motion : kinematics

  force = cause of motion : dynamics

  few- body problem : deterministic from Newton's equation

  <u>molecular mechanics</u> (<u>statistical</u> mechanics)

  many- body problem → continuum approximation sometimes

  stochastic, random , chaotic ≈ thermal noise

  collective behavior can be known, details can be ignored to some extent

- the ensemble concept : one " macro- state " (structure, energy function, ...)

  ≡ many equally accessible " micro- states "

  (backbone chain orientations, bond lengths, ...)

  <u>ensemble</u> = a set of microstates that correspond to a given macrostate in equilibrium.

  molecular force : water and proteins are a closed system of very many microstates.

  $\begin{cases} \text{the macrostate of water is negligibly affected by the protein} \\ \text{water} \equiv \text{heat reservoir} \end{cases}$

  why does a protein bend (kinesin) ? deterministic ≠ random motion

- <u>entropy S</u>

  Let Ω be the number of microstates in an ensemble ;

  why is S defined as a logarithm ? S is additive

  e.g. dice $\Omega_d = 6$, coin $\Omega_c = 2$

  both dice and coin $\Omega_{tot} = \Omega_d \times \Omega_c$

  $S_{tot} = S_d + S_c$

  some additive physical variables : E, N, V, S   extensive quantities

  others are independent of system size : T, P, $\rho = N/V$ intensive quantities

$S \nearrow$ (entropy) when $E \nearrow$ (energy)

$\begin{cases} \Omega = \exp \left(S / k_B \right) \\ S = k_B \ln \Omega \\ k_B = 1.38 \times 10^{-23} \text{J/K} \\ \underline{Boltzmann} \text{'s constant} \end{cases}$

- Boltzmann factor
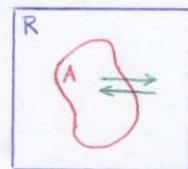
R heat reservoir
A system of interest
⇌ exchange of energy

Let $p_i = p(E_i)$ the probability that A is in a microstate of energy E_i

states
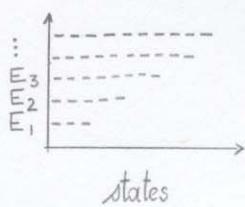
$$p_i \propto \Omega_R (E_{tot} - E_i) \quad \text{with} \quad E_{tot} = E_R + E_i = \text{constant}$$

$$\Omega_R (E_{tot} - E_i) = \exp\left[S_R (E_{tot} - E_i) \div k_B \right] \qquad \text{(closed system)}$$

$$S_R = S_R(E_{tot}) - E_i \left. \frac{\partial S_R}{\partial E} \right|_{E_{tot}} \qquad \text{because} \quad E_i \ll E_{tot}$$

$$\frac{1}{T} \equiv \left. \frac{\partial S_R}{\partial E} \right|_{E_{tot}} \quad \text{defines temperature} , \quad \text{hence} \quad S_R = S_R^{(tot)} - \frac{E_i}{T}$$

$$\boxed{ p_i \propto \exp\left(\frac{-E_i}{k_B T} \right) = \exp(-\beta E_i) } \qquad \text{Boltzmann's factor}$$

after normalization $\qquad p_i = \frac{1}{Z} \exp(-\beta E_i) \quad \text{with} \quad \beta \equiv \frac{1}{k_B T}$

$$\boxed{ Z = \sum_i \exp(-\beta E_i) } \qquad \text{partition function}$$

- free energy G

start with partition function $\quad Z = \sum_i e^{-\beta E_i} = e^{-\beta F} \qquad \text{or} \qquad \ln Z = -\beta F$

Helmholtz free energy $\quad F = -k_B T \ln Z \qquad \text{(in joules J)}$

we can now write $\quad p_i = \frac{1}{Z} \exp(-\beta E_i) \quad \text{(in log:)} \quad \ln p_i = -\beta E_i + \beta F$

and consider a particular state whose energy is equal to the average of the system A:

$$\bar{E} = \langle E \rangle$$

$$E_i = \bar{E} \implies \ln p(\bar{E}) = -\beta \bar{E} + \beta F = \overline{-\beta E + \beta F} = \overline{\ln p_i}$$

$$\langle \ln p_i \rangle = -\beta \bar{E} + \beta F = \sum_i p_i \ln p_i$$

Let Ω_i be the number of microstates with energy E_i ; $\quad p_i = \frac{1}{\Omega_i} = \exp\left(-\frac{S_i}{k_B} \right)$

from $\ln p_i = -\frac{S_i}{k_B}$, we get $\quad \sum_i p_i \ln p_i = -\frac{1}{k_B} \sum_i p_i S_i$

and $\quad \bar{S} = \sum_i p_i S_i \quad$ measurable entropy of the system

$$\frac{1}{k_B T} (-\bar{E} + F) = -\frac{\bar{S}}{k_B} \qquad \text{or} \qquad F = E - TS$$

Molecular mechanics - 3.

$$\boxed{F = E - TS}$$

F: free energy of the system
E: internal energy of the system
S: entropy of the system
T: temperature of the reservoir

- Meaning:

- $F = E - TS$ "free" energy = maximum work the system can do
  let W be the work done by the system $\Delta W \leqslant -\Delta F$
  with $\Delta F \leqslant 0$

- $e^{-\beta F} = \sum_i e^{-\beta E_i}$ is a "representative Boltzmann's factor"
  it lumps up all Boltzmann's factors of the system.

protein conformation
electronic states :
nuclei quarks :

$e^{-\beta F'}$
$e^{-\beta F}$
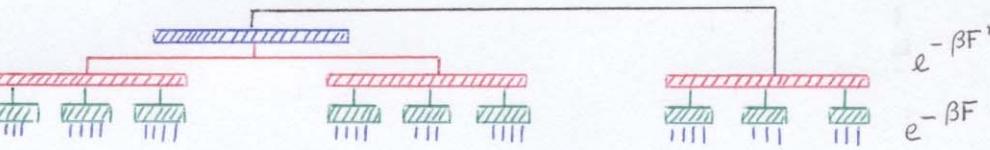
free energy enough to describe the whole system
at many time, length, energy scales

Gibbs free energy $G = H - TS$

- ## Equipartition theorem
in a typical equilibrium situation :   fluctuations near (free) energy at equilibrium
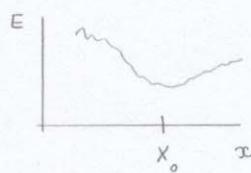
$$E(x) = E(x_0) + (x-x_0) \left. E' \right|_{x=x_0} + \frac{(x-x_0)^2}{2} \left. E'' \right|_{x-x_0}$$

0 at min

$$E(x) = E(x_0) + a(x-x_0)^2$$

$$p(x) = \frac{\exp(-\beta a(x-x_0)^2)}{Z} \quad \text{with} \quad Z = \int_{-\infty}^{+\infty} dx \, e^{-\beta a (x-x_0)^2}$$

average energy :
$$\langle E \rangle = \int dx \, p(x) \, a(x-x_0)^2 = \frac{1}{Z} \int dx \, a(x-x_0)^2 \exp\left[-\beta a(x-x_0)^2\right]$$

$$= \frac{k_B T}{2} \quad \text{using the Gaussian integral} \quad \int_{-\infty}^{+\infty} e^{-ay^2} dy = \sqrt{\frac{\pi}{a}}$$

energy per degree of freedom $= \frac{RT}{2}$ with $R = \mathcal{N}_A \, k_B$

if N degrees of freedom $\boxed{\langle E \rangle = \frac{N}{2} k_B T}$

ex: $N = 3 + 2$
$\bar{E} = \frac{5}{2} k_B T$

Molecular mechanics – 4.

– Generalized force

$$E(x) = \dots \quad , \quad x \text{ is called the control parameter}$$

if the system does work W resulting in $x \to x + dx$


$E' < 0, W > 0$

$$dE = dx \cdot \frac{\partial E}{\partial x} = -W$$

$$= -f \, dx$$

$$\boxed{f = -\frac{\partial E}{\partial x}}$$
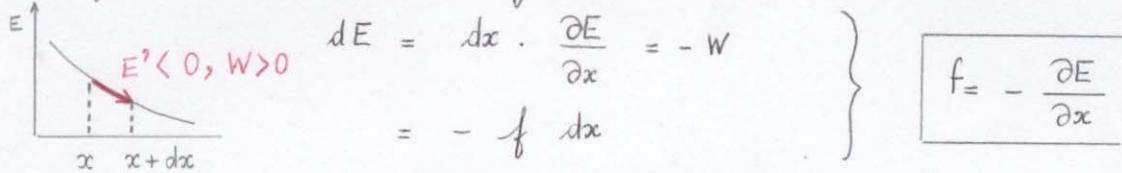
## – So, in a two-state equilibrium

work done ON the system by an external force $f_{ext} \, \Delta x$

"   "   BY   "   " $= -f_{ext} \, \Delta x$
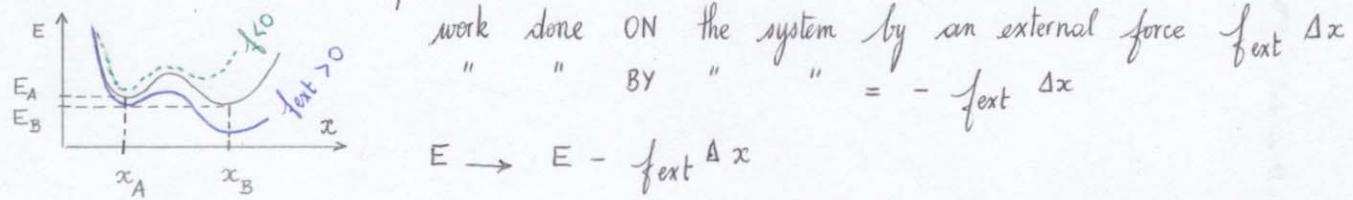
$$E \to E - f_{ext} \, \Delta x$$

equilibrium constant $K_{eq} = \dfrac{[B]}{[A]} = e^{-\beta(\Delta E - f_{ext}\Delta x)}$ <u>force tilts energy profile</u>

in auditory hair cells

tension

F_{ext}

state A   B
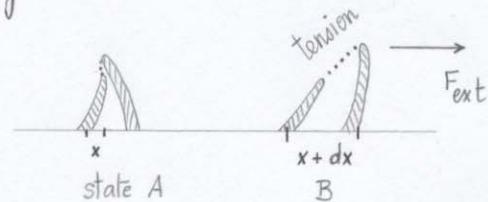
under force, "open state B" is made much more likely.

SUMMARY

statistical mechanics link micro to macro.

Molecular mechanics : ab initio foundation   Ju Li

see Powerpoint handout by instructor.

– <u>quantum mechanics</u> do matter to study hydrogen bonds

haemoglobin / porphyrin interactions need to be understood to parameterize / approximate systems.

– why are electrons treated as quantum objects?

electrons $\{\underline{x}_i\}$ $i = 1 .. n$ of mass m_e

ions $\{\underline{\bar{x}}_I\}$ $I = 1 .. N$ $M_I \approx A_I \, m_p \, , \quad m_p \approx 2000 \, m_e$

$A_I = 12$ and $Z_I = 6$

$$V_{iI} = \frac{e^2}{|x_i - X_I|}$$ electrostatic interaction between electron & ion

de Broglie wavelength $\lambda = \frac{h}{|\bar{p}|}$, how large is it with respect to electrons?

kinetic energy $\frac{1}{2} m_e v_e^2 = \frac{P_e^2}{2 m_e} \sim 13.6 \text{ eV} \sim \frac{P_I^2}{2 m_I}$

for C – C bond $p_e \sim \sqrt{2 m_e * 13.6 \text{ eV}}$

$\lambda_e = 3.3 \text{ Å}$

if $I \equiv C$ $\lambda \sim 0.02 \text{ Å}$ }   position fluctua-

$I \equiv H$ $\lambda \sim 0.08 \text{ Å}$ " smearing factor "   } tions, quantum,

electrons's λ permeates through several bonds   } (not thermal !)

↳ structural units

– <u>Born- Oppenheimer approximation</u> : ions are immobile

electrons fluctuate

$$V_{BO}\left(\{\overline{X_I}\}_{I= 1.. N}\right) \equiv \min_{\Psi(\{x_i\})} E_{tot}\left(\{x_i\}, \{X_I\}\right)$$

energy function $\Psi(\{x_i\}) = \Psi_G\left(\{x_i\} | \{X_I\}\right)$ electrons in ground state

addendum : Newton's law $M_I \ddot{X}_I = - \dfrac{\partial V_{BO}(\{X_I\})}{\partial X_I} = \underline{F}_I$

steepness of energy landscape = force

when does the BO approximation break down?

· if optically excited
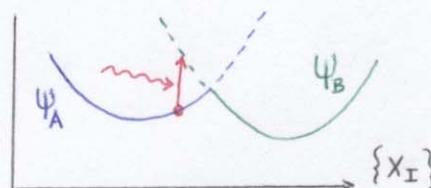
light ⤳ can make
electrons leave the
least energetic state

· during diabatic electron transfer (Marcus) : you stay on Ψ_{cst} most of the

time

when does the addendum break down?

· at low temperature, or for low - weight ions (hydrogen) = quantum too

quantum fluctuation ≈ thermal fluctuation at low T

Eigenfunction problem $\hat{H} \Psi_G (x_1, \ldots, x_n) = V_{BO} \Psi_G (x_1, x_2, \ldots, x_n)$

Hamilton operator $\hat{H} = \sum_{i=1}^{n} \frac{-\hbar^2 \nabla_i^2}{2 m_e} + \sum_{i \neq j} \frac{e^2}{2 |x_i - x_j|}$

$$- \sum_{i, I} \frac{Z_I e^2}{|x_i - X_I|} + \sum_{I \neq J} \frac{Z_I Z_J e^2}{2 |X_I - X_J|}$$

$\{$ solve with only parameters being Plank's constant $\hbar$, m_e, e, Z_I
$\quad$ ab initio approach

but <u>information explosion</u>: how to store $\Psi_G (x_1, y_1, z_1, \ldots, x_n, y_n, z_n)$

if 1 variable $\quad$ 10 spline points of 8 Bytes each = 80 bytes

if 60 variables $\quad 10^{60}$ points

(20 electrons) $\quad$ or 10^{52} CDs of 100 MB each

!!!
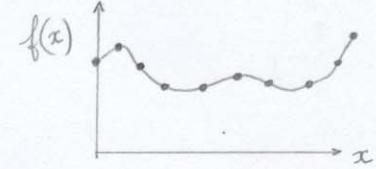
fortunately, symmetry relations, and redundancy in spline functions

quantum chemistry $\quad$ (Pople)

density - functional theory (Kohn) $\Big\}$ address this issue of huge numbers.

$\quad\quad\quad\quad$ error in bond energy calculations $\downarrow$

$\quad\quad \hookrightarrow$ advances in theoretical methodology

$\quad\quad\quad\quad$ computational techniques

$\quad\quad\quad\quad$ computing technology.

- <u>Hartree - Fock theory</u>

approximation of wavefunction by Slater determinant

from $3n$ - dimensional to n 3 - dimensional problem

add electron spin indication : $\uparrow$ or $\downarrow$ only takes 1 bit extra.

$$\Psi \approx S = \frac{1}{\sqrt{n!}} \begin{vmatrix} \tilde{\Psi}_1(x_1) & \cdots & \tilde{\Psi}_n(x_1) \\ & \vdots & \\ \tilde{\Psi}_1(x_n) & \cdots & \tilde{\Psi}_n(x_n) \end{vmatrix}$$

decompose the total energy into sum of pairs

$\sum_{i, j} \Delta E (\Psi_i, \Psi_j)$ $\quad$ with $\quad \Delta E = E_{Hartree} - E_{exchange}$

$E_H (\tilde{\Psi}_i, \tilde{\Psi}_j) \equiv e^2 \iint dx\, dx' \frac{[\tilde{\Psi}_i^*(x) \tilde{\Psi}_i(x)][\tilde{\Psi}_j^*(x') \tilde{\Psi}_j(x')]}{|x - x'|}$ $\quad$ classical

$E_{exchange}$ is quantum !

the $E_{exchange}$ stabilizes occupation of same spin wavefunction
(whereas spin doesn't matter in $E_{Hartree}$)
in fact, use several Slater determinants   and minimize all configurations
                                                                 by optimizing their coefficients

Correlation energy ≡ energy reduction
Configuration interaction (CI) is formally exact, but limited to 10 - 20-atom problems

- Hence <u>density functional theory</u> (DFT)   both exchange & correlation accounted for.
Hohenberg and Kohn (1964) $V_{BO} (\{X_I\}) = V_{BO}$
                                                $\updownarrow$ $\nearrow \rho$
                                $v(\underline{x}) = \sum_I \dfrac{e^2 Z_I}{x_i - X_I}$ $\longleftrightarrow$ $\Psi_G (x_1, ..., x_n)$

two non-trivially different ion-electron potentials   cannot give rise to the same density
and vice versa ⟹   one-to-one marking   } 8×10^{60} bytes ⟷ $8 * 10^3$ bytes
                        compression

Kohn and Sham (1965) : fictitious noninteracting electron system
        $V_{BO} (\rho(x)) = V_{independent} (\rho(x)) + V_{Hartree} (\rho(x)) + V_{exchange} (\rho(x))$
                                                                                                                        correlation
        this is not a "single-determinant" method like H-F.

        → approximate* $V_{exch-corr.}$ by a density with Monte Carlo quantum method.
                                                derived for homogeneous electron gas.
        * LDA: local density approximation
          GGA: generalized gradient approximation.   }   limitations = heterogeneity

hybrid functionals mix DFT and Hartree-Fock (but not really ab initio ?!)
- Semi-empirical electronic-structure methods
                use ab initio or experimental information to fit intrinsic electronic quantities

- <u>interatomic potential</u> / force field method: parameterized V_{BO} surface
                                                O(N) linear scaling algorithms.