10. MOLECULAR SIMULATIONS

10.1. Introduction

- Difficulty in applying statistical mechanics: Despite the existence of a powerful, rigorous formalism for the calculation of macroscopic properties given the Hamiltonian of a system, an analytical solution to this formalism is impossible for most material systems of practical interest.

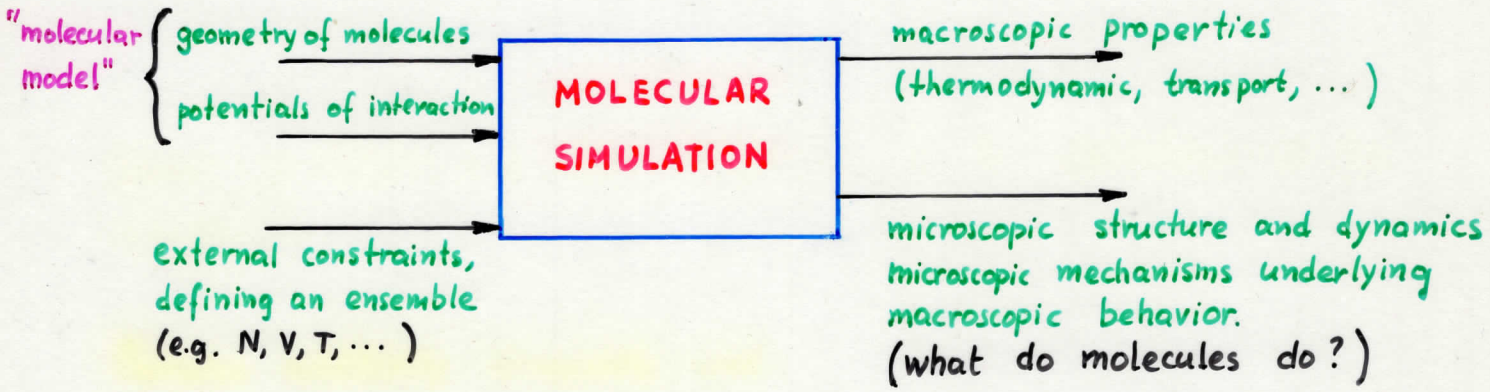
- Analytical theories (e.g. integral equation theories, perturbation theor Bragg-Williams approximation for the lattice), designed to provide closed-form solutions, must necessarily introduce approximations into the statistical mechanical treatment.

- Alternative Approach: Numerical solution of the full statistical mechanics given a model of molecular geometry and energetics.
Simulations can provide, in principle, "exact" results (subject only to numerical error).
In practice, simulations may invoke approximations, which, however, are much less drastic than those invoked by analytical theories.

- Main problems in applying simulations to real-life materials systems:

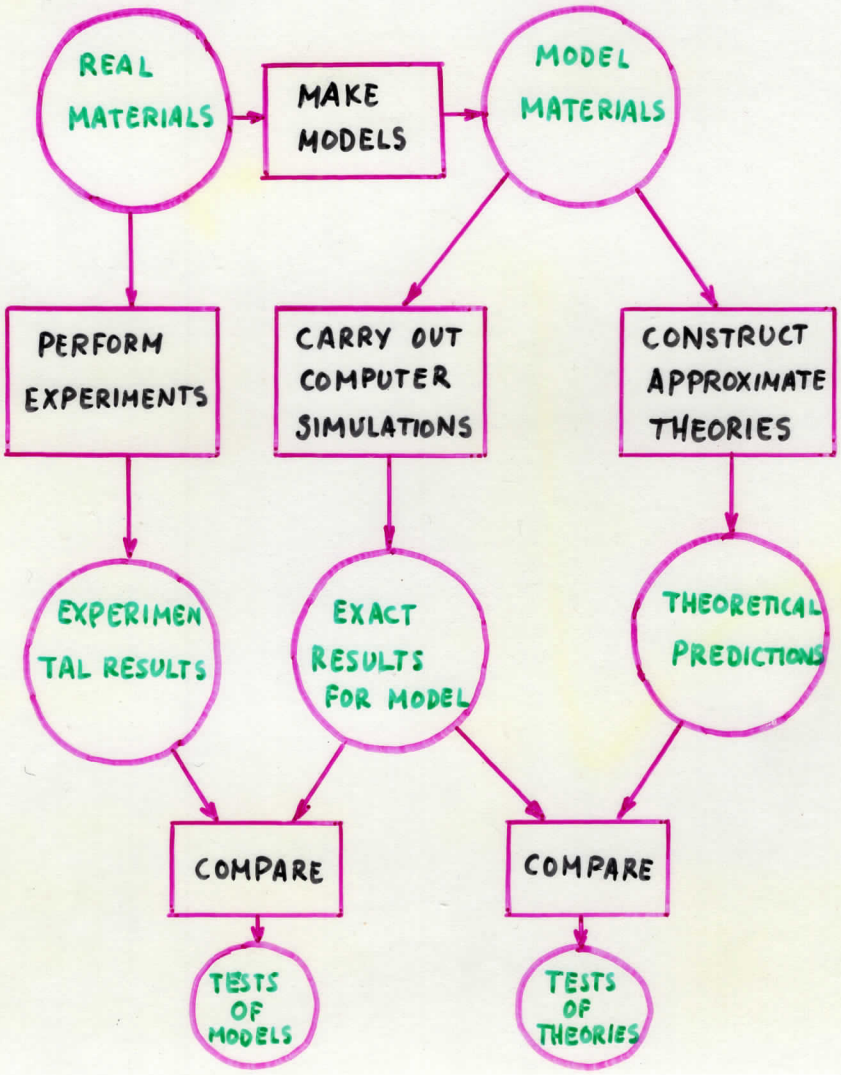
  - Our knowledge of interaction potentials is limited. (Quantum mechanics can help here!)
  - Required computer times are very long (use of supercomputers or dedicated workstations).

Molecular Simulations are Computer Experiments.



Connection between experiment, theory, and computer simulation.

(after Allen & Tildesley)



10.2. Construction of a Molecular Model

10.2.1. Molecular Representation.

Level of detail may vary, depending on the objectives of the investigator.

Pairwise additivity frequently used.

$$\mathcal{V}(\underbrace{\underline{r}_1, \ldots, \underline{r}_N}_{\text{coordinates of interaction sites (e.g. atomic coordinates)}}) = \sum_{i<j} \sum \mathcal{V}_2(\underline{r}_i, \underline{r}_j) + \sum_{i<j<k} \sum \sum \mathcal{V}_3(\underline{r}_i, \underline{r}_j, \underline{r}_k) + \cdots$$

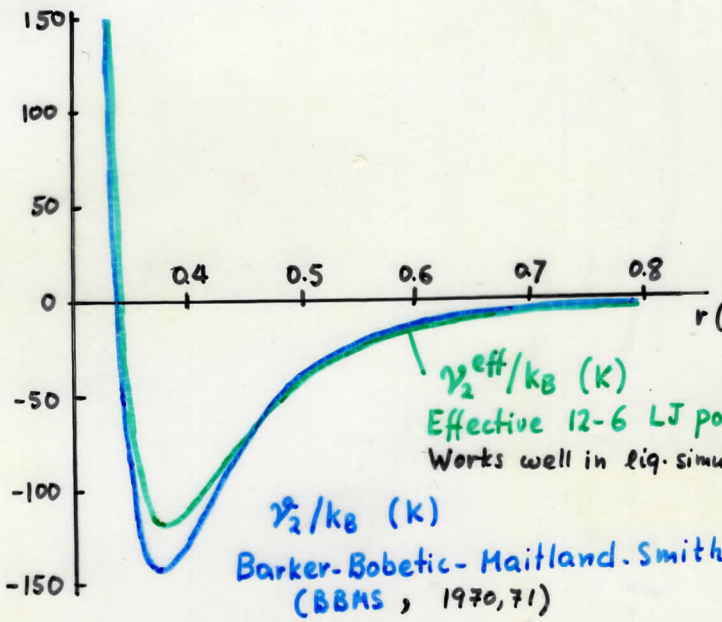
pair potential — three-body potential

Approximate $\mathcal{V}$ as a sum of <u>effective</u> pair potentials:

$$\mathcal{V}(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_N) = \sum_{i<j} \sum \mathcal{V}_2^{\text{eff}}(\underline{r}_i, \underline{r}_j)$$

Three-body term is significant at liquid densities, and even more so in the solid state. (10% of lattice energy of solid argon is due to higher-than-pair interactions).

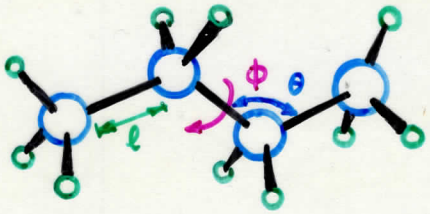
In practice, three-body and higher terms are included in "effective" pair potential, which may thus show some dependence on ρ, T.



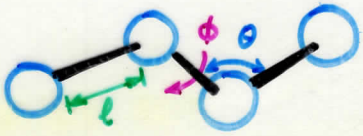
Very accurate, based on molecular beam scattering, spectroscopy of argon dimer, inversion of T-dependence of 2nd virial coefficient, solid state properties, theoretical calculations.

Molecules with no pronounced separation of charge can be represented as sets of LJ interaction sites.

e.g, alkanes:



"explicit representation"
Each C, H atom treated as an individual LJ interaction site.



"united atom representation"
Each CH_3, CH_2 group treated as a single interaction site.

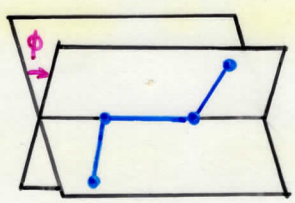
Bond potential: $v_b(\ell) = \frac{1}{2} k_b (\ell - \ell^\circ)^2$, or more elaborate forms.
(ℓ° ← equil. bond length)

It is satisfactory to consider bonds as fixed ($k_b \to \infty$), unless vibrational spectra are desired.

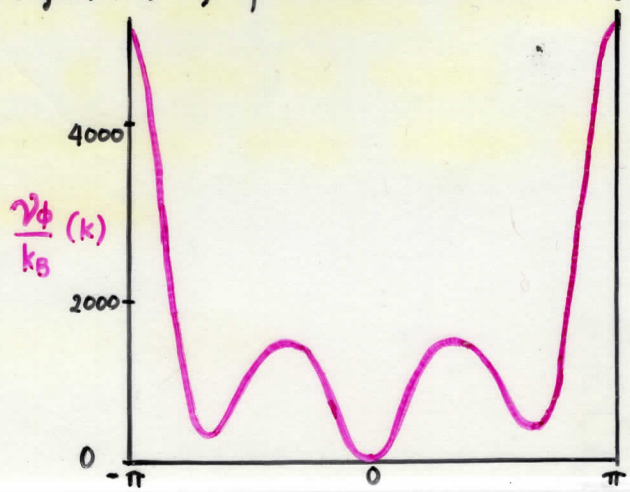
Bond angle potential: $v_\theta(\theta) = \frac{1}{2} k_\theta (\theta - \theta^\circ)^2$, or more elaborate forms
(θ° ← equil. bond angle)

Bond angles frequently taken as fixed. ($k_\theta \to \infty$)
Bond angle variation is thought to be important for the dynamics of flexible molecules. (MD).

Torsional potential:



dihedral (torsion) angle ϕ



(Marechal & Ryckaert, 198

A list of atom-atom LJ interaction parameters that can be used to construct crude first-guess intermolecular potential functions is given by Allen & Tildesley, p 21.

Lorentz-Berthelot combining rules can be used to estimate cross-interaction parameters from self-interaction parameters:

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}) \quad , \quad \epsilon_{ij} = (\epsilon_{ii}\,\epsilon_{jj})^{1/2} \qquad (10.1)$$

Potential parameters can be refined by comparison to experimental data, e.g.
- second virial coefficients $B(T)$
- internal energies at given p, T; ΔH^{vap}; solubility parameter
- pressure at given ρ, T
- lattice energy of crystal
- experimental crystal structure, lattice parameters.

Molecules with expressed separation of charge (polar molecules) are described in terms of LJ interaction sites and partial charges.

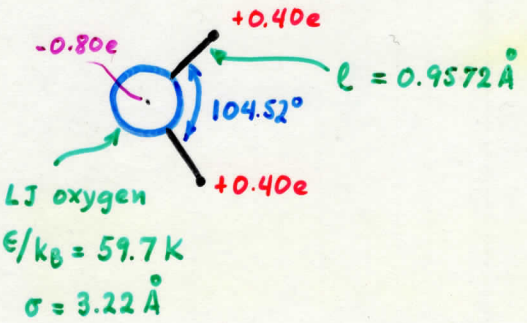
Values of partial charges may be based on:
- experimentally measured dipole moments (quadrupole moments, ...)
- electronic density distribution from ab initio quantum mechanics. (e.g. Mulliken population analysis of Hartree-Fock results).
- fitting ab initio calculated intermolecular energy between two molecules at representative orientations.

Partial charges interact via Coulomb forces:

$$v^{zz}(r_{ij}) = \frac{z_i z_j}{4\pi\epsilon_0 r_{ij}} \qquad (10.2)$$

"transferable intermolecular potentials"

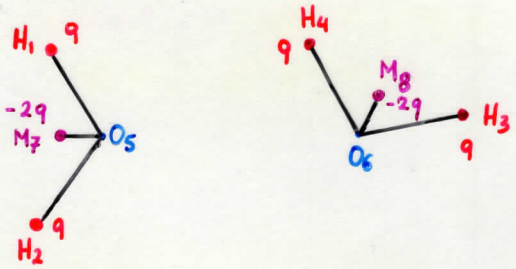
Example: "TIPS" model of water (Jorgensen, 1981)



H_2O represented simply as a set of one LJ center (oxygen) and three partial charges.

Note: Model dipole moment = 2.25 D
expt. dipole moment = 1.85 D

Matsuoka-Clementi-Yoshimine water model (J. Chem. Phys. 1976, 64, 1351)
Fit to ab initio calculation of potential hypersurface of water dimer based on configuration-interaction (CI) method.



Dimer energy:

$$\epsilon = q^2\left(\frac{1}{r_{13}} + \frac{1}{r_{14}} + \frac{1}{r_{23}} + \frac{1}{r_{24}}\right) + \frac{4q^2}{r_{78}}$$

$$-2q^2\left(\frac{1}{r_{18}} + \frac{1}{r_{28}} + \frac{1}{r_{37}} + \frac{1}{r_{47}}\right) + a_1 \exp(-b_1 r_{56})$$

$$+ a_2\left[\exp(-b_2 r_{13}) + \exp(-b_2 r_{14}) + \exp(-b_2 r_{23}) + \exp(-b_2 r_{24})\right]$$

$$+ a_3\left[\exp(-b_3 r_{16}) + \exp(-b_3 r_{26}) + \exp(-b_3 r_{35}) + \exp(-b_3 r_{45})\right]$$

$$- a_4\left[\exp(-b_4 r_{16}) + \exp(-b_4 r_{26}) + \exp(-b_4 r_{35}) + \exp(-b_4 r_{45})\right]$$

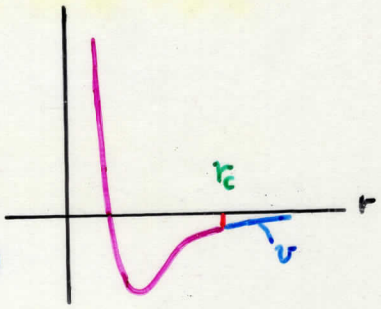
10 potential parameters:

q^2, a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4, $R(O_5\text{-}M_7)$

10.2.2. Summation of the total potential energy function

$$\mathcal{V}(\underset{\sim}{r}_1, \ldots, \underset{\sim}{r}_N) = \sum_{i<j} \sum v(r_{ij})$$

(effective) pair potential, e.g. LJ.



Potentials cannot be summed to infinite distance!

It is practical to use a finite range modification of $v(r)$

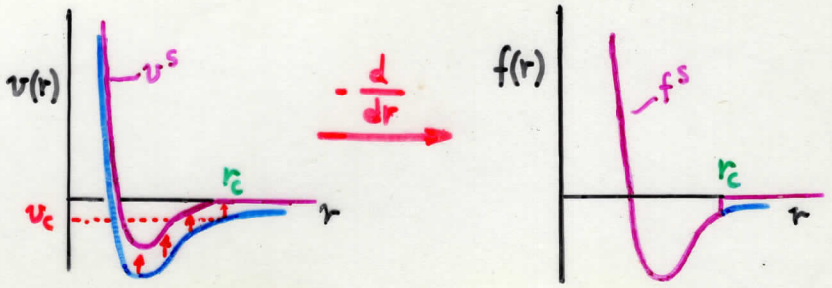
Cutoff distance r_c (e.g. $r_c \simeq 2.5\sigma$).

PROBLEM: Cutoff generates infinite forces at r_c. Energy conservation deteriorates in M

SOLUTION:

a. Shifted potential

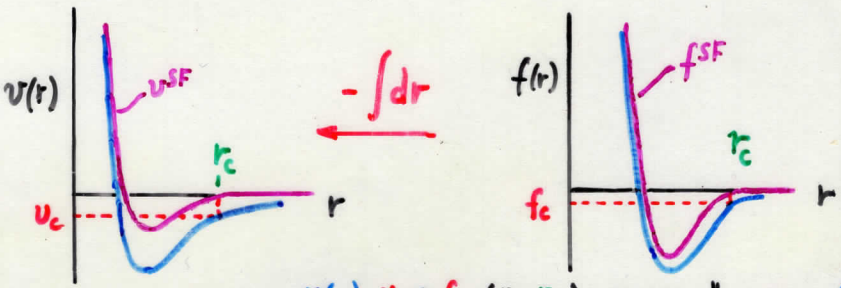
- does not affect forces
- PROBLEM: force discontinuous at r_c (instabilities)



$$v^S(r) = \begin{cases} v(r) - v_c & r \leq r_c \\ 0 & r > r_c \end{cases}$$

b. Shifted force potential

- discontinuity only in the first derivative of force.
- PROBLEM: Interaction has been changed. Thermodynamics of original system recovered by perturbation scheme.

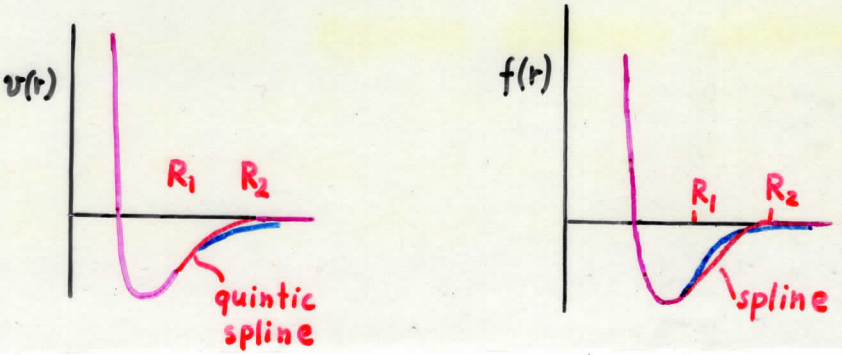


$$v^{SF}(r) = \begin{cases} v(r) - v_c + f_c (r - r_c), & r \leq r_c \\ 0 & r > r_c \end{cases} \qquad f^{SF}(r) = \begin{cases} -\frac{dv}{dr} - f_c \\ 0 \end{cases}$$

c. Other Modifications

(continuous v, $\frac{dv}{dr}$, $\frac{d^2v}{dr^2}$)

e.g. $R_1 = 1.45\sigma$
$R_2 = 2.30\sigma$



The contribution to $\mathcal{V}$ from the omitted "tails" of the pair potential function is calculated by direct integration:

$$\mathcal{V}^{tails} = 2\pi N\rho \int_{r_c}^{\infty} \left[u(r) - u^{modified}(r)\right] g(r)\, r^2 dr$$

($g(r)$: usually $\simeq 1$ at $r \gg r_c$)

($u^{modified}(r)$: for a simple potential cutoff, this is zero)

(r_c: distance above which $u(r)$ and $u^{modified}(r)$ differ.)

NOTE: Potential truncation is a poor approximation for long-range potentials (eg. Coulomb, $\propto 1/r$). Special summation techniques (eg. Ewald summation) must be used in this case.

10.2.3. Model System Size and Periodic Boundary Conditions.

Number of molecules in a common simulation: $10 \leq N \leq 10000$

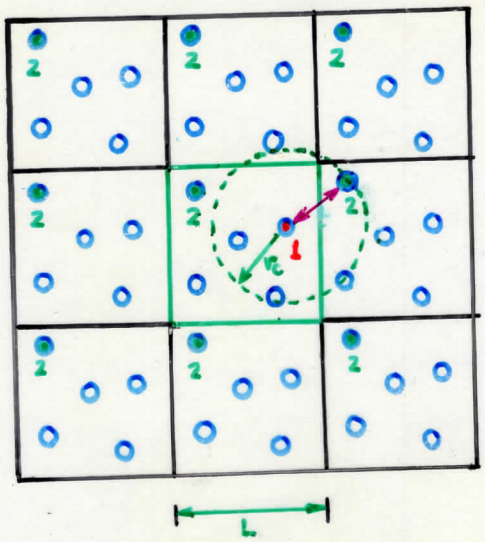
Limitations on system size imposed by:

- storage requirements
- speed: Force/potential calculation loop requires time $\mathcal{O}(N^2)$

System size-independence of results must be confirmed.

For low-molecular weight liquids, simulation averages are very weakly dependent on system size for $N \geq 100$.

Except for cases where the simulation of a very small amount of material is desired (droplet, microcrystal), PERIODIC BOUNDARY CONDITIONS are used. In this way, the strong surface effects that would be present due to the small size of the sample are avoided. (For 1000 molecules arranged in a 10×10×10 cube, 488 molecules appear on cube faces!)

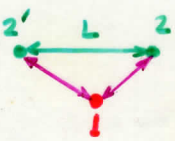


Primary box envisioned as surrounded by images of itself (26 images around central box in 3-d).
Primary box plays a role of a unit cell.
Whenever, in the course of simulation, an atom exits the box, an identical atom enters through opposite face.
Periodicity does not perturb the thermodynamics significantly, if potential is short-ranged (For LJ atoms, if $L \simeq 6\sigma$, a particle doesn't "sense" the periodic geometry.)

Periodic boundary conditions cut off long-wavelength fluctuations (density waves with wavelength $> L$). Thus, simulation near a critical point is a very difficult task. PBC's may also introduce spurious info. on time correlations for times $\gtrsim \frac{L}{c}$. System size-independence of results must always be confirmed.
(c ↑ sound velocity)

Consider two atoms, 1 and 2, in the primary box.
In three dimensions, there are 27 images of 2 in the primary box and the surrounding boxes that could interact with 1

If $L/2 \geq r_c$ (potential cutoff distance (potential 0 beyond r_c)), then at most one of the images of 2 can interact with 1.



(If 1 interacted with two images 2, 2', then $12' + 12 < r_c + r_c = 2r_c$ but by triangle inequality $12' + 12 \geq L \geq 2r_c$, which is not possible)

The image of 2 lying closest to 1 is called minimum image.
When $L/2 \geq r_c$, the simulation obeys the MINIMUM IMAGE CONVENTION.

Simple code for calculation of minimum image distance:

```
DO      I = 1, N-1
   RXI = RX(I)
   RYI = RY(I)
   RZI = RZ(I)
   DO    J = I+1, N
      RXIJ = RXI - RX(J)
      RYIJ = RYI - RY(J)
      RZIJ = RZI - RZ(J)
      RXIJ = RXIJ - BOXL * ANINT(RXIJ/BOXL)
      RYIJ = RYIJ - BOXL * ANINT(RYIJ/BOXL)
      RZIJ = RZIJ - BOXL * ANINT(RZIJ/BOXL)
      R1JSQ = RXIJ**2 + RYIJ**2 + RZIJ**2
      IF (RIJSQ.LT. RCUTSQ) THEN
      ... compute I-J interaction ...
      ... accumulate energy and forces...
      ENDIF
END DO
```

actual distance vector J to I

minimum image vecto J to I

screening of squared minimum image distance against squared cutoff distance.

Note: $\text{ANINT}(x) = \begin{cases} +1, & \text{if } 0.5 < x < 1 \\ 0, & \text{if } -0.5 < x < 0.5 \\ -1, & \text{if } -1 < x < -0.5 \end{cases}$

Note: IF statement in innermost loop prevents vectorization. On a CRAY, can use "conditional vector merge" statement to eliminate IF (see Allen and Tildesley, p 32).

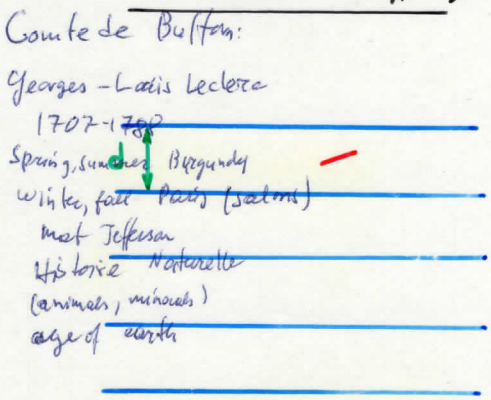
10.3. MONTE CARLO METHODS

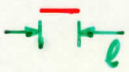
Ref. Allen and Tildesley, p 110 (Chapter 4)
Kalos, M.H.; Whitlock, P.A. Monte Carlo Methods, Vol. I : Basics, Wiley, 198

Name Monte Carlo applied to a class of mathematical methods first by scientists working on the development of nuclear weapons in Los Alamos, in 1940s. Essence of the method is the invention of games of chance, whose behavior and outcome can be used to find the solution to a mathematical problem.

10.3.1. The roots of stochastic experiments: "Hit or Miss Monte Carlo".

G. Comte de Buffon, 1777 discovered a theorem in geometrical probability.





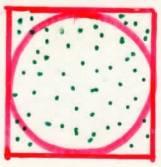
A needle of length ℓ is thrown randomly onto a set of equally spaced parallel lines, d apart ($d > \ell$). The probability that the needle will cross a line is $\frac{2\ell}{\pi d}$

Lazzerini, 1901: Performed a stochastic experiment, by spinning round and dropping a needle $N_{trials} = 3407$ times and counting the number of times N_{hits} that the needle crossed a line. He used this result to estimate π:

$$\pi = \frac{2\ell}{d} \frac{N_{trials}}{N_{hits}}$$

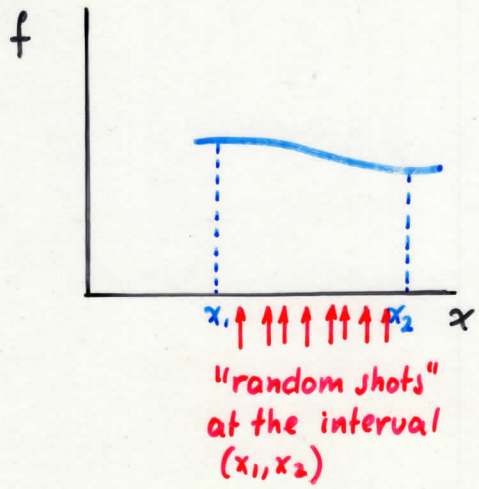
A machine can now estimate π to 3 significant figures by performing 10^7 trials on the Buffon needle experiment. Error of estimate is $O(N_{trials}^{-1/2})$

Another "hit and miss" experiment:
Estimate π by counting the raindrops collected in a circular pan inscribed in a square pan.



$$\pi = 4 \cdot \frac{N_{hits}}{N_{trials}}$$

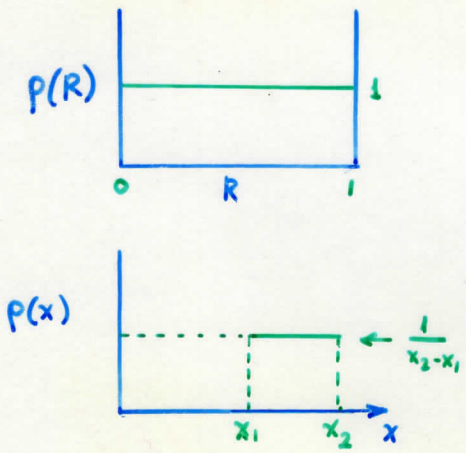
10.3.2. Monte Carlo as an Integration Method.



Problem: Compute the integral

$$I = \int_{x_1}^{x_2} f(x)\,dx$$

One Solution: "Monte Carlo Sample Mean Integration"



Generate a random number $R_m \in (0,1)$

↓

Form $x_m = x_1 + R_m (x_2 - x_1)$

↓

Calculate $f(x_m)$

↓

Accumulate sum $\sum_m f(x_m)$

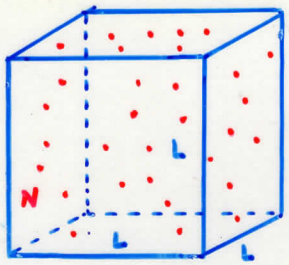
After performing the experiment N_{trials} times, one has an estimate of the integral as:

$$I \simeq \frac{(x_2 - x_1)}{N_{trials}} \sum_{m=1}^{N_{trials}} f(x_m)$$

(Theoretical justification: Mean value theorem).

By applying this method to the function $f(x) = (1-x^2)^{1/2}$ between $x_1 = 0$ and $x_2 = 1$, one can obtain π to within 4 significant figures in 10^7 trials.

Sample mean integration cannot compete with conventional numerical integration methods (Simpson, Gauss) in one or two dimensions. It becomes superior to them when one has to perform multidimensional integrations.



Example:

Calculate the configurational integral

$$Z = \int d^{3N}r \underbrace{\exp\left[-\beta \mathcal{V}(\underline{r}^N)\right]}_{f(\underline{r}^N) \equiv f(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_N)}$$

for N molecules in volume V at temperature T.
This is a $3N$-dimensional integral.

For $N = O(10)$, an estimate of Z can be obtained as

$$Z \simeq \frac{V^N}{N_{trials}} \sum_{m=1}^{N_{trials}} \exp\left[-\beta \mathcal{V}\left(\underbrace{\underline{r}_1^{(m)}, \ldots, \underline{r}_N^{(m)}}_{\text{a randomly selected point in the } 3N\text{-dimensional configuration space. (i.e. } N \text{ triples of random numbers within } (0, L))}\right)\right]$$

A Simpson technique would require evaluating $f(\underline{r}^N)$ at all nodes of a regular grid through the configuration space. If 10 points (nodes) per coordinate are used, this would entail 10^{3N} function evaluations. With Monte Carlo integration, one can get a reasonable estimate for N_{trials} much less than that.

Such a direct calculation of Z by Monte Carlo integration becomes impractical for $N = O(100)$.

2.3.3. Importance Sampling in the calculation of Ensemble Averages.

We have seen that important thermodynamic properties can be expressed as ensemble averages, of the form

$$\langle f \rangle_{ens} = \frac{\int d^{3N}r \; f(\underset{\sim}{r}^N)\, \rho^{ens}(\underset{\sim}{r}^N)}{\int d^{3N}r \; \rho^{ens}(\underset{\sim}{r}^N)} \qquad (10.3)$$

e.g., with $\rho^{ens}(\underset{\sim}{r}^N) = \rho^{NVT}(\underset{\sim}{r}^N) = \exp[-\beta \mathcal{V}(\underset{\sim}{r}^N)]$

$f = \mathcal{V} \rightarrow \langle f \rangle_{NVT} = U^{ex}$, excess internal energy, i.e. $U(\rho,T) - U^{ig}(\rho,T)$

$f = \frac{1}{3}\sum_i \underset{\sim}{r}_i \cdot \underset{\sim}{F}_i^{int} \rightarrow \langle f \rangle_{NVT} = PV - Nk_BT$ virial theorem for pressure

$f = \exp(-\beta \mathcal{V}_{test}) \rightarrow \langle f \rangle_{NVT} = e^{-\beta\mu^{ex}}$, excess chemical potential, i.e. $\mu(\rho,T) - \mu^{ig}(\rho,T)$

One could try to estimate $\langle f \rangle$ by applying Monte Carlo integration to both the numerator and the denominator, i.e.

$$\langle f \rangle \simeq \frac{\sum_{m=1}^{N_{trials}} f(m)\, \rho^{ens}(m)}{\sum_{m=1}^{N_{trials}} \rho^{ens}(m)}$$

where m labels randomly generated configurations (points in configuration space)

This, however, would be <u>very inefficient</u>.

<u>Why</u>? Some regions of the multidimensional space (in NVT ensemble, the regions of low $\mathcal{V}$) contribute <u>much more</u> to the average than others. If we pick the sampled points by shooting randomly at configuration space, we have an overwhelming probability of hitting states (i.e. configurations) with very high energy, which would contribute almost nothing to the average.

Instead, we should sample our configuration space nonuniformly, so that more probable states (low energy configurations) are sampled more frequently than less probable states (high energy configurations). Such a sampling is called IMPORTANCE SAMPLING.

In fact, we should sample states so that
Frequency of sampling a state $\propto \rho^{ens}$ at that state.

If we manage to do this, then calculating $\langle F \rangle$ amounts to taking a simple arithmetic average over the sampled microstates.

$$\langle F \rangle = \frac{1}{N_{trials}} \sum_{m=1}^{N_{trials}} F(m). \qquad (10.4)$$

An ingenious algorithm for sampling a multidimensional space according to a (nonuniform) probability distribution, and for calculating averages with respect to that distribution, has been introduced by Metropolis et al. Before describing the Metropolis algorithm we will say a few things about Markov chains.

10.3.4. Markov Chains.

Markov chain: A sequence of trials (stochastic process) that satisfies two conditions:

a. Outcome of each trial belongs to a finite set of outcomes, called the state space.

b. Outcome of each trial depends only on outcome of immediately preceding trial. (memory goes only one step back).

Central in Markov chains: Concept of **transition probability**.

P_{nm} = Probability that a trial produces the state n, given that the previous trial has resulted in state m.

(n: to, m: from)

P_{nm} depends only on the values m, n but is independent of where within the sequence the considered trial lies. (i.e. on "time", or number of trials).

By collecting values of P_{nm} for all possible n and m we can form a transition probability matrix. This matrix contains all information about the "dynamics" governing the "evolution" of the Markov chain.

Example (after Allen + Tildesley):

The reliability of our computer follows a certain pattern:

- If the computer is up and running on one day, it has a 60% chance of running correctly on the next day.
- If it is down, it has 70% chance of also being down the next day.

Two states in state space: U, D.

Transition probabilities: $P_{UU} = 0.60$, $P_{DU} = 0.40$

(P_{UU}: to ↑ U, from ↖ U)

$P_{DD} = 0.70$, $P_{UD} = 0.30$

Transition probability matrix:

$$\underline{\underline{P}} = \begin{array}{c} \\ U \\ D \end{array} \overset{\begin{array}{cc} U & \text{from } D \end{array}}{\begin{bmatrix} 0.60 & 0.30 \\ 0.40 & 0.70 \end{bmatrix}}$$

(to: ↓ U, D)

Important note: $\sum_n P_{nm} = 1 \quad \forall\, m$ (10.5)

Columns of transition probability matrix sum to one (the system must be in some state at the end of the next trial).
This property makes $\underline{\underline{P}}$ a STOCHASTIC MATRIX.

Suppose that, when we receive it, the computer has 50% probability of being up. We can then write for the state probabilities in this initial state:

$$\underset{\sim}{q}^{(0)} = \begin{bmatrix} q_U^{(0)} \\ q_D^{(0)} \end{bmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$$

What are the probabilities the computer will be up/down on next day?

$$q_U^{(1)} = q_U^{(0)} P_{UU} + q_D^{(0)} P_{UD}$$
$$q_D^{(1)} = q_U^{(0)} P_{DU} + q_D^{(0)} P_{DD}$$

, or $\underset{\sim}{q}^{(1)} = \begin{pmatrix} P_{UU} & P_{UD} \\ P_{DU} & P_{DD} \end{pmatrix} \begin{pmatrix} q_U^{(0)} \\ q_D^{(0)} \end{pmatrix} = \underline{\underline{P}}\, \underset{\sim}{q}^{(0)}$

Similarly, on the next day $\underset{\sim}{q}^{(2)} = \underline{\underline{P}}\, \underset{\sim}{q}^{(1)} = \underline{\underline{P}}^2\, \underset{\sim}{q}^{(0)}$

and, in general, on the t^{th} day: $\underset{\sim}{q}^{(t)} = \underline{\underline{P}}\, \underset{\sim}{q}^{(t-1)} = \underline{\underline{P}}^t\, \underset{\sim}{q}^{(0)}$ (10.6).

($\underline{\underline{P}}^t$: exponent ↙ t)

Note: $\underset{\sim}{q}$ is a vector of a priori probabilities, while $\underline{\underline{P}}$ is a matrix of conditional probabilities.

How does $\underset{\sim}{q}$ evolve?

	t=0	t=1	t=2	t=3	t=4	t=5	t=6	...	t=20	...
q_U	0.50	0.45	0.435	0.4305	0.4292	0.4287	0.4286	...	0.4286	...
q_D	0.50	0.55	0.565	0.5695	0.5709	0.5713	0.5714		0.5714	

Observation: The Markov chain approaches a limiting probability distributi

$$\underset{\sim}{\Pi} = \lim_{t\to\infty} \underset{\sim}{q}^{(t)} \qquad (10.7)$$

That limiting probability distribution satisfies the equation:

$$\underline{\underline{P}}\ \underset{\sim}{\Pi} = \underset{\sim}{\Pi} \qquad \text{(steady state condition)} \qquad (10.8)$$

This means that $\underset{\sim}{\Pi}$ is an eigenvector of the stochastic matrix $\underline{\underline{P}}$, corresponding to an eigenvalue of 1.

Notice that $\underset{\sim}{\Pi}$ is completely determined by $\underline{\underline{P}}$, and is not at all influenced by where we started (i.e., by $\underset{\sim}{q}^{(0)}$). All memory of the initial state has been effaced.

Markov chains in which one can ultimately go from any state to any other state are called ergodic, or irreducible. Such chains always have a limiting probability distribution (Perron - Frobenius theorem).

If transition probability matrix is block diagonal, chain is not ergodic.
If transition probability matrix is full, chain is ergodic.

10.5. The Metropolis Monte Carlo algorithm: A numerical method for calculating averages with respect to a multivariate probability density function.

The Mathematical Problem.

Given: A probability distribution in a multidimensional space. Distribution defined by probabilities Π_m for being at each point m of the multidimensional space.

[**Note:** multidimensional space may be *discrete* (Π_m = a probability), or *continuous*. ($\Pi_m = \rho(\underset{\sim}{X}_m)\, d^{3N}\underset{\sim}{X}$, where $\rho(\underset{\sim}{X}_m)$ is a probability density, and $d^{3N}\underset{\sim}{X}$ an elementary volume in the multidimensional space centered at $\underset{\sim}{X}_m$].

Sought: An efficient numerical procedure for *sampling* the multidimensional space, according to the probability distribution $\{\Pi_m\}$.

By "sampling" we mean picking a finite set of points (states)

$$m_1, m_2, \ldots, m_t, \ldots, m_{N_{trials}}$$

where a given state may appear more than once in the sequence, such that the probability of finding each state m_t in the sequence is practically equal to Π_{m_t}

If we manage to accomplish this sampling (generate such a sequence) then we can immediately calculate the average of any function defined on the space of interest as:

$$\langle F \rangle \simeq \frac{1}{N_{trials}} \sum_{t=1}^{N_{trials}} F(m_t) \qquad (10.9)$$

true average — average over finite sample of generated states.

Solution Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H.; Teller, E. J. Chem. Phys. 1953, 21, 1087-92, Los Alamos MANIAC computer. MR^2T^2

Generate the representative set of points as a sequence, each from the previous one, according to certain stochastic rules: a Markov chain.

$$m_1 \to m_2 \to \cdots \to m_t \to \cdots \to m_{N_{trials}}$$

Select the transition matrix of the Markov chain as:

$$P_{nm} = \begin{cases} C_{nm}, & \text{if } \Pi_n \geq \Pi_m, \quad n \neq m \\ C_{nm} \dfrac{\Pi_n}{\Pi_m}, & \text{if } \Pi_n < \Pi_m, \quad n \neq m \end{cases} \qquad (10.10)$$

$$P_{nn} = 1 - \sum_{n \neq m} P_{nm}$$

where $\underline{\underline{C}}$ is a symmetric stochastic matrix (underlying matrix of Markov chain)

$$C_{nm} = C_{mn} \qquad \sum_n C_{nm} = 1 \qquad (10.11)$$

C_{nm} = probability of attempting a move from state m to state n.

Usually, moves are attempted only between close-lying states in the multidimensional space. That is, most of the elements of $\underline{\underline{C}}$ are zero,

except a few corresponding to pairs of states which lie very close to each other. Usually, C_{nm} is uniform in a small region of the space around state m. (Attempt a move from m to any of the nearby states n with equal probability).

$$\min\left(1, \frac{\Pi_n}{\Pi_m}\right) = \text{probability of accepting a move.} \qquad (10.12)$$

(to: n, from: m)

"METROPOLIS SELECTION CRITERION"

- If new state, n, is more probable than the state m from which we started, then "accept" new state n as the next state in the chain

- If new state, n, is less probable than the state m from which we started, then "accept" new state n with probability $\Pi_n / \Pi_m < 1$.

If new state, n, is not accepted, then state m is retained as the next state in the chain, and its characteristics are entered again in the averaging procedure.

Properties of the transition matrix $\underline{\underline{P}}$

- $\underline{\underline{P}}$ is stochastic: $\sum_n P_{nm} = 1$ (10.13)
  (By construction, see eq. 10.10)

- $\underline{\underline{P}}$ satisfies the condition of "microscopic reversibility".

$$\Pi_m P_{nm} = \Pi_n P_{mn} \qquad (10.14)$$

(P_{nm}: to n, from m)

Probability of occurrence of move $m \to n$ | Probability of occurrence of move $n \to m$

PROOF: Assume that, for the two states in question, $\Pi_n \gg \Pi_m$. Then, from eq (10.10

$$\Pi_m P_{nm} = \Pi_m C_{nm}$$

$$\Pi_n P_{mn} = \Pi_n \left(C_{mn} \frac{\Pi_m}{\Pi_n} \right) = C_{mn} \Pi_m = \Pi_m C_{mn}$$

($\uparrow$ [$\Pi_m < \Pi_n$])

$\Rightarrow \Pi_m P_{nm} = \Pi_n P_{mn}$

From the symmetry of $\underline{\underline{C}}$, eq (10.11), $C_{nm} = C_{mn}$

Proof for $\Pi_n < \Pi_m$ proceeds identically (role of m,n reversed).

- $\underset{\sim}{\Pi}$ is the limiting distribution corresponding to the transition matrix $\underline{\underline{P}}$

PROOF: For $\underset{\sim}{\Pi}$ to be the limiting distribution corresponding to $\underline{\underline{P}}$, one must have $\underline{\underline{P}}\,\underset{\sim}{\Pi} = \underset{\sim}{\Pi}$.

The i^{th} element of $\underline{\underline{P}}\,\underset{\sim}{\Pi}$ is

$$\sum_m P_{nm} \Pi_m \overset{\text{microscopic reversibility, (10.14)}}{=} \sum_m P_{mn} \Pi_n = \Pi_n \sum_m P_{mn} \overset{\text{stochastic nature of } \underline{\underline{P}}, \text{ (10.13}}{=} \Pi_n \, 1 = \Pi_n$$

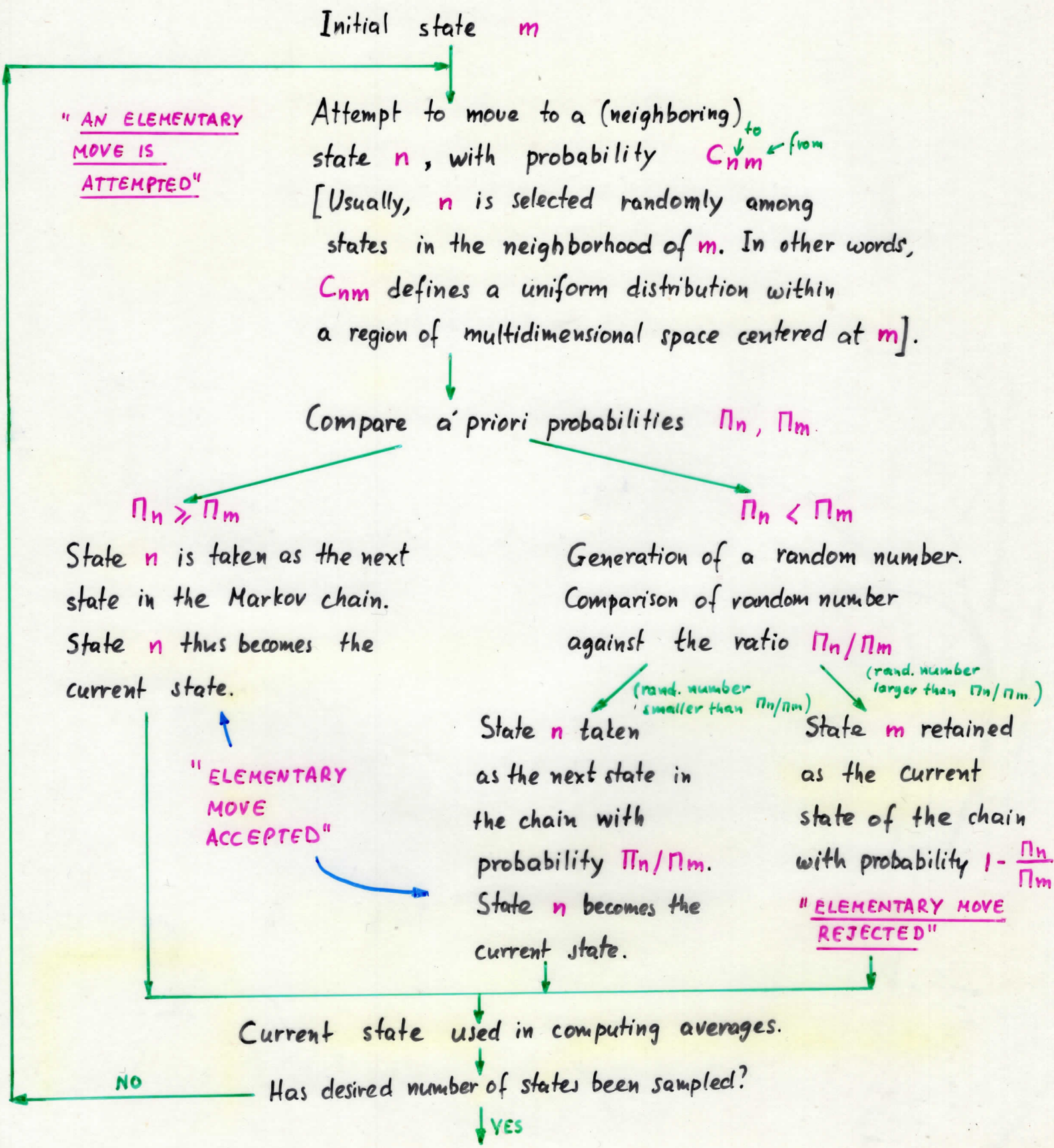
Therefore, combining the equations for all elements, $\underline{\underline{P}}\,\underset{\sim}{\Pi} = \underset{\sim}{\Pi}$. (10.15)

Eq. (10.15) guarantees that, whatever the starting state in the generated Markov chain, a chain long enough will asymptotically sample the probability distribution of interest, defined by $\underset{\sim}{\Pi}$.

Note: In the definition of the transition matrix $\underline{\underline{P}}$ by equations (10.15), we only need to know $\underset{\sim}{\Pi}$ up to a multiplicative constant. This is because only probability ratios appear in the Metropolis scheme.

This aspect of MR^2T^2 algorithm is particularly convenient in Statistical Mechanics applications.

Flow of calculations in Metropolis Monte Carlo



10.3.6. NVT MC of a simple liquid by the Metropolis algorithm.

(Assume liquid argon, represented by pairwise LJ interactions)

OBJECTIVE: Generate a sequence of configurations of N molecules in volume V that asymptotically samples the probability density of the canonical (NVT) ensemble.

Transcription

Multidimensional space → Configuration space $(\underset{\sim}{r}_1, \ldots, \underset{\sim}{r}_N)$

States (points) → Configurations.

$\Pi_m \rightarrow \rho^{NVT}(\underset{\sim}{r}_1^{(m)}, \ldots, \underset{\sim}{r}_N^{(m)}) d^{3N}r$ ($d^{3N}r$ = elementary volume element in config. space)

$\Pi_n / \Pi_m \rightarrow \rho^{NVT}(\underset{\sim}{r}_1^{(n)}, \ldots, \underset{\sim}{r}_N^{(n)}) / \rho^{NVT}(\underset{\sim}{r}_1^{(m)}, \ldots, \underset{\sim}{r}_N^{(m)}) \equiv \rho_n^{NVT} / \rho_m^{NVT}$

Note:

$$\rho^{NVT}(\underset{\sim}{r}_1^{(m)}, \ldots, \underset{\sim}{r}_N^{(m)}) \equiv \rho_m^{NVT} = \frac{\exp[-\beta \mathcal{V}(\underset{\sim}{r}_1^{(m)}, \ldots, \underset{\sim}{r}_N^{(m)})]}{\mathcal{Z}(N,V,T)} \equiv \frac{\exp(-\beta \mathcal{V}_m)}{\mathcal{Z}} \quad (10.16)$$

and

$$\frac{\rho_n^{NVT}}{\rho_m^{NVT}} = \frac{\exp[-\beta \mathcal{V}_n]}{\exp[-\beta \mathcal{V}_m]} = \exp[-\beta(\mathcal{V}_n - \mathcal{V}_m)] = \exp[-\beta \Delta\mathcal{V}_{m \to n}] \quad (10.17)$$

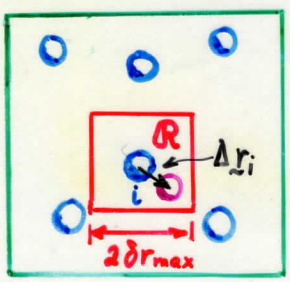
Therefore, acceptance probability

$$\min\left(1, \frac{\rho_n^{NVT}}{\rho_m^{NVT}}\right) = \begin{cases} 1, & \text{if } \mathcal{V}_n \leq \mathcal{V}_m \quad \text{(downhill move accepted always)} \\ \exp[-\beta \Delta\mathcal{V}_{m \to n}], & \text{if } \mathcal{V}_n > \mathcal{V}_m \quad \text{(uphill move accepted sometimes)} \end{cases} \quad (10.18)$$

(ρ_n^{NVT}: new; ρ_m^{NVT}: old)

METROPOLIS SELECTION CRITERION: Involves only potential energy.

Elementary Move: Translation of a randomly chosen molecule by a random amount. Components of random displacement vector uniformly distributed within the interval $(-\delta r_{max}, \delta r_{max})$. Distribution of attempted new positions is uniform within a cube of edgelength $2\delta r_{max}$.

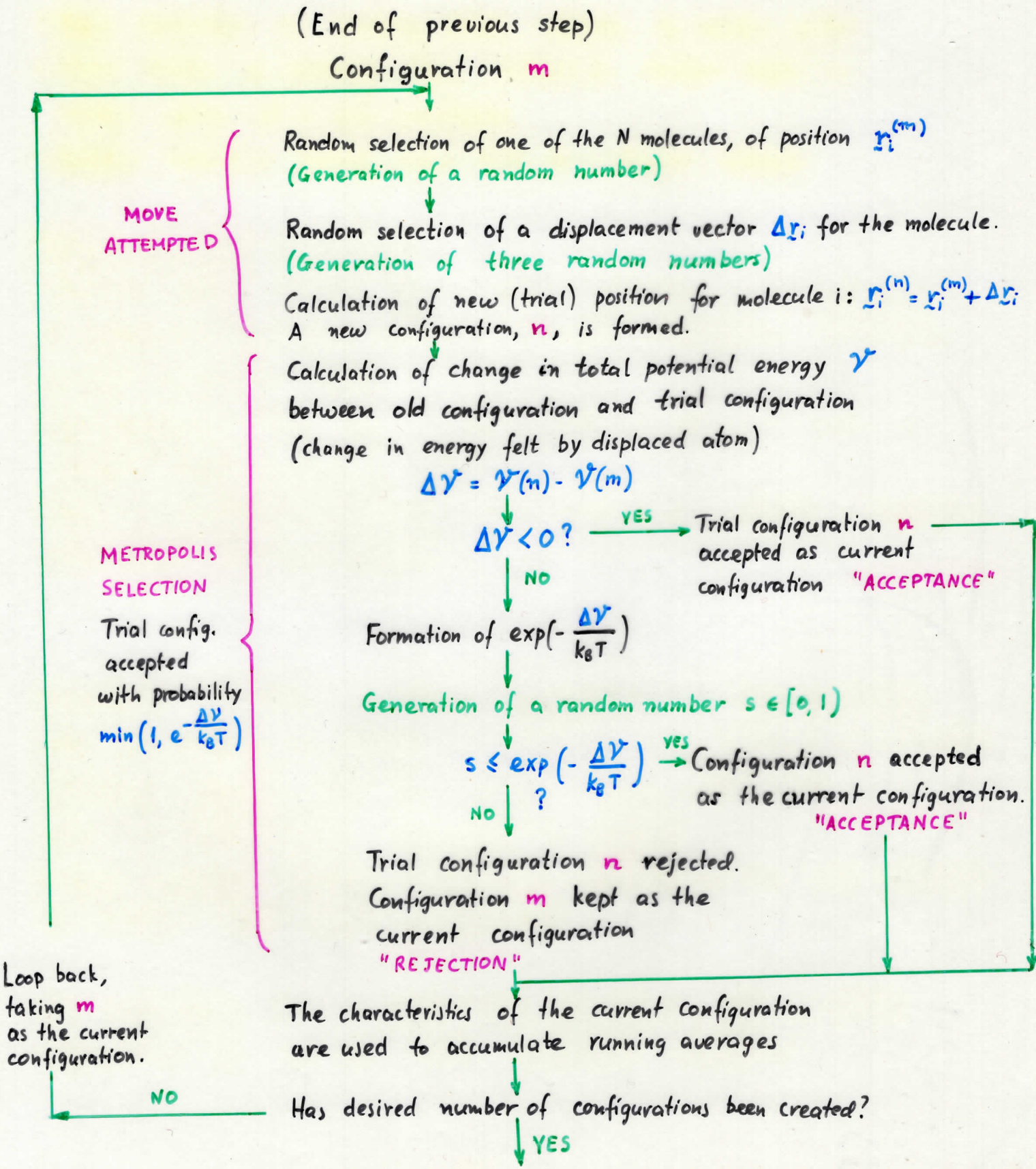


$$C_{nm} = \begin{cases} \frac{1}{N} \cdot \frac{1}{N_R}, & \Delta\underset{\sim}{r}_i \in R \\ 0, & \Delta\underset{\sim}{r}_i \notin R \end{cases} \quad (10.19)$$

(number of molecules; Number of possible new positions in cube R)

BY CONSTRUCTION, C IS SYMMETRIC.

Flow of Calculations in NVT Monte Carlo of Liquid Argon.



In molecular systems, elementary moves should change all configurational degrees of freedom:

- Rigid translation of randomly chosen molecule by random vector
- Rigid rotation of randomly chosen molecule by random angle, or random choice of a new orientation.
- Rotation around a randomly chosen bond for a flexible molecule.

Boldness of moves (e.g., δr_{max})

Should be chosen so as to give maximal efficiency in the exploration of configuration space (e.g., maximal mean squared displacement of molecules for given number of moves.)

Rule of thumb: Choose δr_{max} so that 50% of attempted moves are accepted (50% acceptance ratio).

There are indications that bolder moves (with lower probability of acceptance than 50%) are more efficient.

Starting Configuration

Advantageous to choose a starting configuration of high probability under the conditions of interest, to minimize time spent in equilibration.

Traditional approach: Start from an fcc lattice, and "melt" it to obtain a liquid. Ensures non-overlap of molecules.

Another approach: Random insertion of molecules in box, and relaxation of overlaps with energy minimization.

Last configuration from a previous MC run can be used as starting configuration for new run under comparable conditions.

Equilibration

Initially, the probability distribution of the sampled microstates in configuration space is $\underset{\sim}{q}^{(0)} = (0, 0, \ldots, 1, \ldots, 0)$ [system in initial configuration].

As simulation goes on, the distribution of sampled microstates approaches the stationary distribution of the Metropolis transition matrix, given by the simulated ensemble: $\lim_{t \to \infty} \underset{\sim}{q}^{(t)} = \underset{\sim}{\pi}$

(Generated Markov chain of microstates asymptotically samples ensemble of interest).

Initial period of the simulation is an equilibration period, and must not be taken into account in the calculation of averages.

At the end of the equilibration period, all memory of the initial configuration should have been lost.

Monitoring equilibration:

- Monitor potential energy and pressure: Equilibration period should be extended until they show no systematic drift, but oscillate about steady mean values.
- In the case of a lattice starting configuration, ensure that all vestiges of order have disappeared. (translational and orientational order parameters should show no order in fluid).
- For a fluid simulation, make sure that mean squared displacement of molecules grows roughly linearly with time (diffusive behavior).

Rule of thumb: For low-molecular weight systems, 500N - 1000N steps are sufficient for equilibration. More time is required if one starts from a lattice, or if a phase transition is close.

10.3.7. Isothermal - Isobaric (NPT) Monte Carlo (Wood, 1968)

Convenient: Corresponds to the way thermodynamic properties are measured experimentally.
Not subject to the problem of two phases appearing in the same simulation cell.

Probability density: (fluctuates)

$$\rho(\underset{\sim}{r}^N; V) = \frac{\exp[-\beta(\mathcal{V}(\underset{\sim}{r}^N) + PV)]}{\mathcal{Z}_{NPT}}$$

A problem in using $(\underset{\sim}{r}^N, V)$ as configurational degrees of freedom is that the limits between which the components of $\underset{\sim}{r}^N$ take on values depend on V. To avoid this interdependence among degrees of freedom, introduce scaled coordinates $\underset{\sim}{s} = (\underset{\sim}{s}_1, \underset{\sim}{s}_2, \ldots, \underset{\sim}{s}_N)$, where $\underset{\sim}{s} = \frac{1}{L}\underset{\sim}{r}$. $0 \le s_{\alpha i} < 1$ (10.20)

$$\rho(\underset{\sim}{s}^N; V)\, d^{3N}s\, dV = \rho(\underset{\sim}{r}^N, V)\, d^{3N}r\, dV = \rho(\underset{\sim}{r}^N, V)\, L^{3N}\, d^{3N}s\, dV \text{, or}$$

$$\rho(\underset{\sim}{s}^N; V) = \frac{\exp[-\beta(\mathcal{V}(\underset{\sim}{s}^N; V) + PV)]}{\mathcal{Z}_{NPT}} V^N = \frac{\exp[-\beta(\mathcal{V}(\underset{\sim}{s}^N; V) + PV) + N\ln V]}{\mathcal{Z}_{NPT}} \quad (10.21)$$

Ratio of state probabilities:

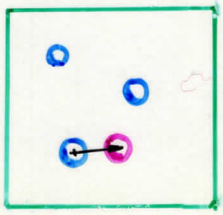
$$\frac{\rho_n}{\rho_m} = \frac{\exp[-\beta(\mathcal{V}_n + PV_n) + N\ln V_n]}{\exp[-\beta(\mathcal{V}_m + PV_m) + N\ln V_m]} = \exp\left[-\beta\left(\mathcal{V}_n - \mathcal{V}_m + P(V_n - V_m) - \frac{N}{\beta}\ln\frac{V_n}{V_m}\right)\right]$$

$$= \exp\left[-\beta\left(\Delta\mathcal{V} + P\Delta V - \frac{N}{\beta}\ln\frac{V_n}{V_m}\right)\right] \quad (10.22)$$

used in place of $\Delta\mathcal{V}$ in Metropolis selection criterion

Note: "Tail" contribution to $\mathcal{V}$ changes upon volume fluctuation, and therefore must be computed at each step.

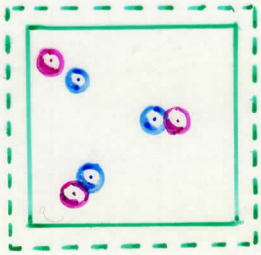
Moves



Translation of randomly selected molecule.

ACCEPTANCE PROBABILITY

$$\min\left(1, \exp[-\beta \Delta \mathcal{V}]\right)$$



$$\min\left(1, \exp\left[-\beta\left(\Delta \mathcal{V} + P\Delta V - \frac{N}{\beta} \ln \frac{V_{new}}{V_{old}}\right)\right]\right)$$

Dilation/contraction of box, under preservation of all scaled coordinates (affine transformation of atom centers)

Calculation of energy change upon volume fluctuation:

$$\mathcal{V}_m = 4\epsilon \sum_i \sum_{j>i} \left(\frac{\sigma}{L_m s_{ij}^{(m)}}\right)^{12} - 4\epsilon \sum_i \sum_{j>i} \left(\frac{\sigma}{L_m s_{ij}^{(m)}}\right)^{6} = \mathcal{V}_m^{(12)} + \mathcal{V}_m^{(6)} \qquad (10.23)$$

↑ stored separately. ↑

$$\mathcal{V}_n = \mathcal{V}_m^{(12)} \left(\frac{L_m}{L_n}\right)^{12} + \mathcal{V}_m^{(6)} \left(\frac{L_m}{L_n}\right)^{6} \qquad \text{[scaled coordinates not changed]} \qquad (10.24)$$

(The above presupposes that there is only one characteristic length in the potential function. For molecular systems, whole $\mathcal{V}$ must be recalculated).

Relative frequency of translation (rotation, conformational change) and volume (edgelength) fluctuation moves must be set so that equilibration is fastest for a given amount of CPU time. Jorgensen: 1 volume fluctuation every $6N$ moves.

10.8. Grand Canonical Monte Carlo (μVT) (Norman and Filinov, 1969)

$$p(\underset{\sim}{r}^N; N) = \frac{1}{N!} \frac{\exp(N\beta\mu)}{\Lambda^{3N}} \cdot \exp[-\beta V(r_1, \dots, r_N)] \cdot \frac{1}{\Xi}$$

Again, $\underset{\sim}{r}^N$ and N are interdependent. Introduce a set of scaled coordinates, $(s_1, \dots, s_N)$, where $s_{\alpha i} = V^{-1/3} r_{\alpha i}$, $0 \le s_{\alpha i} \le 1$

$$p(\underset{\sim}{r}^N; N)\, d^{3N}r = p(\underset{\sim}{s}^N; N)\, d^{3N}s \quad \Rightarrow \quad p(\underset{\sim}{s}^N; N) = V^N p(\underset{\sim}{r}^N; N)$$

$$\text{or} \quad p(\underset{\sim}{s}^N; N) = \frac{\exp[-\beta(V(\underset{\sim}{s}) - N\mu) - \ln N! - 3N\ln\Lambda + N\ln V]}{\Xi} \qquad (10.25)$$

Whereas the dimensions of $p(\underset{\sim}{r}^N; N)$ depend on N, $p(\underset{\sim}{s}^N; N)$ is dimensionless. Taking the ratio of two $p(\underset{\sim}{s}^N; N)$:

$$\frac{p_n}{p_m} = \frac{\exp[-\beta(V_n - N_n\mu) - \ln N_n! - 3N_n\ln\Lambda + N_n\ln V]}{\exp[-\beta(V_m - N_m\mu) - \ln N_m! - 3N_m\ln\Lambda + N_m\ln V]} =$$

$$= \exp\left[-\beta(V_n - V_m) + \beta\mu(N_n - N_m) - \ln\frac{N_n!}{N_m!} - 3(N_n - N_m)\ln\Lambda + (N_n - N_m)\ln V\right] \qquad (10.26)$$

In particular, if $N_n - N_m = 1$ (creation of a molecule),

$$\frac{p_n}{p_m} = \exp[-\beta(V_n - V_m) + \beta\mu - \ln(N_m + 1) - 3\ln\Lambda + \ln V] =$$

$$= \exp\left[-\beta(V_n - V_m) + \ln\left(\frac{z V}{N_m + 1}\right)\right] \qquad \text{where } z = \frac{\exp(\beta\mu)}{\Lambda^3}, \qquad (10.27)$$

a constant of the simulation.

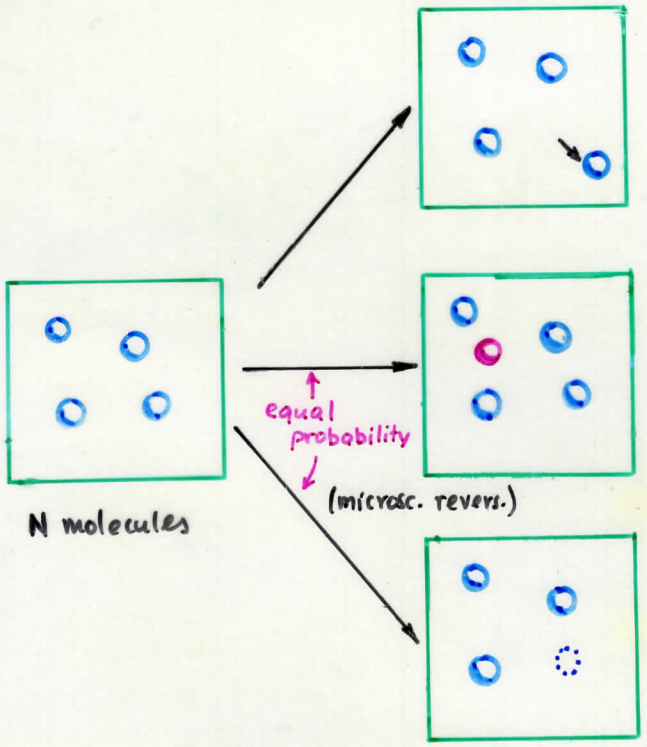
If $N_n - N_m = -1$ (destruction of a molecule)

$$\frac{p_n}{p_m} = \exp\left[-\beta(V_n - V_m) - \beta\mu + \ln N_m + 3\ln\Lambda - \ln V\right] =$$

$$= \exp\left[-\beta(V_n - V_m) + \ln\left(\frac{N_m}{z V}\right)\right] \qquad \text{where } z = \frac{\exp(\beta\mu)}{\Lambda^3} \quad (10.28)$$

Moves

ACCEPTANCE PROBABILITY



Molecule translation: $\min\left(1, \exp(-\beta\Delta V)\right)$

Molecule Creation: $\min\left(1, \exp\left(-\beta\Delta V + \ln\left(\frac{zV}{N+1}\right)\right)\right)$

Molecule Destruction: $\min\left(1, \exp\left(-\beta\Delta V + \ln\left(\frac{N}{zV}\right)\right)\right)$

Adams modification (1974):

Separate μ into ideal and excess parts:

$$\mu = \mu^{ex} + \mu^{id} = \mu^{ex} + \frac{1}{\beta} \ln\left(\frac{\langle N\rangle_{\mu VT}}{V}\Lambda^3\right) = \underbrace{\mu^{ex} + k_B T \ln\langle N\rangle_{\mu VT}}_{k_B T B} + k_B T \ln\left(\frac{\Lambda^3}{V}\right) \qquad (10.29)$$

Simulation is performed at constant B, V, T.
μ can be calculated by calculating $\langle N\rangle$ during the run, and then using the definition of B.

Note: $z = \frac{\exp(\beta\mu)}{\Lambda^3} = \frac{\exp(B).\,\Lambda^3/V}{\Lambda^3} = \frac{\exp(B)}{V}$, or $zV = \exp(B)$. (10.30)

Hence the ratio of probabilities appearing in the selection criteria becomes

$$\exp\left[-\beta\Delta V + \ln\left(\frac{e^B}{N+1}\right)\right] \qquad \text{molecule creation} \qquad (10.31)$$

$$\exp\left[-\beta\Delta V + \ln\left(\frac{N}{e^B}\right)\right] \qquad \text{molecule destruction.} \qquad (10.32)$$

An advantage of GCMC method is that it permits the direct calculation of free energy quantities, as μ is set.
Method is very appropriate for the study of sorption equilibria (surfaces, pores).