# *Intermolecular Forces and Monte-Carlo Integration*

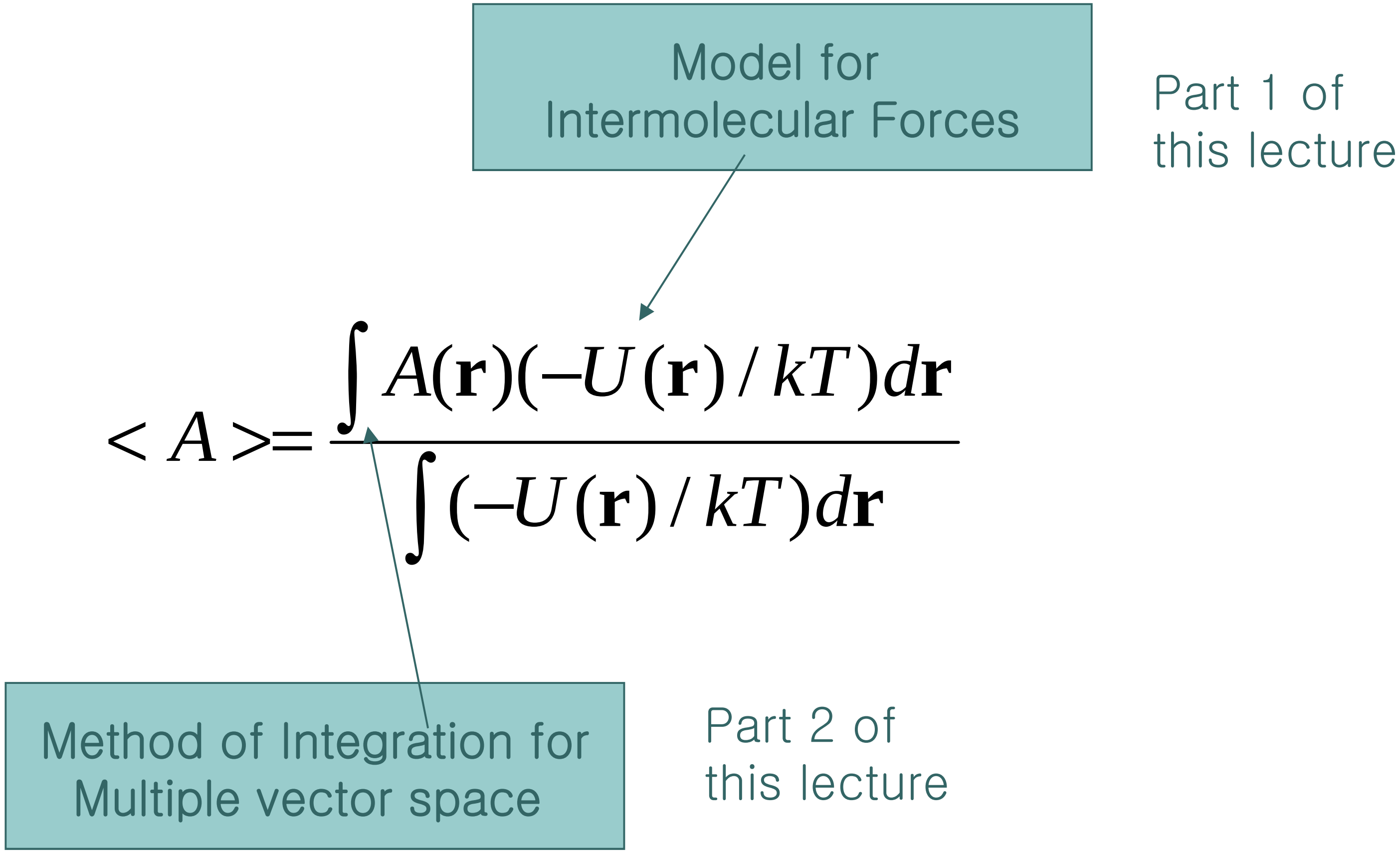
열역학 특수 연구

**2003.3.28**

# *Source of the lecture note.*

- **J.M.Prausnitz and others, "Molecular Thermodynamics of Fluid Phase Equiliria"**
- **Atkins, "Physical Chemistry"**
- **Lecture Note, Prof. D.A.Kofke, University at Buffalo**
- **Lecture Note, R.J.Sadus, Swineburn University**

# Tasks of Molecular Simulation



Model for
Intermolecular Forces

Part 1 of
this lecture

$$< A >= \frac{\int A(\mathbf{r})(-U(\mathbf{r})/kT)d\mathbf{r}}{\int(-U(\mathbf{r})/kT)d\mathbf{r}}$$

Method of Integration for
Multiple vector space

Part 2 of
this lecture

# *Intermolecular Forces*

- Intermolecular forces
  - Force acting between the molecules of given mixture or pure species
- It is essential to understand the nature of intermolecular forces for the study of molecular simulation
- Only simple and idealized models are available (approximation)
- Our understanding of intermolecular forces are far from complete.

# *Types of intermolecular forces*

- **Electrostatic forces**
  - **Charged particles and permanent dipoles**
- **Induced forces**
  - **Permanent dipole and induced dipole**
- **Force of attraction between nonpolar molecules**
- **Specific forces**
  - **Hydrogen bonding, association and complex formation**

# *Potential Energy Function and Intermolecular Forces*

- **Potential Energy : Energy due to relative position to one another**

$$F = -\frac{d\Gamma}{dr}$$

- **If additional variables are required for potential energy function …**

$$F(r,\theta,\phi,\ldots) = -\nabla\Gamma(r,\theta,\phi,\ldots)$$

# *1. Electrostatic Force*

- **Due to permanent charges (ions,…)**
- **Coulomb's relation (inverse square law)**
  - **Two point charges separated from distance r**

$$F = \frac{e_i e_j}{r^2} \qquad F = -\frac{d\Gamma}{dr} \qquad \Gamma_{ij} = \frac{e_i e_j}{r} + const.$$

- **For two charged molecules (ions) ,**

$$\Gamma_{ij} = \frac{z_i z_j \varepsilon^2}{\underset{\downarrow}{D}r}$$

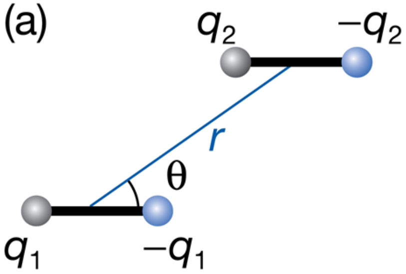
Dielectric constant of given medium

# *Nature of Electrostatic forces*

- **Dominant contribution of energy ….**
- **Long range nature**
  - **Force is inversely proportional to square of the distance**
  - **Major difficulties for concentrated electrolyte solutions**

# *Electrostatic forces between dipoles*

- **Dipole**
  - **Particles do not have net electric charge**
  - **Particles have two electric charges of same magnitude *e* but opposite sign.**
  - **Dipole moment** $\mu = el$
- **Potential Energy between two dipoles**



$$\Gamma_{ij} = \frac{\mu_i \mu_j}{r^3}\left[2\cos\theta_i \cos\theta_j - \sin\theta_i \sin\theta_j \cos(\phi_i - \phi_j)\right]$$

# *Energies of permanent dipole, quadrupoles*

- **Orientations of molecules are governed by two competing factors**
  - **Electric field by the presence of polor molecules**
  - **Kinetic energy → random orientation**

- **Dipole-Dipole** $$\Gamma_{ij} = -\frac{2}{3}\frac{\mu_i^2 \mu_j^2}{r^6 kT} + \ldots$$

- **Dipole-Quadrupole** $$\Gamma_{ij} = -\frac{\mu_i^2 Q_j^2}{r^8 kT} + \ldots$$

- **Quadrupole-Quadrupole** $$\Gamma_{ij} = -\frac{Q_i^2 Q_j^2}{40 r^{10} kT} + \ldots$$

# *2. Induced Forces*

- **Nonpolar molecules can be induced when those molecules are subjected to an electric field.**

$$\mu^{i} = \alpha E$$

Electric Field Strength

Polarizability

# *Mean Potential Energies of induced dipoles*

- **Permanent Dipole + Induced Dipole**

$$\Gamma_{ij} = -\frac{\alpha_i \mu_j^{\ 2}}{r^6}$$

- **Permanent Dipole + Permanent Dipole**

$$\Gamma_{ij} = -\frac{\left(\alpha_i \mu_j^2 + \alpha_j \mu_i^2\right)}{r^6}$$

- **Permanent Quadrupole + Permanent Quadrupole**

$$\Gamma_{ij} = -\frac{3\left(\alpha_i Q_j^2 + \alpha_j Q_i^2\right)}{2r^8}$$

# *3. Intermolecular Forces between Nonpolar Molecules*

- **1930, London**
  - **There was no adequate explanation for the forces between nonpolar molecules**
  - **Instant oscillation of electrons → Distortion of electron arrangement was sufficient to caus temporary dipole moment**
  - **On the average, the magnitude and direction averages zero, but quickly varying dipoles produce an electric field. → induces dipoles in the surrounding molecules**
  - **Induced dipole-induced dipole interaction**

# London dispersion force

$$\Gamma_{ij} = -\frac{3}{2}\frac{\alpha_i\alpha_j\left(I_iI_j\right)}{r^6(I_i+I_j)}$$

Potential energy between two nonpolar molecules are :

<u>independent of temperature and</u>

<u>varies inversely as sixth power of the distance  between them .</u>

$$\Gamma_{ii} = -\frac{B}{r^6} \qquad \Gamma_{ij} = \sqrt{\Gamma_{ii}\Gamma_{jj}}$$

# *Repulsive force and total interaction*

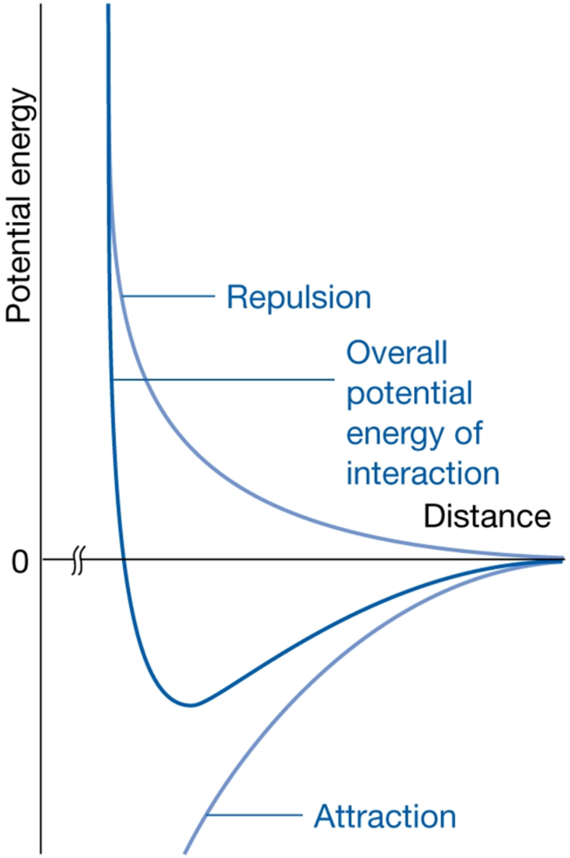
- When molecules are squeezed, electronic replusion and rising of eletronic kinetic energy began to dominate the attractive force
- The repulsive potential can be modeled by inverse-power law

$$\Gamma = \frac{A}{r^m}$$

- The total potential is the sum of two separate potential

$$\Gamma_{ij} = \frac{A}{r^m} - \frac{B}{r^n}$$

# *General form of intermolecular potential curve*



- **Mie's Potential**

$$\Gamma = \frac{A}{r^m} - \frac{B}{r^n} = \frac{\varepsilon(n^n/m^m)^{1/(n-m)}}{n-m}\left[\left(\frac{\sigma}{r}\right)^n - \left(\frac{\sigma}{r}\right)^m\right]$$

- **Lennard-Jones Potential**

$$\Gamma = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6\right]$$

→ The parameters for potential models can be estimated from variety of physical properties (spectroscopic and molecular-beam experiments)

# *Specific (Chemical) Forces*

- **Association : The tendency to from polymer**
- **Solvation : The tendency to form complexes from different species**

- **Hydrogen Bond and Electron Donor-Acceptor complexes**

- **The models for specific forces are not well established.**

- **The most important contribution in bio-molecules (proteins, DNA, RNA,…)**

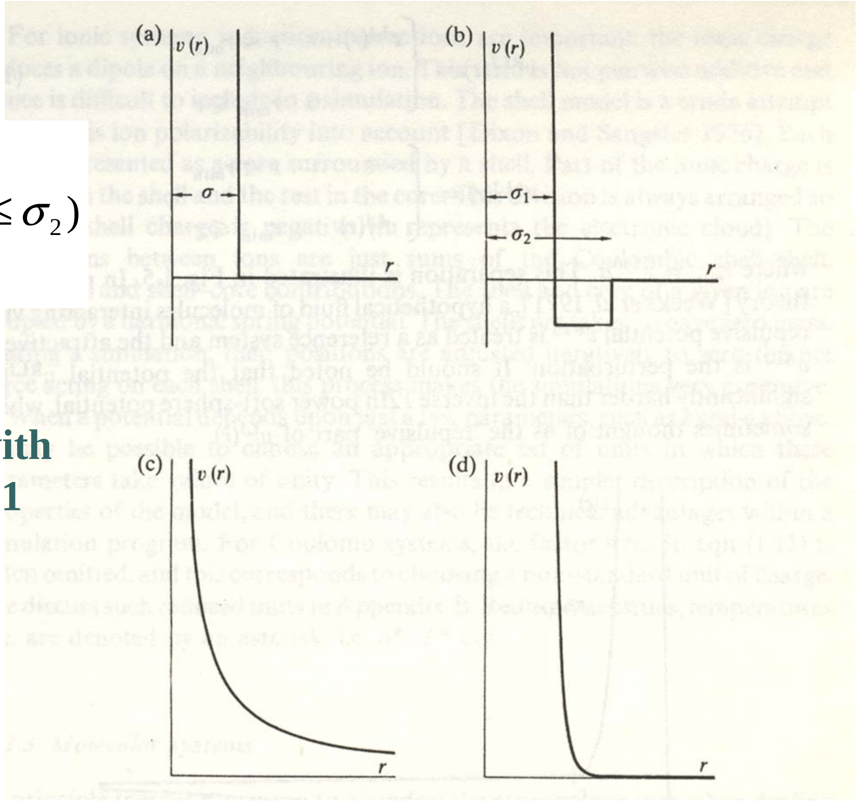
# Simplified Potential Models for Molecular Simulations

**Hard Sphere Potential**

$$\Gamma_{SW}(r) = \infty \qquad (r < \sigma)$$
$$\Gamma_{SW}(r) = -\varepsilon \qquad (\sigma_1 \le r \le \sigma_2)$$
$$\Gamma_{SW}(r) = 0 \qquad (r \ge \sigma_2)$$

**Square Well Potential**

$$\Gamma_{HS}(r) = \infty \qquad (r < \sigma)$$
$$\Gamma_{HS}(r) = 0 \qquad (r \ge \sigma)$$



**Soft-Sphere Potential with Repulsion parameter = 1**

$$\Gamma_{SS}(r) = \frac{a}{r^{v}}$$

**Soft-Sphere Potential with Repulsion parameter = 12**

$$\Gamma_{SS}(r) = \frac{a}{r^{v}}$$

# *Calculation of Potential in Molecular Monte Carlo Simulation*

- **There are no contribution of kinetic energy in MMC simulation**
  - **Only “configurational” terms are calculated**

$$U = \sum \Gamma_1(r_i) + \sum \sum \Gamma_2(r_i, r_j) + \sum \sum \sum \Gamma_3(r_i, r_j, r_k) + \ldots$$

Potential between particles of triplets

Potential between pairs of particles

Effect of external field

# *Using reduced units…*

- **Dimensionless units are used for computer simulation purposes**

$$\rho^* = \rho\sigma^3$$

$$T^* = kT / \varepsilon$$

$$E^* = E / \varepsilon$$

$$P^* = P\sigma^3 / \varepsilon$$

# *Contribution to Potential energy*

- **Two-body interactions are most important term in the calculation**
- **For some cases, three body interactions may be important.**
- **Including three body interactions imposes a very large increase in computation time.**

$$t \propto N^m$$

**m: number of interactions**

# *Short range and long range forces*

- ❒ Short range force
  - Dispersion and Replusion
- ❒ Long range force
  - Ion-Ion and Dipole-Dipole interaction

| Interaction Type | Dependence | Typical E (kJ/mol) | Comment |
|---|---|---|---|
| Ion-Ion | 1/r | 250 | |
| Ion-Dipole | $1/r^2$ | 15 | |
| Dipole-Dipole | $1/r^3$ | 2 | Stationary |
| Dipole-Dipole | $1/r^6$ | 0.6 | Rotating |
| London | $1/r^6$ | 2 | |

# *Short range and long range interactions*

- **Computation time-saving devices for short range interactions**
    - **Periodic boundary condition**
    - **Neighbor list**
- **Special methods are required for long range interactions. (The interaction extends past the length of the simulation box)**

# Naïve energy calculation

$$U = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \Gamma_2(r_{ij})$$

↓

Summation are chosen to avoid “self” interaction

**Pseudo Code**

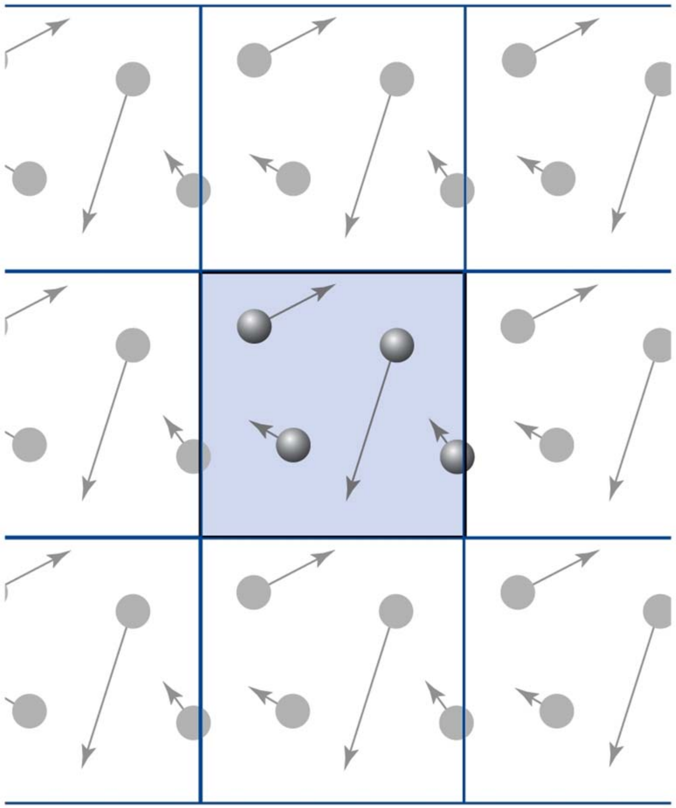
```
Loop i = 1, N−1
          Loop j = i+1,N
                    Evaluate $r_{ij}$
                    Evaulate $U_{ij}$
                    Accumulate Energy
          End j Loop
End j Loop
```

# *Problems*

- **Simulations are performed typically with a few hundred molecules arranged in a cubic lattice.**
  - **Large fraction of molecules can be expected at the surface rather than in the bulk.**

- **Periodic Boundary Conditions (PBC) are used to avoid this problem**

# *Periodic Boundary Condition*

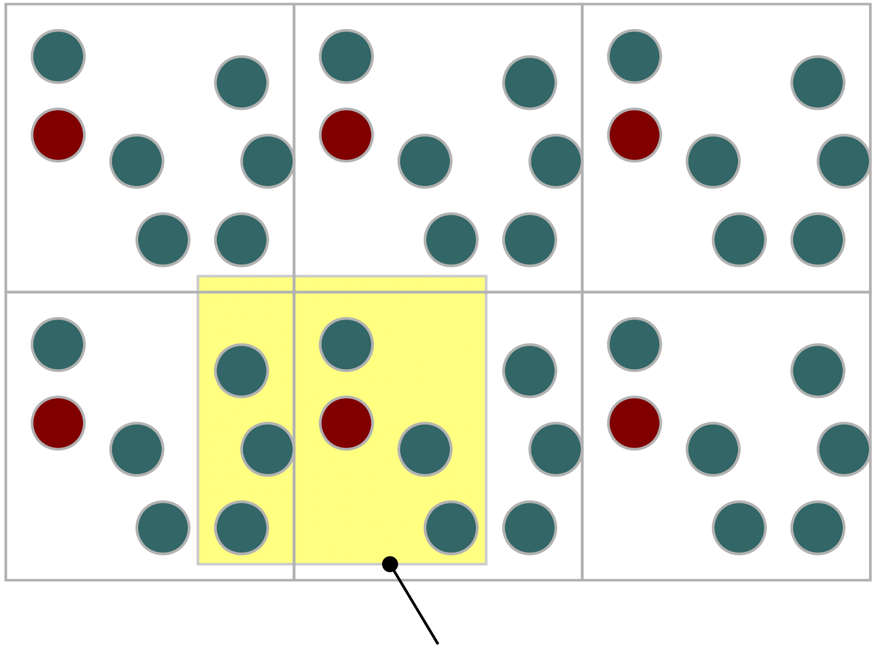


- **Infinite Replica of the lattice of the cubic simulation box**
- **Molecules on any lattice have a mirror image counter part in all the other boxes**
- **Changes in one box are matched exactly in the other boxes → surface effects are eliminated.**

# *Another difficulty…*

- **Summation over infinite array of periodic images**
  - **→ This problem can be overcame using Minimum Image Convention (MIC)**

# *Minimum Image Convention (MIC)*



Nearest images of colored sphere

***For a given molecule, we position the molecule at the center of a box with dimension identical to the simulation box.***

***Assume that the central molecule only interacts with all molecules whose center fall within this region.***

***All the coordinates lie within the range of ½ L and – ½ L***

# *Implementing PBC & MIC*

## Two Approaches

- **Decision based : if statement**
- **Function based : rounding, truncation, modulus**

**Decision**

```
BOXL2 = BOXL/2.0
IF(RX(I).GT.BOXL2) RX(I)=RX(I)−BOXL
IF(RX(I).LT.−BOXL2) RX(I) = RX(I) + BOXL
```

**Function**

```
RX(I) = RX(I) − BOXL * AINT(RX(I)/BOXL)
```

Nearest integer

# *Implementing PBC & MIC*

Pseudo CODE

```
Loop i = 1, N −1
        Loop j = I + 1, N
                Evaluate $r_{ij}$
                Convert $r_{ij}$ to its periodic image ($r_{ij}$’)
                if ($r_{ij}$’ < cutOffDistance)
                        Evaluate U($r_{ij}$)
                        Accumulate Energy
                End if
        End j Loop
End i Loop
```

# *Improvement due to PBC & MIC (Compared with naïve calculation)*

- **Accumulated energies are calculated for the periodic separation distance.**
- **Only molecules within cut-off distance contribute to the calculated energy.**

- **Caution : cut-off distance should be smaller than the size of the simulation box → Violation to MIC**
- **Calculated potential → Truncated potential**

# *Long range correction to PCB*

- **Adding long range correction…**

$$X_{full} = X_c + X_{lrc}$$

$$E_{lrc} = 2\pi N\rho \int_{r_c}^{\infty} r^2 u(r) dr$$

For NVT ensemble, density and no. of particles are const.
→LRC and be added after simulation
For other ensembles, LRC terms must be added during simulation

# *Technique to reduce computation time → Neighbor List*

- **1967, Verlet proposed a new algorithm.**
- **Instead of searching for neighboring molecules, the neighbor of the molecules are stored and used for the calculation.**

# *Neighbor List*

```
Part 1       $topOfList \leftarrow 0$            //start with empty list
Part 1.1     loop $i \leftarrow 1 \ldots N - 1$   //select molecule $i$
                 $listEntry_i \leftarrow 0$
Part 2           loop $j \leftarrow i + 1 \ldots N$   //look for neighbours of $i$
Part 2.1             Evaluate $rx_{ij}$, $ry_{ij}$ and $rz_{ij}$.
                     Evaluate periodic images ( $rx_{ij}'$, $ry_{ij}'$ and $rz_{ij}'$)
                     $r^2 \leftarrow rx_{ij}'^2 + ry_{ij}'^2 + rz_{ij}'^2$
Part 2.2             if ($r^2 < (rCut^2 + d)$)  //neighbour found
                         $topOfList \leftarrow topOfList + 1$
                         $listEntry_i \leftarrow topOfList$ //position of $j$ on list
                         $list_{topOfList} \leftarrow j$      //enter $j$ on list
                     end if
                 end $j$ loop
             end $i$ loop
```

# *Neighbor List*

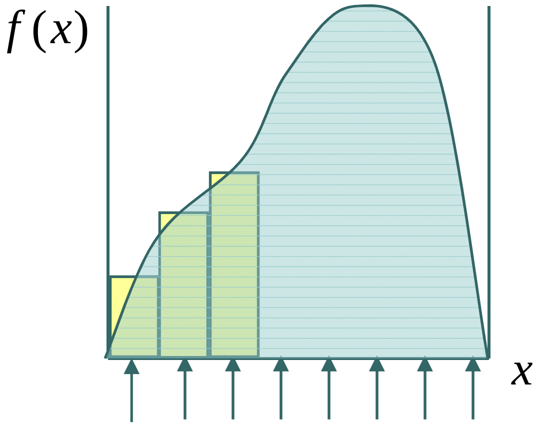
- **Variable d is used to encompass molecules slightly outside the cut-off distance (buffer).**
- **Update of the list**
  - **Update of the list per 10-20 steps**
  - **Largest displacement exceed d value.**

# *Algorithm for Integration*

# *Method of Integration*

- **Methodological Approach**
  - **Rectangular Rule, Triangular Rule, Simpson's Rule**

$$I = \int_a^b f(x)\,dx$$



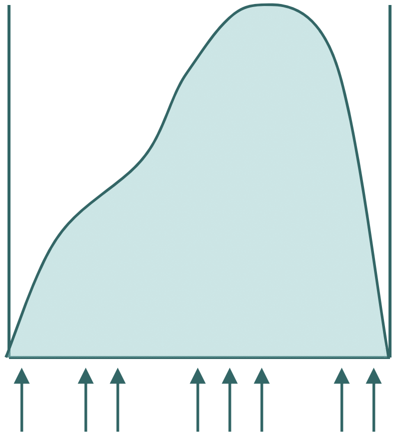
Uniformly separated points

*Quadrature Formula*

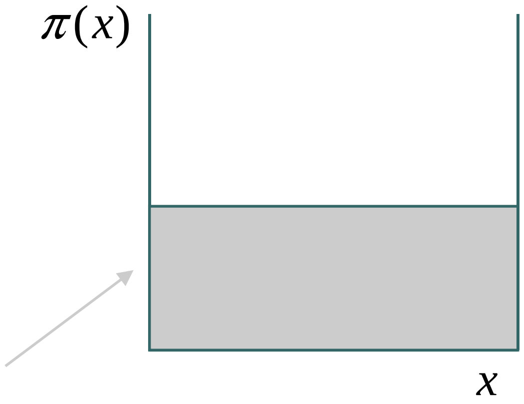
$$I \approx \Delta x \sum_{i=1}^{n} f(x_i) = \frac{b-a}{n} \sum_{i=1}^{n} f(x_i)$$

# *Monte Carlo Integration*

- **Stochastic Approach**
- **Same quadrature formula, different selection of points**



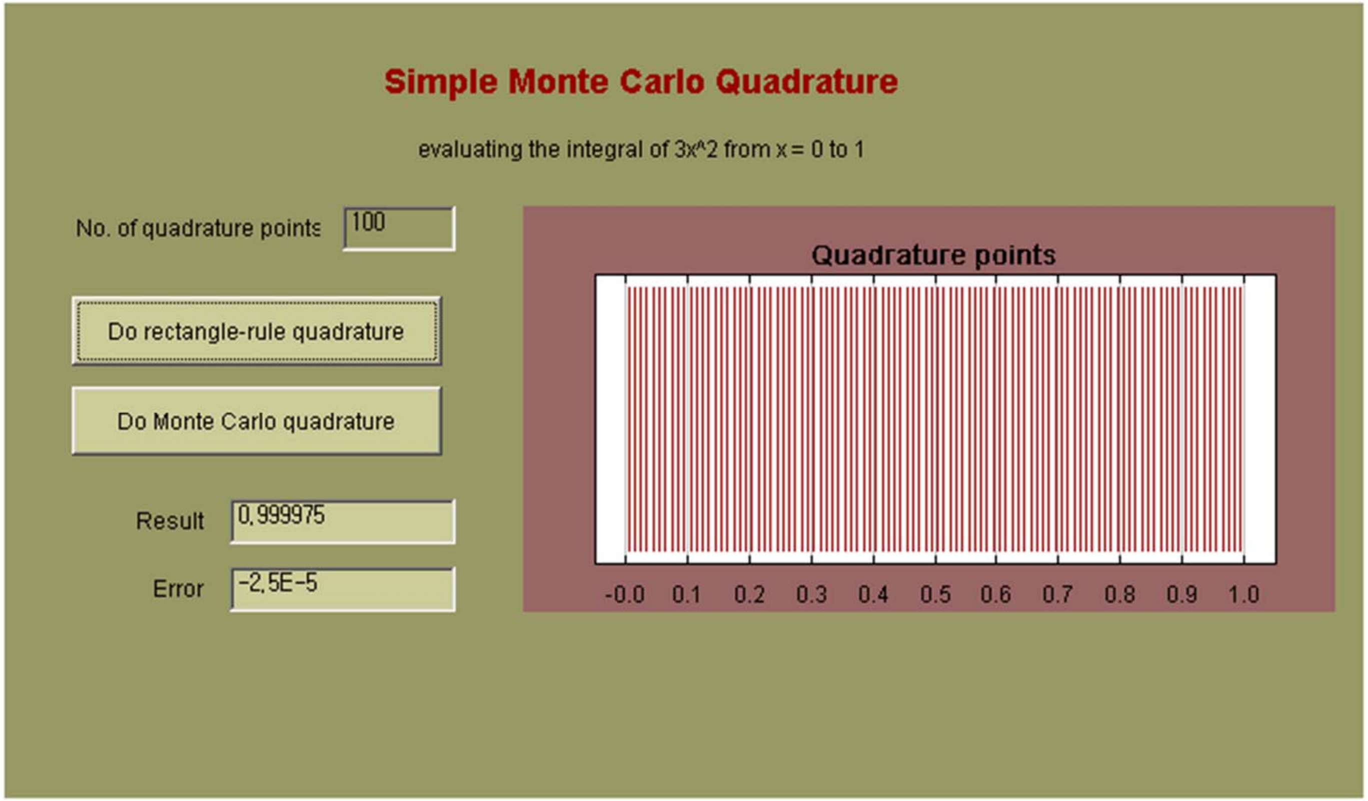
$$I \approx \frac{b-a}{n}\sum_{i=1}^{n} f(x_i)$$



Points are selected from uniform distribution $\pi(x)$

# Example …

*(from Univ. at Buffalo, School of Eng. And Appl. Science,  Prof. David Kofke)*



**Simple Monte Carlo Quadrature**

evaluating the integral of 3x^2 from x = 0 to 1

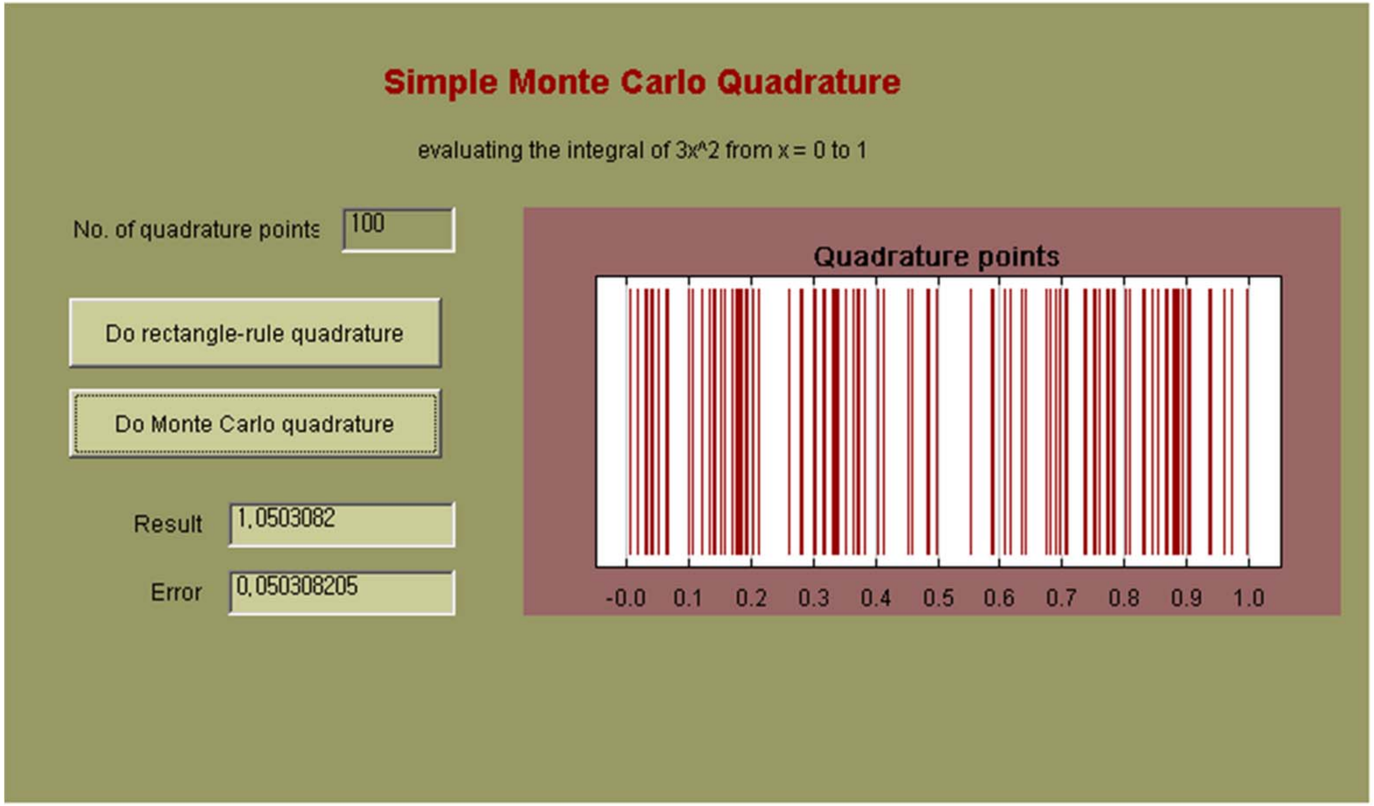
No. of quadrature points: 100

Result: 0,999975

Error: -2,5E-5

# *Example …*

*(from Univ. at Buffalo, School of Eng. And Appl. Science, Prof. David Kofke)*



**Simple Monte Carlo Quadrature**

evaluating the integral of 3x^2 from x = 0 to 1

No. of quadrature points: 100

Result: 1,0503082

Error: 0,050308205

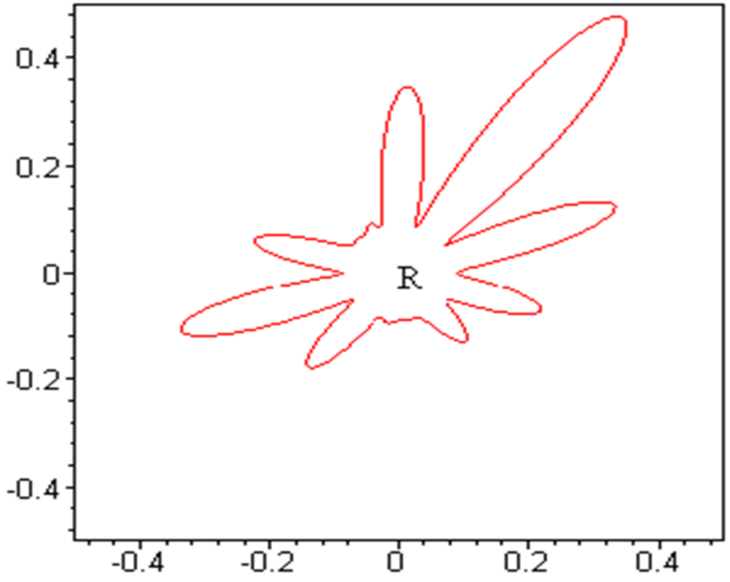
# *Why Monte Carlo Integration ?*

- **Comparison of errors**
  - **Methodological Integration** $E \propto \Delta x^2 / n^2$
  - **Monte Carlo Integration** $E \propto 1/n^{1/2}$
- **MC error vanishes much slowly for increasing n**
- **For one-dimensional integration, MC offers no advantage**
- **The conclusion changes when dimension of integral increases**
  - **Methodological Integration** $E \propto \Delta x^2 / n^{2/d}$
  - **Monte Carlo Integration** $E \propto 1/n^{1/2}$

→ *MC "wins" about d = 4*

# *Shape of High Dimensional Region*

- **Two (and Higher) dimensional shape can be complex**
- **How to construct weighted points in a grid that covers the region R ?**

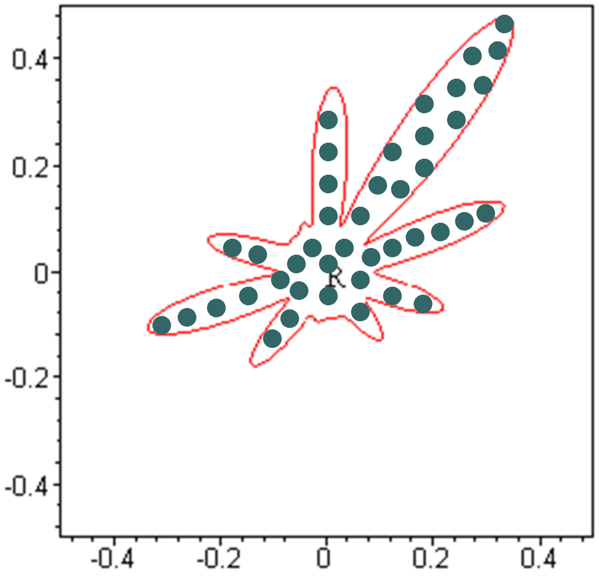


Problem :
mean-square distance from the origin

$$< r^2 >= \frac{\int\int (x^2 + y^2) dxdy}{\int\int dxdy}$$

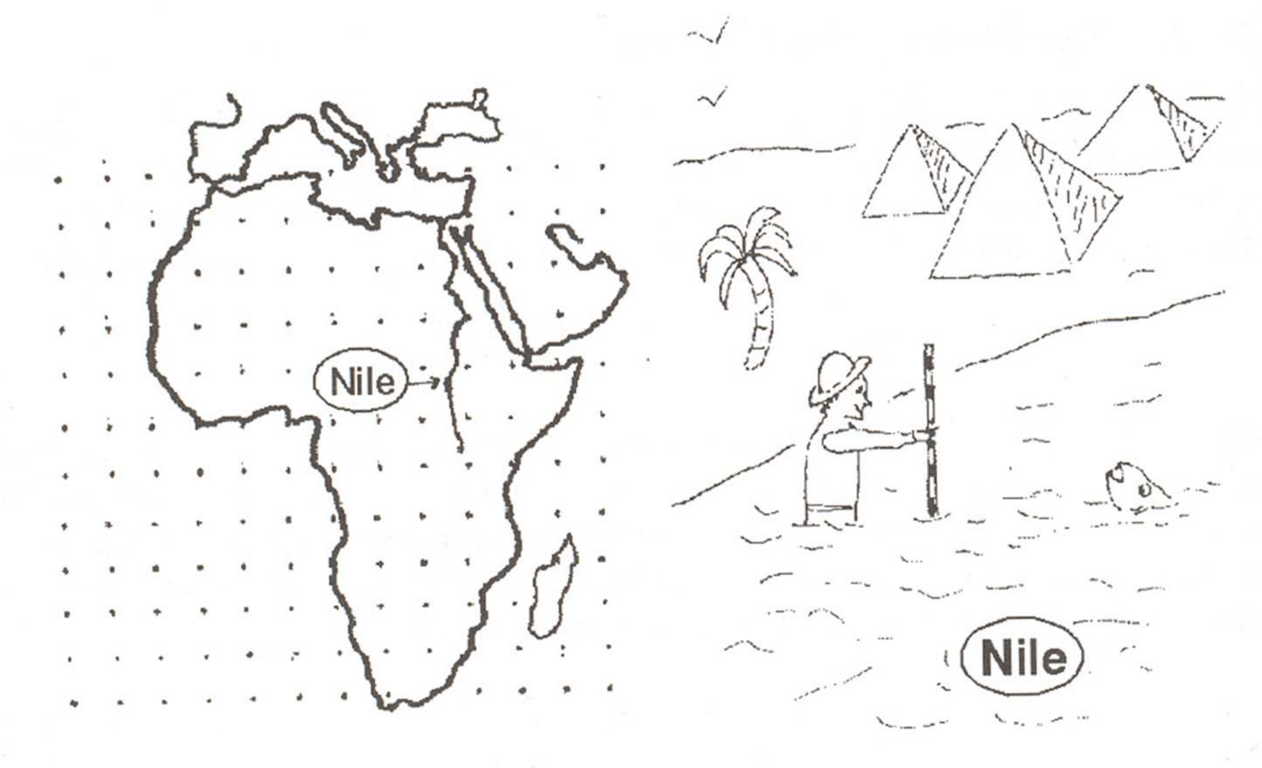
# *Shape of High Dimensional Integral*

- It is hard to formulate methodological algorithm in complex boundary
- Usually we do not have analytical expression for position of boundary
- Complexity of shape can increase unimaginably as dimension of integral grows
- We want 100 + dimensional integrals



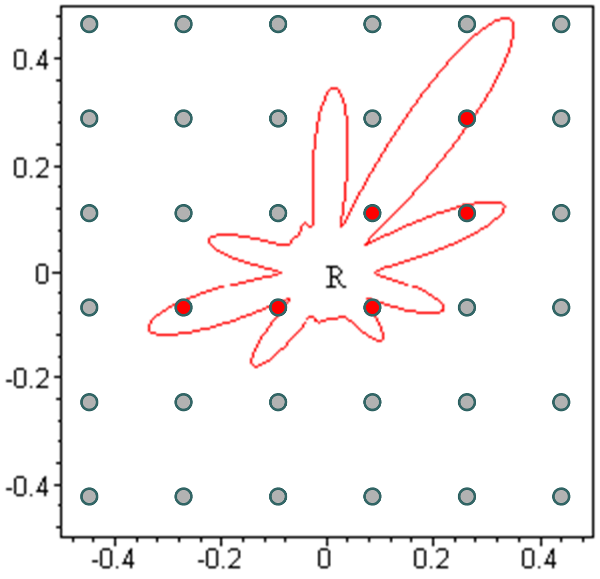
$$< A >= \frac{\int A(\mathbf{r})(-U(\mathbf{r})/kT)d\mathbf{r}}{\int(-U(\mathbf{r})/kT)d\mathbf{r}}$$

# *Nature of the problem …*



# Integration over simple shape ?

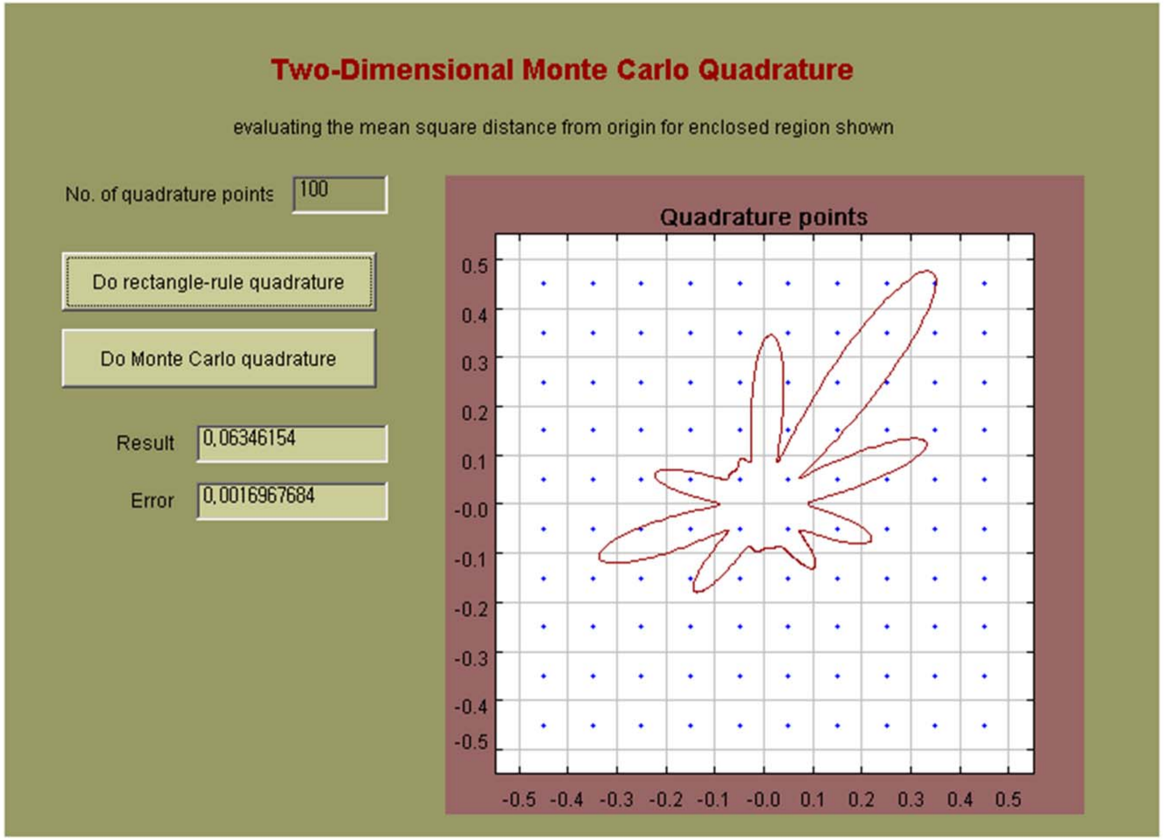
$$s = \begin{cases} 1 & \text{inside R} \quad ● \\ 0 & \text{outside R} \quad ● \end{cases}$$



$$\langle r^2 \rangle = \frac{\int_{-0.5}^{+0.5} dx \int_{-0.5}^{+0.5} dy (x^2 + y^2) s(x,y)}{\int_{-0.5}^{+0.5} dx \int_{-0.5}^{+0.5} dy s(x,y)}$$

→ Grid must be fine enough !

# Sample Integration



**Two-Dimensional Monte Carlo Quadrature**

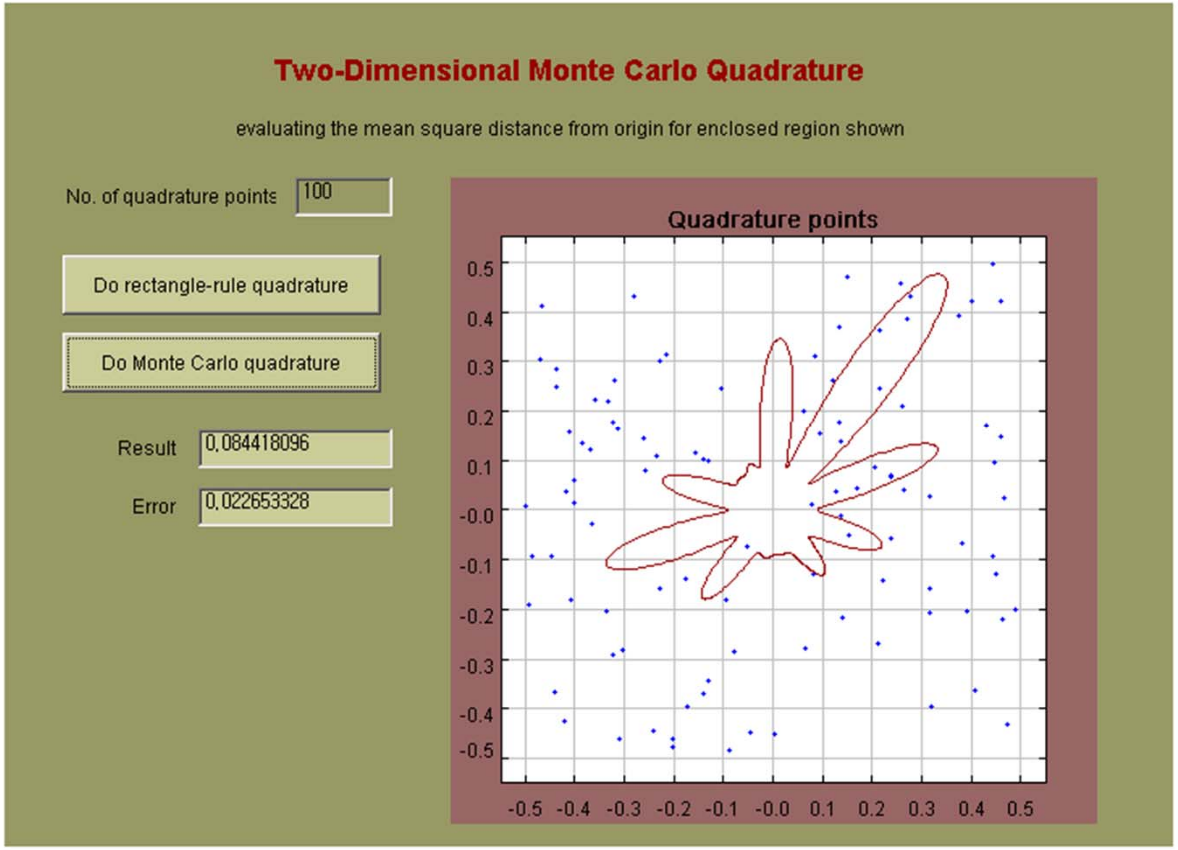
evaluating the mean square distance from origin for enclosed region shown

No. of quadrature points: 100

Result: 0,06346154

Error: 0,0016967684

# *Sample Integration*



**Two-Dimensional Monte Carlo Quadrature**

evaluating the mean square distance from origin for enclosed region shown

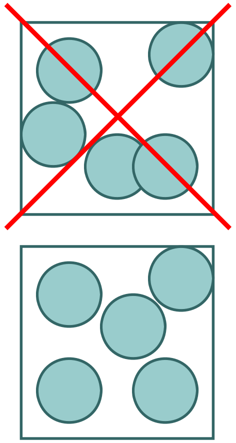
No. of quadrature points: 100

Result: 0,084418096

Error: 0,022653328

# *Integration over simple shape ?*

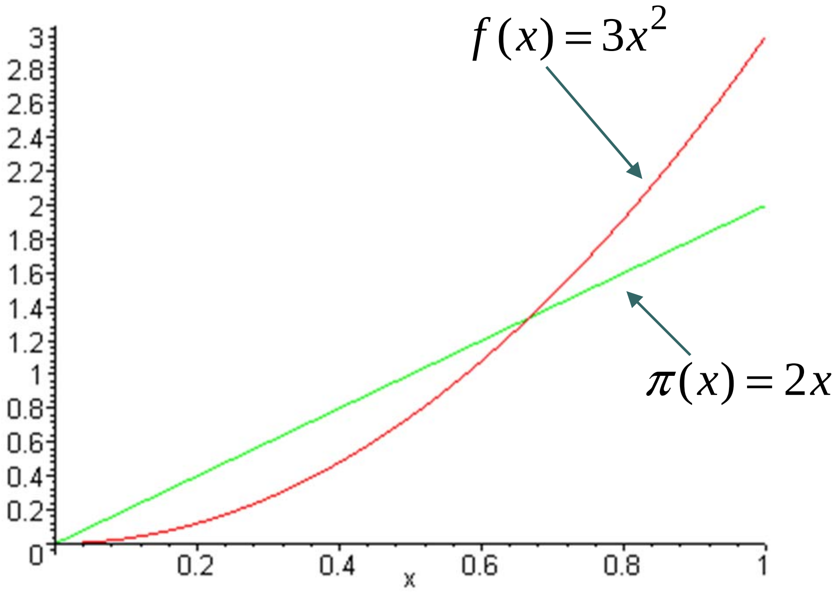
- **Statistical mechanics integrals typically have significant contribution from miniscule regions of the integration space.**
- **Ex ) 100 spheres at freezing fraction = $10^{-260}$**



# *Importance Sampling*

- **Put more quadrature points in the region where integral recieves its greatest contribution**
- **Choose quadrature points according to some distribution function.**

# A sample integration.



$f(x) = 3x^2$

$\pi(x) = 2x$

# *Importance Sampling Integral*

## ❐ Using Rectangular-Rule :

- **Use unevenly spaced intervals**

$$I \approx \sum_{i=1}^{n} f(x_i)\Delta x_i$$

$$\Delta x_i = \frac{b-a}{n}\frac{1}{\pi(x_i)}$$