Molecular Modeling I.

Molecular Perspectives in Chemical Engineering

James F. Ely
Colorado School of Mines
ASEE Summer School
July, 2002

Contributors to this project and/or presentation

Project Contributors:

- Robert Baldwin - CSM
- David Wu – CSM
- Molecular Modeling Task Force

Web Pages and Short Courses of:

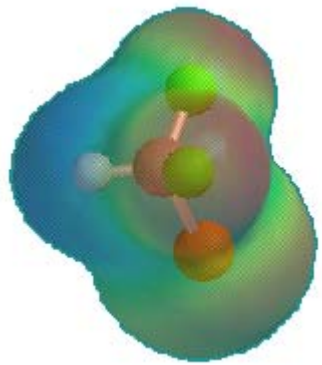
- Dave Kofke – UB/SUNY
- Francis Starr – NIST
- Peter Cummings -- VU

What is Molecular Simulation?

- Molecular simulation is a computational "experiment" conducted on a molecular model
  - *Could be a single molecule (computational quantum chemistry)*
  - *Could involve O(10)-O(10^6) molecules*

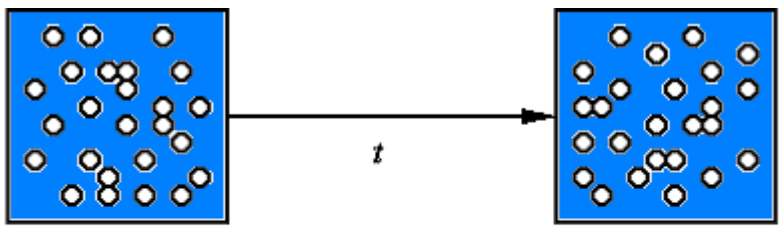
What is Molecular Simulation?

- Computational quantum chemistry generally provides results for isolated or pairs of molecules
  - *Geometry*
  - *Thermochemistry*
  - *Frequencies*
  - *Anything associated with electronic structure*



What is Molecular Simulation?

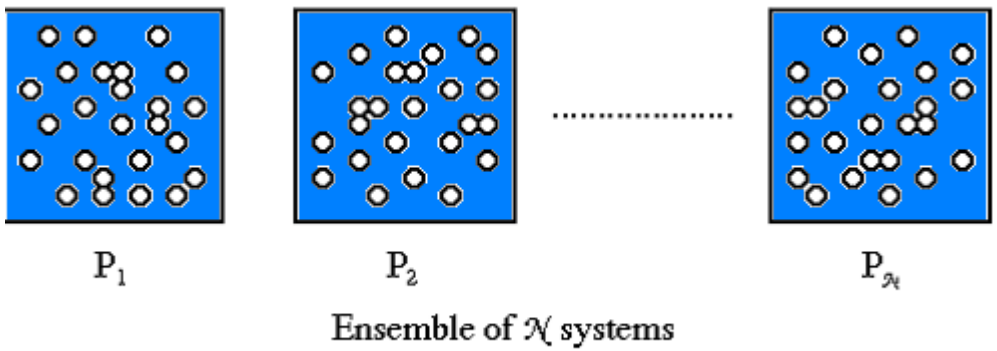
- Molecular dynamics provides results for a system of molecules undergoing dynamic (deterministic) motion
  - *Generates many configurations which are averaged to provide measurements*



Single System Evolving in Time

What is Molecular Simulation?

- Monte Carlo provides results for a system molecules undergoing stochastic motions
  - *Generates an ensemble average with no element of time*

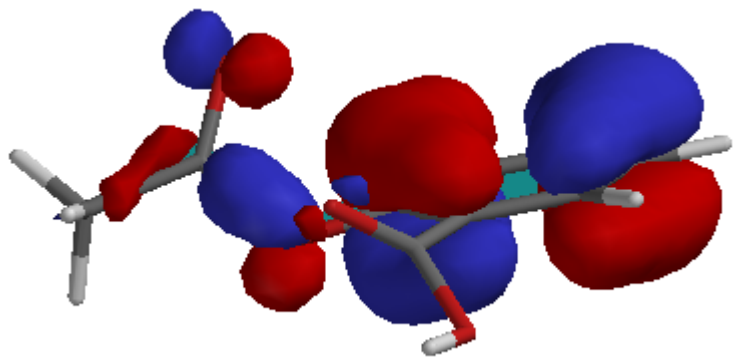


Ensemble of $\mathcal{N}$ systems

What is a Molecular Model?

- Computational Quantum Chemistry
  - *Theory underlying the molecular orbital model that goes into the many-body Schrödinger equation*

$$H\Psi_N = E_N\Psi_N$$



What is a Molecular Model?

- In molecular dynamics and Monte Carlo it is the model for the interactions between molecules, e.g., the force field
  - *Intramolecular Contributions*
  - *Intermolecular Contributions*
- Method for handling finite size of a simulation system

Forces

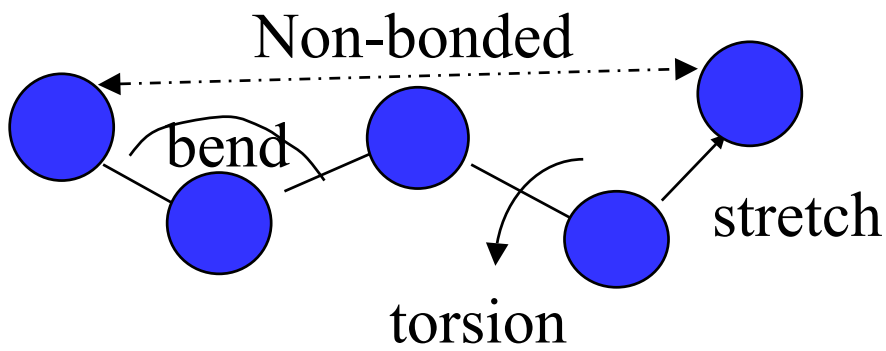
- To begin we start with a categorization of forces (relative strengths)
  - *Strong interactions (10)*
  - *Coulombic force (0.1)*
  - *Weak interactions (10^{-3})*
  - *Gravity (10^{-19})*
- A research area in physics is to unify these forces

Forces

- For our purposes, we are only interested in Coulombic forces
  - *ionic*
  - *covalent*
  - *molecular*
  - *van der Waals*
- All material properties are determined by this force

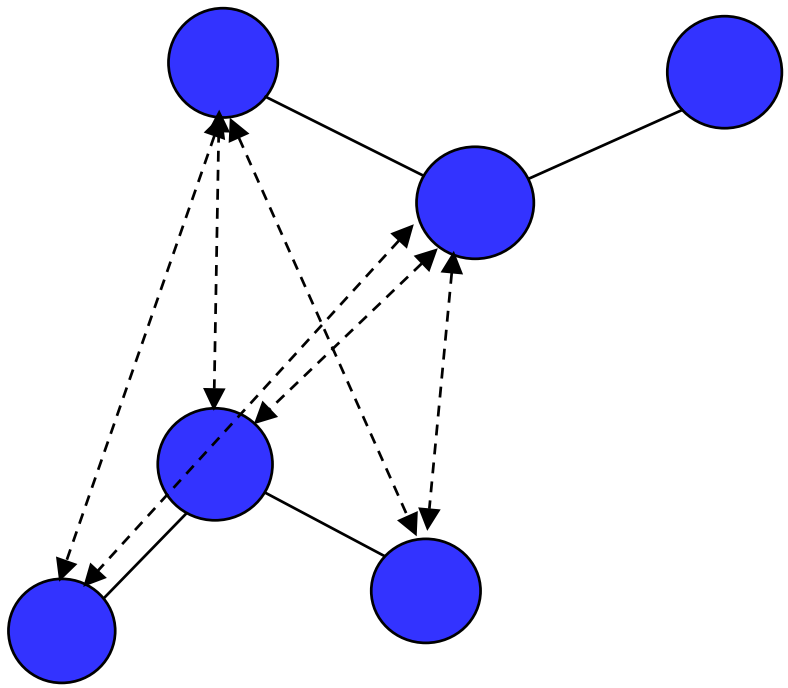
Simulation Force Fields
Intramolecular Forces

$$U^{\text{intra}} = \sum_{\substack{\text{bonds} \\ \text{bond} \\ \text{angles}}} U_i^{stretch} + \sum_{\substack{\text{dihedral} \\ \text{angles}}} U_i^{bend} + \sum_{i} \sum_{j} U_i^{torsion} + \sum_{ij} U_{ij}^{non-bonded}$$



Simulation Force Fields
Intermolecular Forces

$$U^{inter} = U^{repulsion} + U^{lr,attraction} + U^{electrostatic} + U^{multibody}$$



MD and MC Modeling

- It isn't possible to simulate a macroscopic system containing ~10^{23} molecules
  - *No computer exists that could store the positions and momenta of more than 10^9 molecules (~6 terabytes)*
- Solution is paradoxical—we simulate an infinite system!
  - *It is infinite in a special way—it is periodic, i.e. a smaller set of molecules are confined to a regular space filling geometry like a cube which is duplicated in all directions*
  - *applet*

Periodic Boundaries

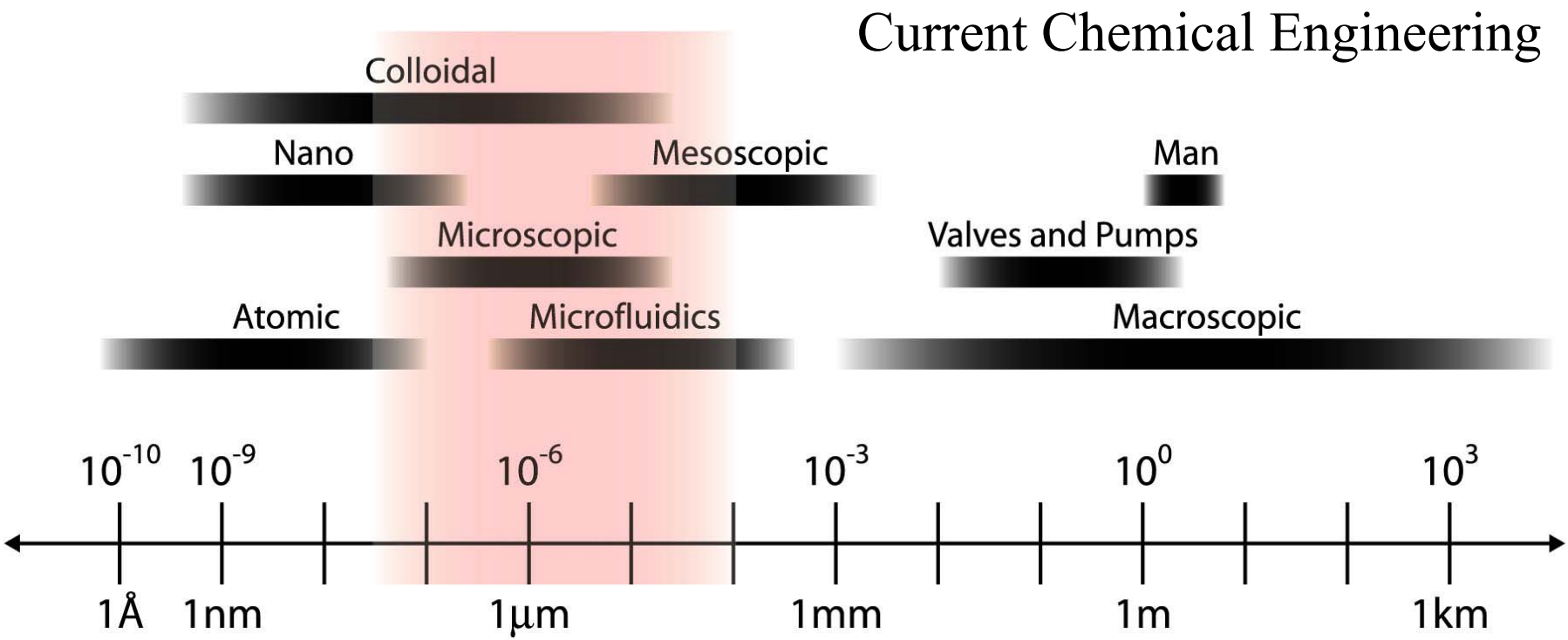
- In the periodic system, each molecule in principle interacts with an infinite number of other molecules.
  - *Not a problem for systems whose forces are short-ranged, e.g., 0.5 nm or so*
  - *For electrostatic systems, where the potential is very long ranged, other techniques must be used*
    - » Ewald sums or reaction fields

Molecular Perspectives in Chemical Engineering

- What is this course about?
  - *Arguably, chemical engineering is at a crossroads as a profession. The traditional chemical processing and petrochemical industries are yielding their dominance to newer, dynamic and fast growing ones like microelectronics and biotechnology.*

Length Scales

Emerging Chemical Engineering



Molecular Simulation as a Tool for Chemical Engineers

- Modern materials science is built upon metallurgy and alchemy
- Today its scope is far beyond its origins
  - *amorphous materials*
  - *ceramics*
  - *polymers*
  - *nanoscale and nanocomposites*
- Modern alchemy is the molecular design of new materials

Molecular Simulation as a Tool for Chemical Engineers

- Since materials are complex, we cannot treat them directly by theory
  - *Historically, experimental methods have been the stalwart of materials research*
- Unfortunately, experimental methods are also inadequate

Progress in Computer Technology is Opening a New Era

- MO, MD and MC calculations are now being used to supplement experiments
  - *New chemicals and materials are being designed by first applying simulations and then being confirmed in experimental tests*
- Computer simulations enhance our understanding of natural phenomena

Industrial Examples

Workshop on Predicting the Thermophysical Properties of Fluids by Molecular Simulation

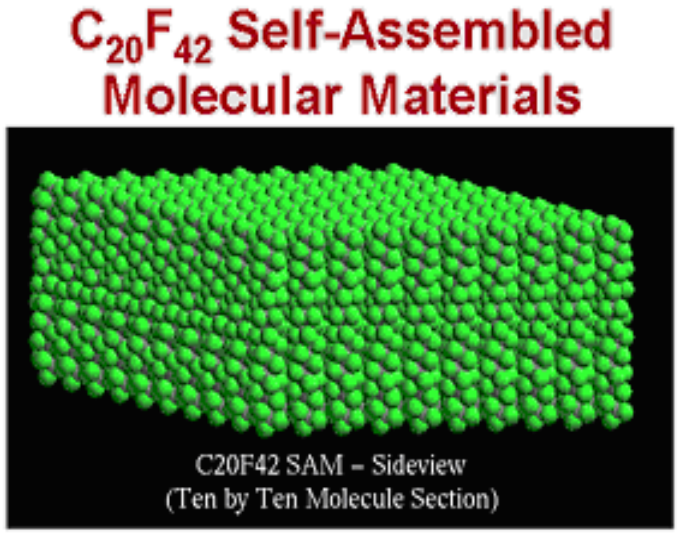
N.I.S.T., Gaithersburg, MD

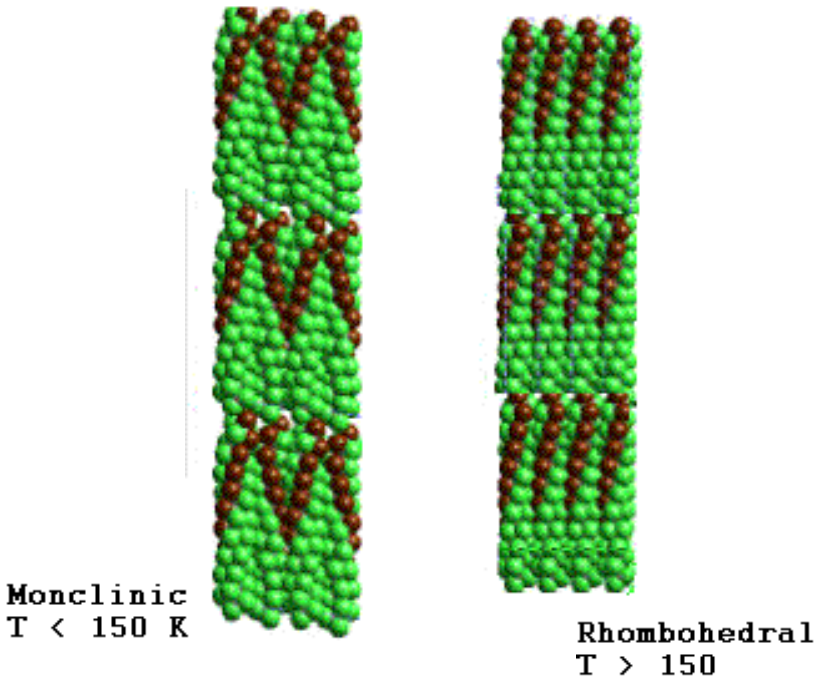
June 18 and 19, 2001

www.ctcms.nist.gov/~fstarr/ptpfms/home.html

Industrial Examples (C. Thomas – 3M)

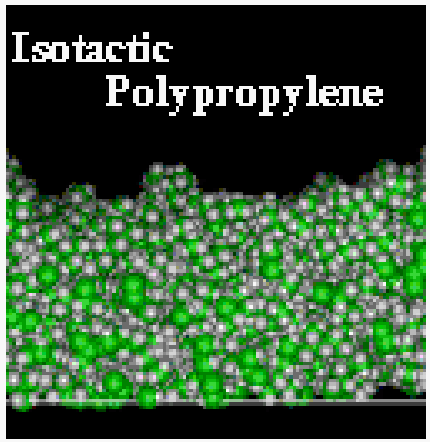
- Molecular Modeling
  - *3M Fluoro-organic molecules*





Industrial Examples (C. Thomas – 3M)

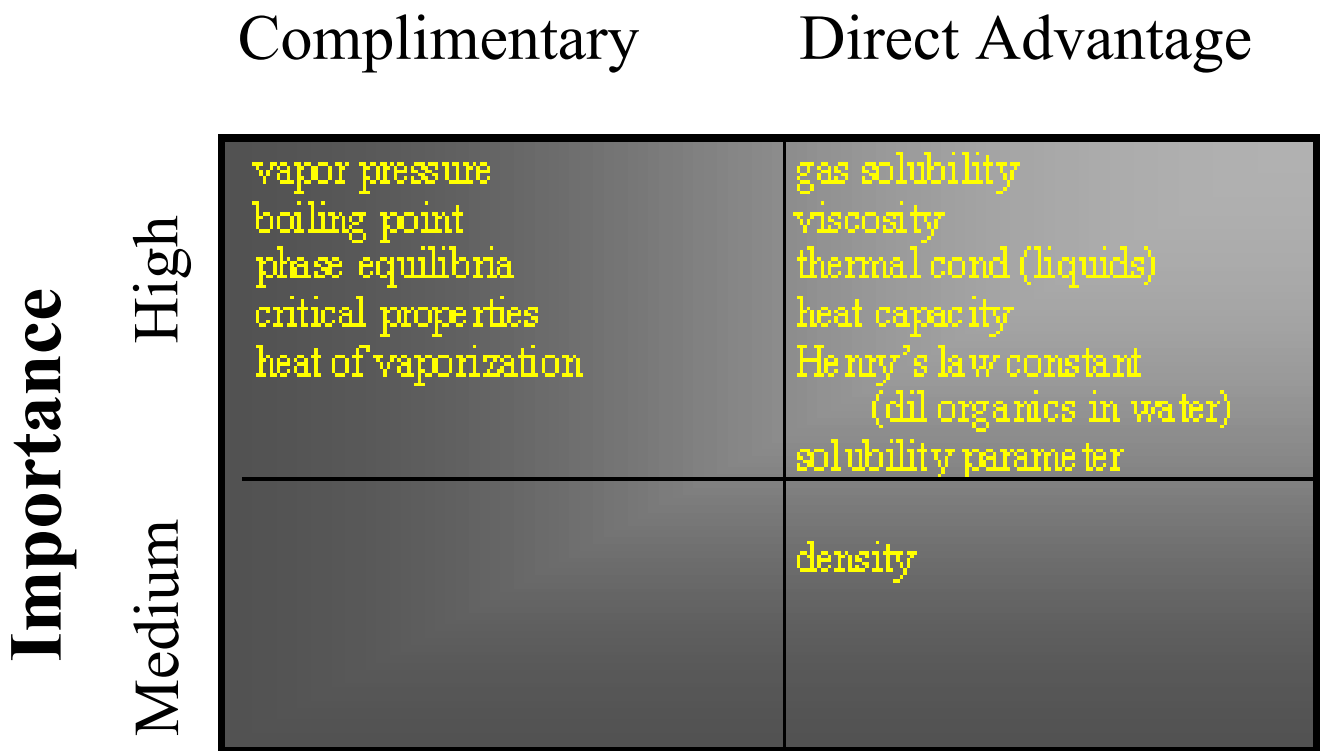
Surface Energies of Polymers



Personal Care Industry (F. Case - Colgate)

- Surfactant/surfactant & surfactant/water interaction parameters
  - *Driving forces for miscibility or phase separation*
  - *Input parameters for mesoscale models (used to predict structure and properties)*
- Absorption onto "real" surfaces (skin, cotton, grout)
- Chaotropic effects - water ordering

Chemical Industry (T. Thompson - Dow)



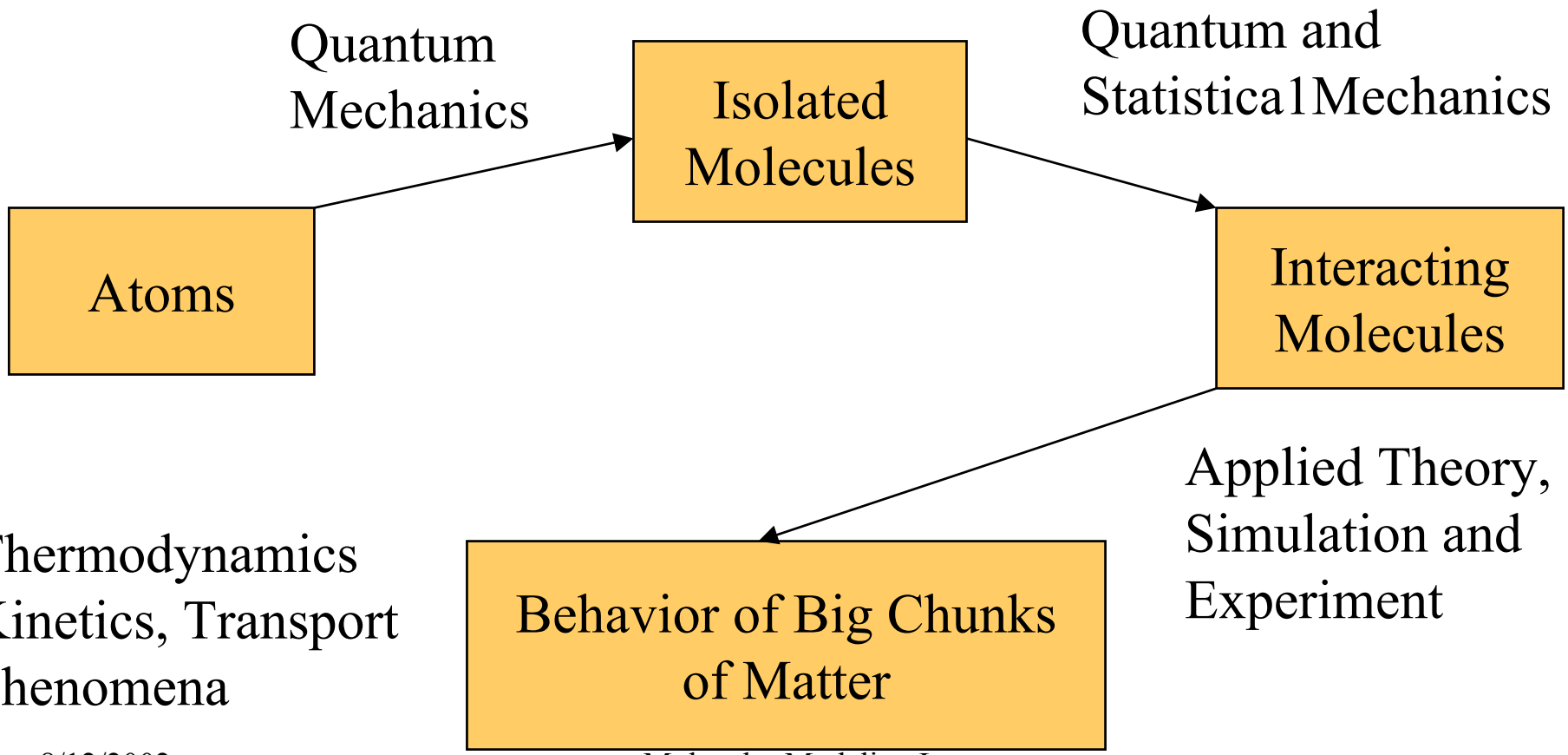
Curriculum Challenge

Maintain quality program in "classical" process oriented chemical engineering while simultaneously introducing molecular engineering concepts

Course Outline

- Review computational quantum chemistry
  - *How does Spartan Work?*
- Review Statistical Thermodynamics
  - *Structure and Ensemble Averages*
- Monte Carlo Simulation Techniques
- Molecular Dynamics Simulation

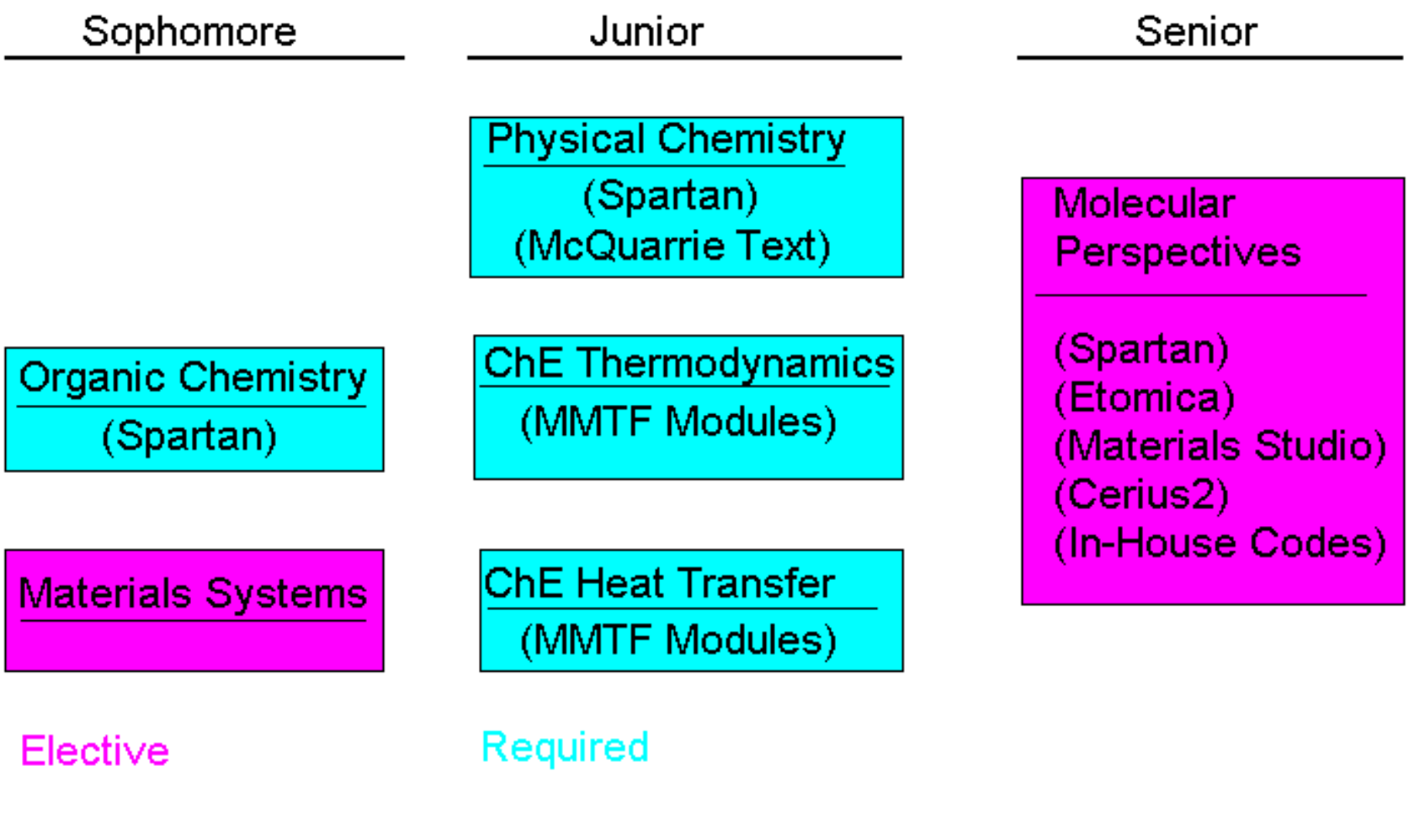
The Big Picture



The Curriculum

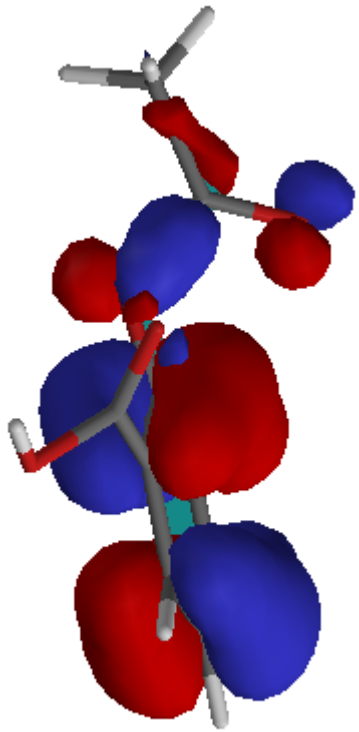
Sophomore	Junior	Senior
	Physical Chemistry (Spartan) (McQuarrie Text)	Molecular Perspectives (Spartan) (Etomica) (Materials Studio) (Cerius2) (In-House Codes)
Organic Chemistry (Spartan)	ChE Thermodynamics (MMTF Modules)	
Materials Systems	ChE Heat Transfer (MMTF Modules)	

The Curriculum



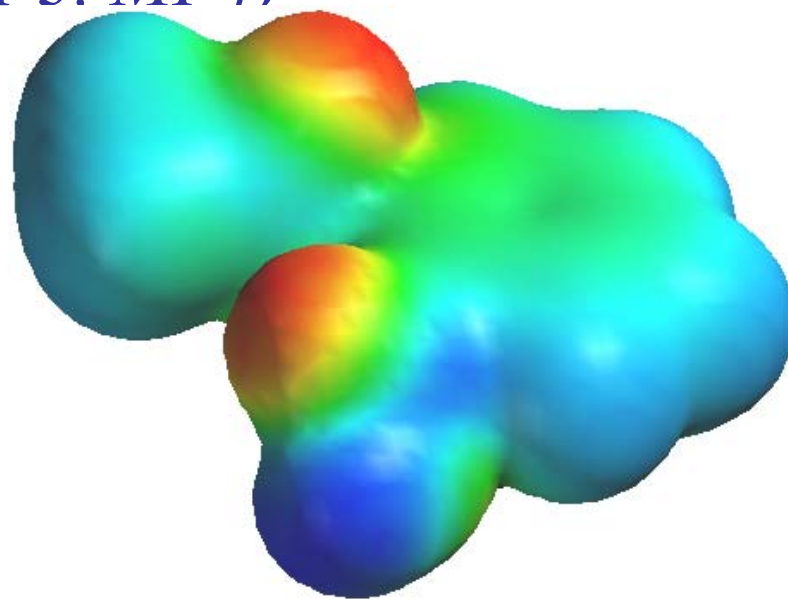
The Tools – Spartan (Wavefunction)

- Methods
  - *Molecular Mechanics (SYBL, MMFF94)*
    - » Up to 1,000 atoms
  - *Semi-empirical Molecular Orbital (MNDO, AM1, PM3, MNDO/d)*
    - » Up to 200 atoms
  - *Hartree-Fock Molecular Orbital (STO-3G, 3-21G, 6-31G*, 6-311G*)*
    - » Up to 100 atoms



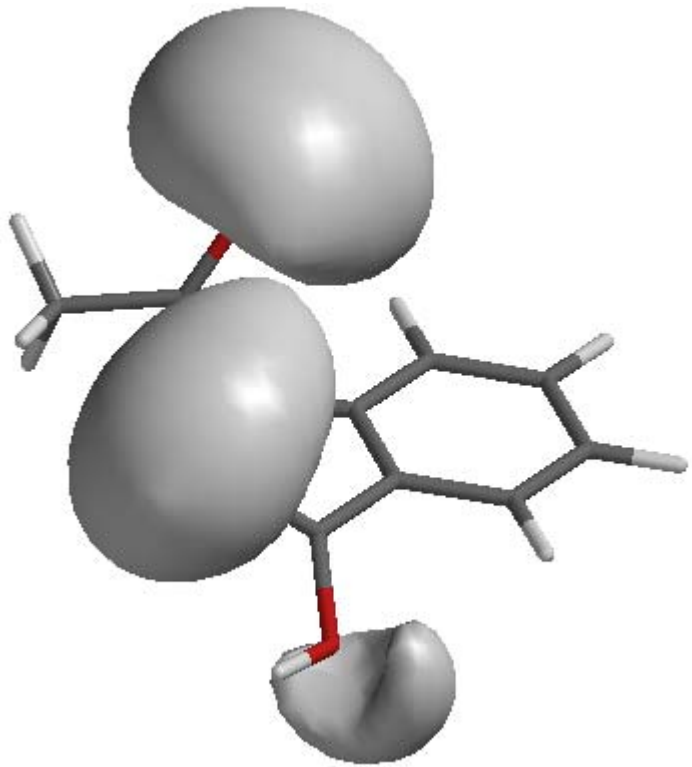
The Tools – Spartan (Wavefunction)

- Methods (continued)
  - *Density Functional (Local density models and BP, BLYP, EDF1 and B3LYP)*
    - » Up to 50 atoms
  - *Møller-Plesset (MP2, MP3. MP4)*
    - » Up to 20 atoms



The Tools – Spartan (Wavefunction)

- Tasks
  - *Energy*
  - *Equilibrium Geometry*
  - *Transition State Geometry*
  - *Normal Mode Frequencies*
  - *Conformation*
  - *Energy Profile*



The Tools - Spartan (Wavefunction)

- Properties
  - *Atomic Charges (Mullikin and Bond Orbital Charges)*
  - *Thermodynamics (Enthalpies, entropies, free energies,isotope effects, based on calculated geometries and vibrational frequencies)*
  - *Electrical (Dipole, quadrupole and higher moments, polarizabilities and hyper-polarizabilities)*
  - *Solvation (Aqueous solvation energies from SM5.4 model)*

The Tools – MMTF Modules

- CACHE – Computer Aids for Chemical Engineering (www.che.utexas.edu/cache/)
- CACHE Molecular Modeling Task Force
- Purpose
  - *Foster integration of molecular modeling in chemical engineering curricula*
- Initiatives
  - *WWW based textbook on molecular simulation*
  - *Molecular-simulation based teaching modules*
  - *FOMMS 2003*

CACHE Molecular Modeling Task Force

Peter T. Cummings, UTK
Arup Chakraborty, UCB
Ariel A. Chialvo, ORNL
Henry D. Cochran, Jr., ORNL
Juan J. DePablo, UWisc
James F. Ely, CSM
David Ford, TAMU
David A. Kofke, SUNY-Buffalo
Sanat K. Kumar, PSU
Daniel Lacks, Tulane
Edward J. Maginn, ND
A. Z. Panagiotopoulos, Princeton
Richard Rowley, BYU
Nigel Seaton, Edinburgh
Warren Seider, Penn
Randall Q. Snurr, Northwestern
Phillip R. Westmoreland, UMass

Module Creators

The Tools
- MMTF Simulation Modules

- Teaching of molecular origins of macroscopic behavior
  - *vapor pressure, viscosity, adsorption, nonideal gases, etc.*
- Web accessible, platform independent (Java)
- Production of multiple modules
  - *Infrastructure developed and first modules produced by collaborators*
  - *Additional modules to be produced by larger community*
    - » Need to provide authoring support!

The Tools – MMTF Modules

- Dynamic, interactive molecular simulation
- Supplemental material
  - *Lecture notes*
  - *Example and homework problems*
  - *Narrative description*
- Evaluation

The Tools – Etomica (Dave Kofke, UB-SUNY)

- GUI-based development environment
  - *Simulation is constructed by piecing together elements*
  - *No programming required*
  - *Results can be exported run stand-alone applet*
- Application Programming Interface (API)
  - *Library of components used to assemble a simulation*
  - *Can be used independent of development environment*

The Tools – Materials Studio (Accelrys)

- Integrated GUI for molecular dynamics, Monte Carlo and mesoscale systems interacting with state of the art force fields
  - *Intended to be a commercial rather than educational tool*
  - *Moderately steep learning curve*

Tools – Materials Studio (Accelrys)

- Can calculate a wide variety of properties
  - *Fluctuation properties*
    - » NVE, NPT and NVT ensembles
  - *Dynamic Properties*
    - » Mean squared displacement
    - » Dipole, velocity, rotational, space-time and stress-time autocorrelation functions
  - *Mechanical Properties*
    - » Static and fluctuating elastic properties

Tools - Materials Studio (Accelrys)

- Properties (continued)
  - *Structural Properties*
    - » Length, angle and dihedral distributions
    - » Pair and orientational correlation functions
    - » Concentration Profiles
    - » Voronoi tessellation
    - » X-ray and neutron scattering

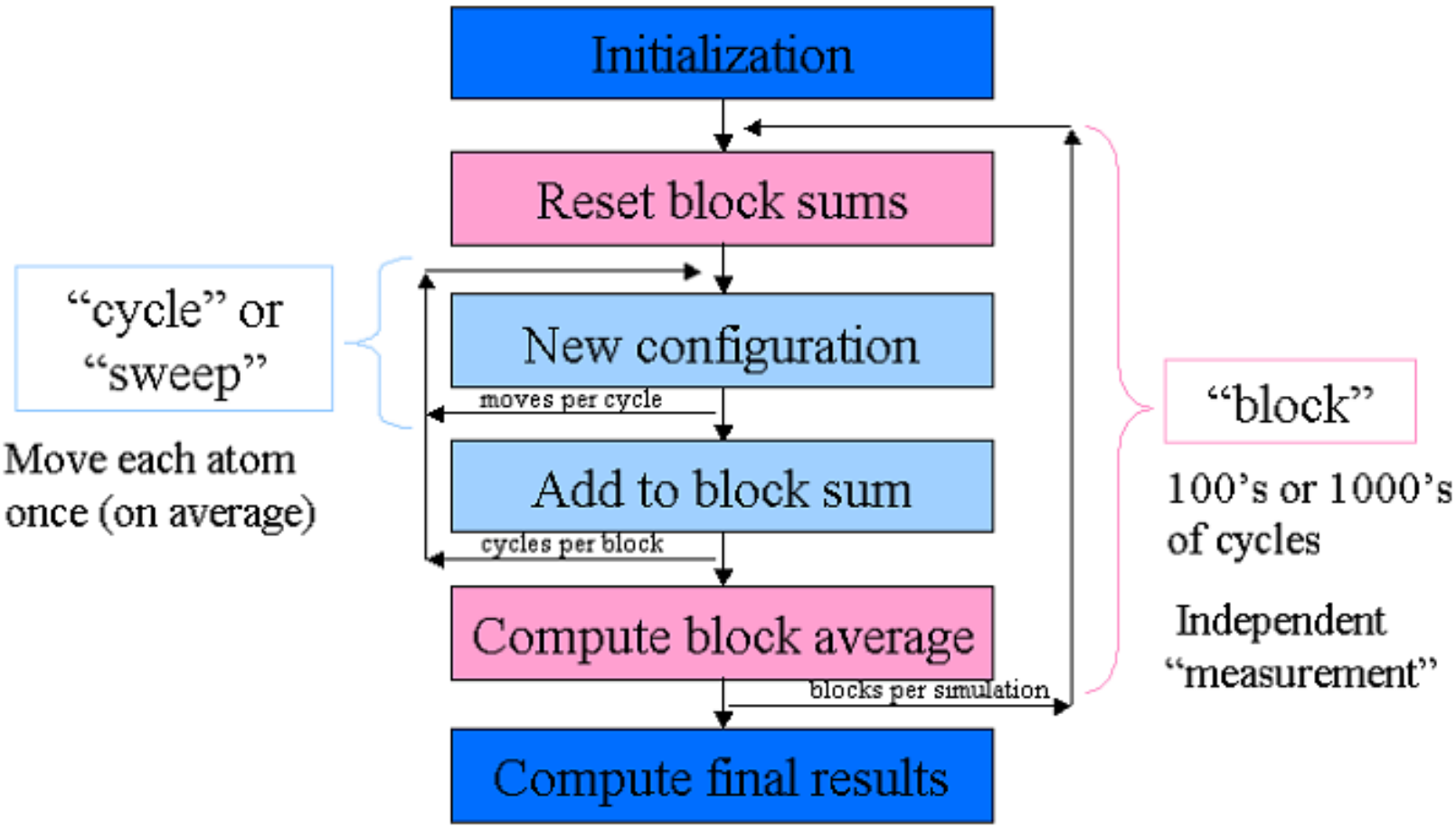
Tools – Materials Studio (Accelrys)

- Properties (continued)
  - *Energetic*
    - » Energy evolution
    - » Temperature profiles

Short(est) Course on Molecular Simulation

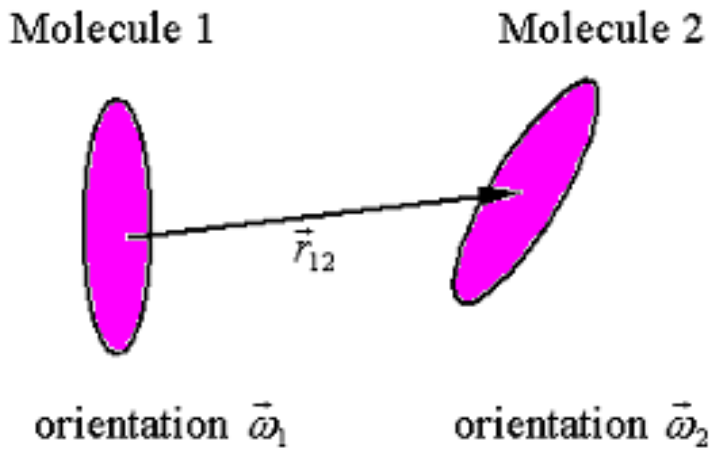
- Focus will be on molecular dynamics and Monte Carlo
- Adapted from short courses originally developed by Peter Cummings (UTK) and David Kofke (UB)

Structure of a Molecular Simulation



Background

- Assume interactions are given by pair potentials
  - *Not necessary assumption, but frequently invoked*



Background - MD

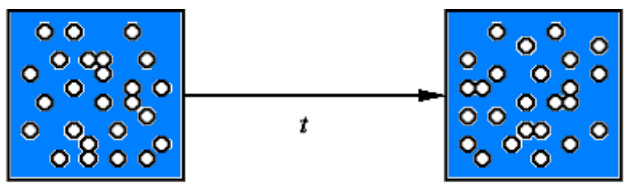
- Calculate time average of a mechanical property

$$\langle P \rangle = \lim_{t \to \infty} \frac{1}{T} \int_0^T P(t) dt$$

  - *Configurational energy*

$$E^{conf} = \sum_{1 \le i < j \le N} u(\mathbf{r}_{ij} \boldsymbol{\omega}_i \ \boldsymbol{\omega}_j)$$

- Compute by *molecular dynamics*



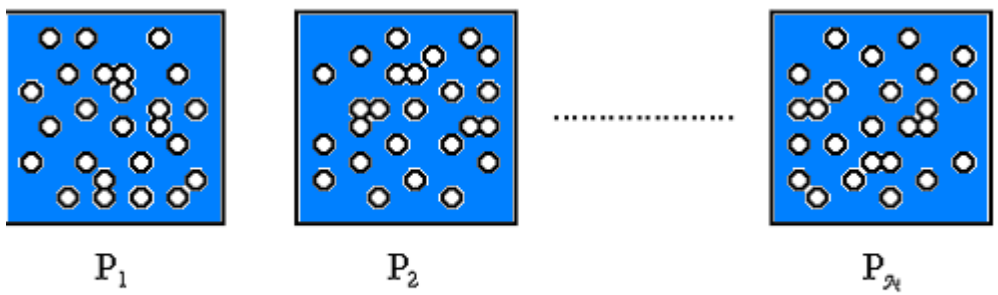
Single System Evolving in Time

Background - MC

- Alternative to time average
  - *Instantaneous average from large number (ensemble) of systems*
  - *Compute ensemble average*

$$\langle P \rangle = \frac{1}{\mathrm{N}} \sum_{i=1}^{\mathrm{N}} P_i$$

- Compute by *Monte Carlo*



Ensemble of $\mathcal{N}$ systems

Background

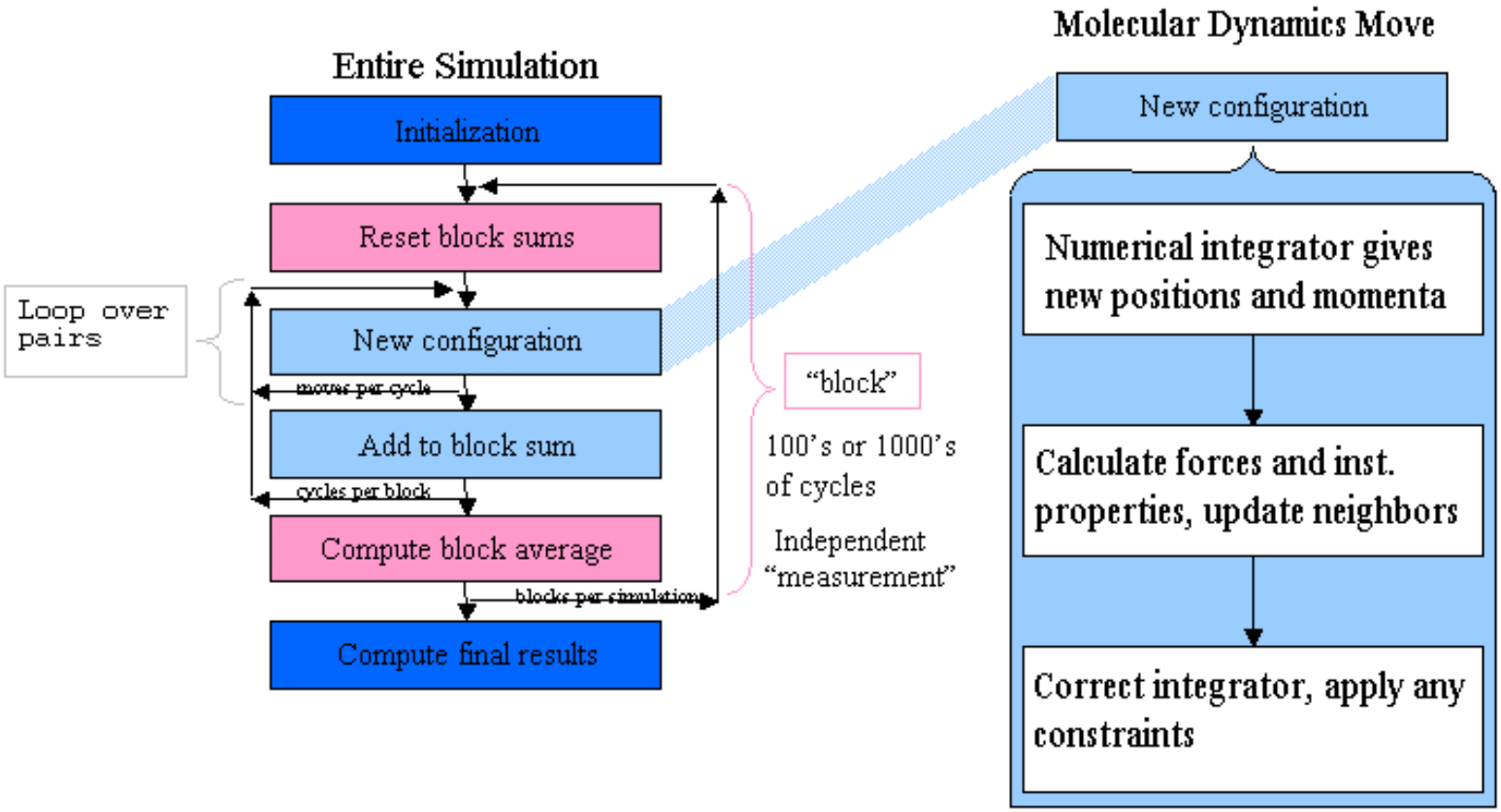
- Gibbs postulate:

  time average = ensemble average

  (MD results) = (MC results)

$$\lim_{t\to\infty}\frac{1}{T}\int_0^T P(t)dt = \frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}} P_i$$

Equilibrium Molecular Dynamics



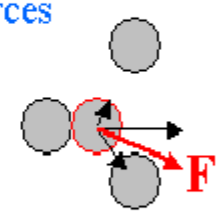
Equilibrium Molecular Dynamics

- Equations of motion in cartesian coordinates

$$\frac{d\mathbf{r}_j}{dt} = \frac{\mathbf{p}_j}{m}$$
$$\frac{d\mathbf{p}_j}{dt} = \mathbf{F}_j$$

$\mathbf{r} = (r_x, r_y)$
$\mathbf{p} = (p_x, p_y)$
} 2-dimensional space (for example)

$$\mathbf{F}_j = \sum_{\substack{i=1 \\ i \neq j}}^{N} \mathbf{F}_{ij}$$ pairwise additive forces



- Desirable features of an integrator
  - minimal need to compute forces (a very expensive calculation)
  - good stability for large time steps
  - good accuracy
  - conserves energy and momentum

Equilibrium Molecular Dynamics

- Common Inegrators
  - *Gear 4^{th} or 5^{th} order predictor corrector*
  - *Verlet*
  - *Velocity Verlet*
  - *Beeman*
  - *RESPA (multi-time step)*
- Integrators should be symplectic
  - *Conserve energy and phase space volume*

Equilibrium Molecular Dynamics

- Predictor-Correctors are easiest to constrain
  - *Gaussian constraints for temperature, pressure, bond length, ...*
- Other constraint methods for temperature and pressure are common
  - *Nose-Hoover*
  - *Anderson*

Equilibrium Molecular Dynamics

- (LJ) Force calculation

$$u(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right], \qquad r < R_c$$

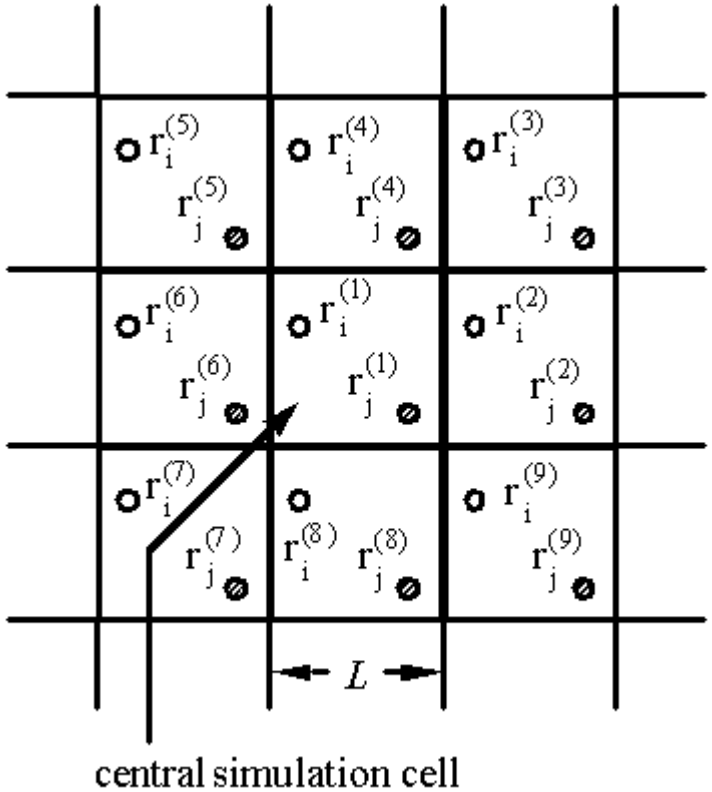
$$= 0, \qquad r > R_c$$

$$\mathbf{F}(\mathbf{r}) = -\nabla u(r) = \frac{24\varepsilon}{\sigma}\left[2\left(\frac{\sigma}{r}\right)^{14} - \left(\frac{\sigma}{r}\right)^{8}\right]\frac{\mathbf{r}}{\sigma}$$

$$\mathbf{F}^*(\mathbf{r}_{ij}) \equiv \mathbf{F}^*_{ij} = 24\left[\frac{2}{r^{*12}_{ij}} - \frac{1}{r^{*6}_{ij}}\right]\frac{1}{r^{*2}_{ij}}\mathbf{r}^*_{ij}$$

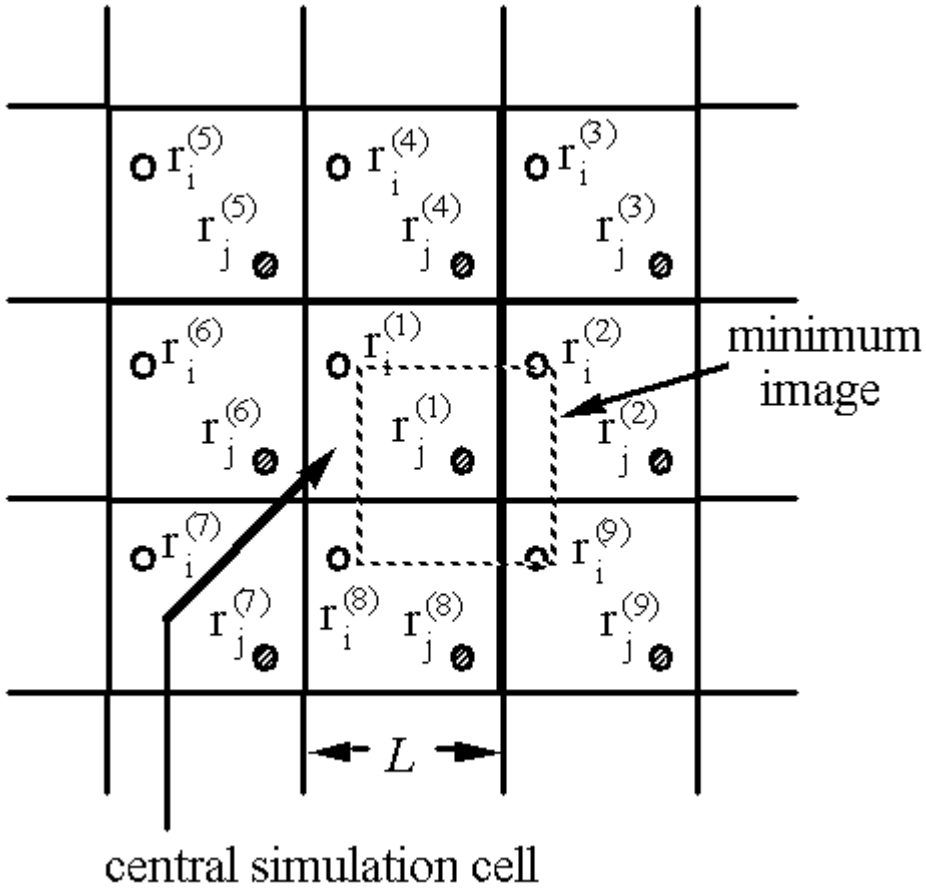
$$\mathbf{F}^*_{ji} = -\mathbf{F}^*_{ij} \qquad \mathbf{F}^*_i = \sum_{j\neq i}\mathbf{F}^*_{ij}$$

Equilibrium Molecular Dynamics



Periodic Cell Geometry

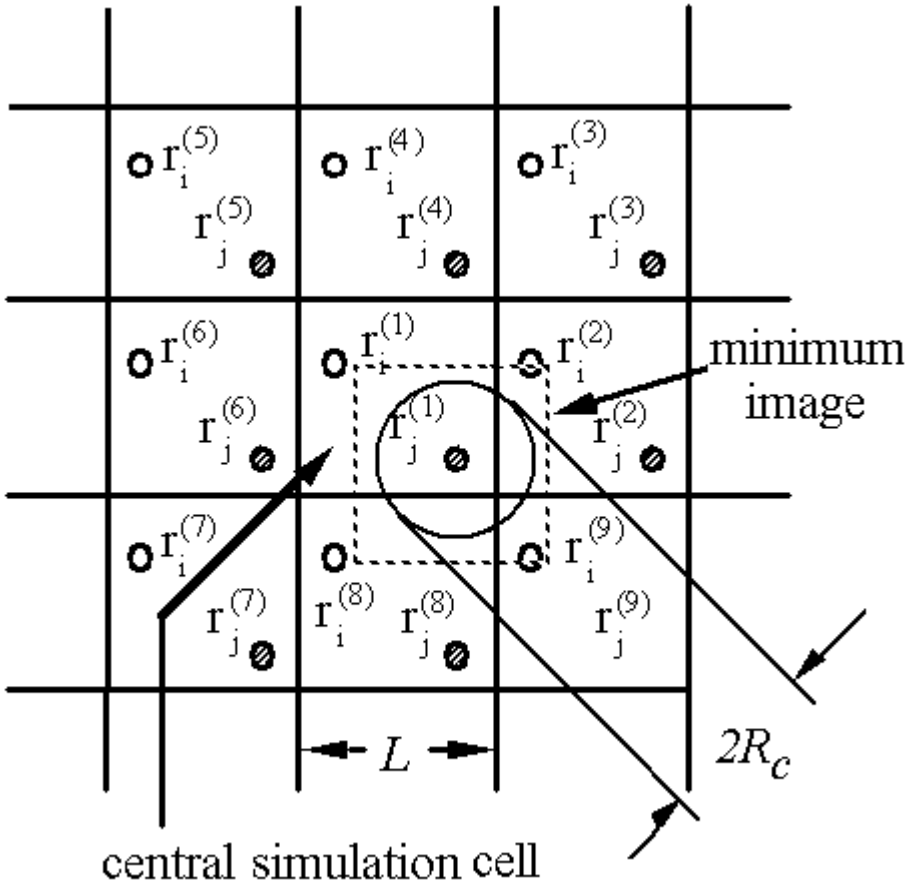
Equilibrium Molecular Dynamics



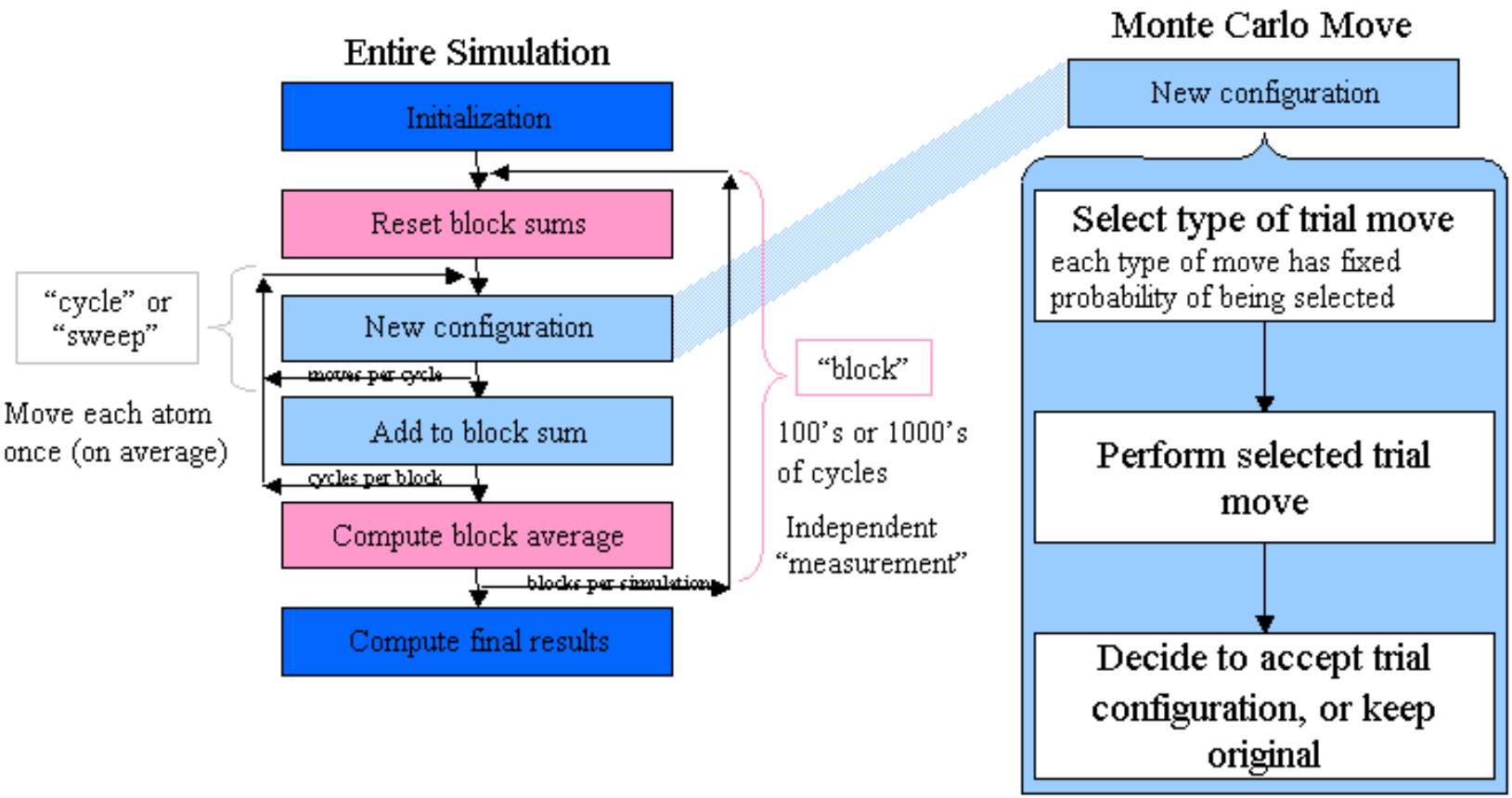
Minimum Image Convention

Equilibrium Molecular Dynamics

Cut-off radius, R_c



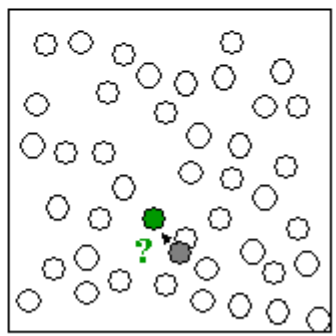
Monte Carlo Simulation



Monte Carlo Simulation

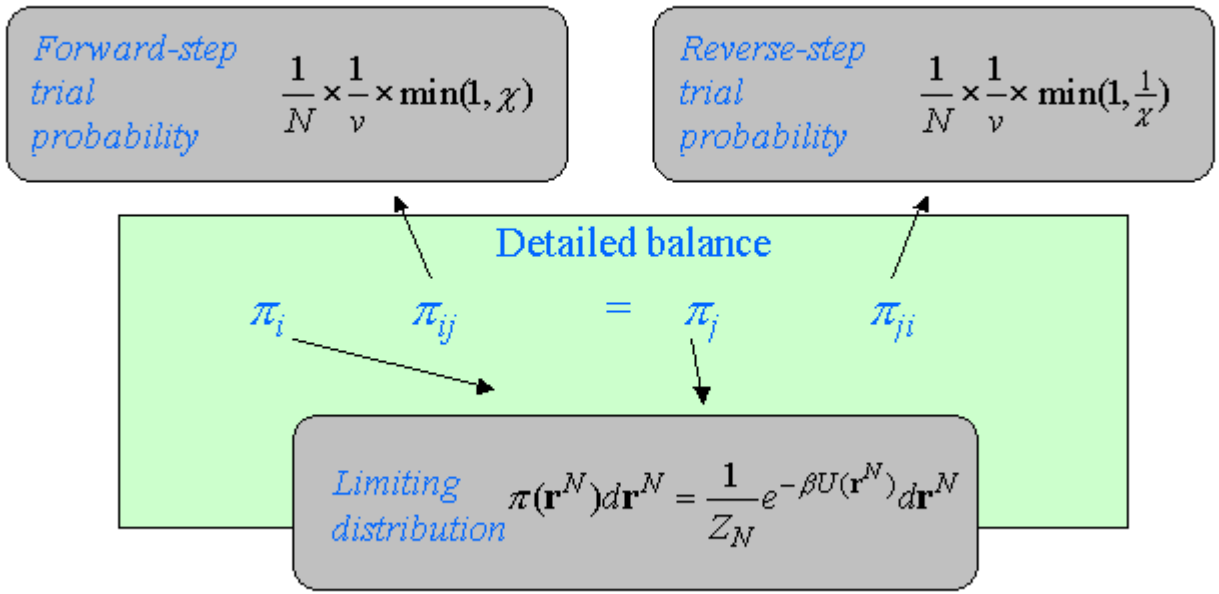
- Displacement
  - Gives new configuration of same volume and number of molecules
  - Basic trial:
    - displace a randomly selected atom to a point chosen with uniform probability inside a cubic volume of edge 2d centered on the current position of the atom

Consider acceptance of new configuration



Monte Carlo Simulation

- Detailed Balance



Monte Carlo Simulation

- Detailed Balance

Forward-step trial probability $\frac{1}{N} \times \frac{1}{v} \times \min(1, \chi)$

Reverse-step trial probability $\frac{1}{N} \times \frac{1}{v} \times \min(1, \frac{1}{\chi})$

Detailed balance

$\pi_i \quad \pi_{ij} \quad = \quad \pi_j \quad \pi_{ji}$

$$\frac{e^{-\beta U^{old}} d\mathbf{r}^N}{Z_N} \times \frac{1}{N} \times \frac{1}{v} \times \min(1, \chi) = \frac{e^{-\beta U^{new}} d\mathbf{r}^N}{Z_N} \times \frac{1}{N} \times \frac{1}{v} \times \min(1, \frac{1}{\chi})$$

Limiting distribution $\pi(\mathbf{r}^N) d\mathbf{r}^N = \frac{1}{Z_N} e^{-\beta U(\mathbf{r}^N)} d\mathbf{r}^N$

Monte Carlo Simulation

- Detailed Balance

Forward-step trial probability $\frac{1}{N} \times \frac{1}{v} \times \min(1, \chi)$

Reverse-step trial probability $\frac{1}{N} \times \frac{1}{v} \times \min(1, \frac{1}{\chi})$

Detailed balance

$$\pi_i \quad \pi_{ij} \quad = \quad \pi_j \quad \pi_{ji}$$

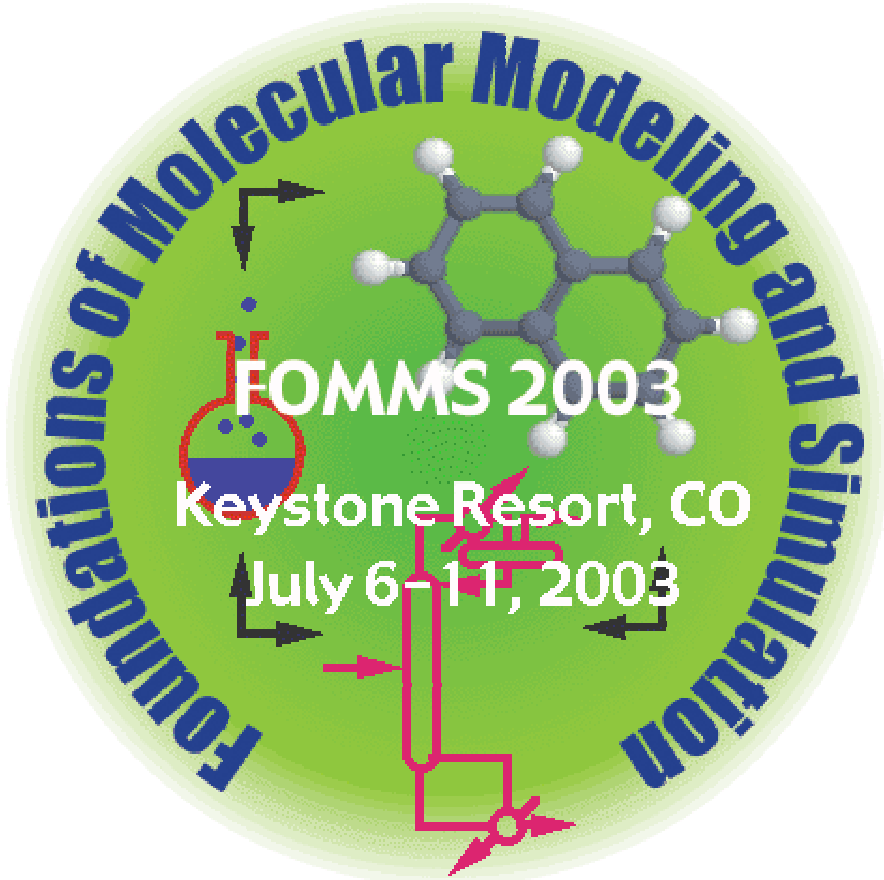
$$\frac{e^{-\beta U^{old}} d\mathbf{r}^N}{Z_N} \frac{1}{N} \frac{1}{v} \min(1, \chi) = \frac{e^{-\beta U^{new}} d\mathbf{r}^N}{Z_N} \frac{1}{N} \frac{1}{v} \min(1, \frac{1}{\chi})$$

$$e^{-\beta U^{old}} \chi = e^{-\beta U^{new}}$$

$$\chi = e^{-\beta (U^{new} - U^{old})}$$ Acceptance probability

Foundations of Molecular Modeling and Simulation (FOMMS)

- Sponsored by:
  - *CACHE*
  - *CoMSEF*
- Second International Meeting
  - *J. F. Ely, Chair*
  - *G. A. Jackson, co-Chair*
  - *P. T. Cummings, Senior Advisor*
  - *P. R. Westmoreland, Senior Advisor*



FOMMS 2000 Movie

Molecular Modeling and Simulation

Break Time

Molecular Perspectives

- Outcomes (hoped for!)
  - *Synthesis of chemistry and chemical engineering*
  - *Higher order thinking*
  - *Use of important new tools*

Molecular Perspectives Sophomore Year

- Organic Chemistry
  - *Introduction to computational quantum chemisty*
- Is the time now?

  "We are perhaps not far removed from the time when we shall be able to submit bulk of chemical phenomena to calculation."

  Joseph Louis Gay-Lussac 1778-1850

Molecular Perspectives Sophomore Year

- Objectives
  - *Structural Visualization*
  - *Isosurfaces*
    - » Molecular Orbitals
    - » Electron Densities
    - » Electrostatic Potentials
  - *Property Maps (color coded isosurfaces)*
  - *Animations*
  - *Intro to molecular modeling*

Molecular Perspectives Junior Year

- MMTF Modules
  - *VLE Module (Richard Rowley)*
  - *Heat Transfer (Randy Snurr)*
- Introduction to Materials Studio
  - *Virial Coefficients*
  - *VLE*
- Physical Chemistry
  - *Spartan Applications (e.g., HF potential)*
  - *Heavy dose of quantum chemistry*

Molecular Perspectives Senior Year

- Capstone Course
  - *Properties of fluids and solids: molecular structure prediction methods, QSAR/QSPR*
  - *Thermochemical computational quantum chemistry*
  - *Intermolecular forces and configurational properties*
  - *Equilibrium molecular dynamics*
  - *Monte Carlo techniques*

Molecular Perspectives Senior Year

- Resources
  - *22 Seat Computer Lab (dual Athlon 1800)*
  - *21 Node IBM SP2*
  - *22 Seat Site License for Materials Studio and Cerius^2*
  - *50 Seat Site License for Spartan*
  - *50 node PIII Beowulf (on-line 9/1/02)*
  - *20 node Athlon 1800 Beowulf (on-line 10/1/02)*

Molecular Perspectives Senior Year

- Example Projects
  - *Group Contributions*
    - » Empirical – Lyderson/Joback (Crit, BP, MP)
    - » Empirical – Benson
    - » Quantum – CQC Calculation of Benson Groups
    - » Quantum – QSAR/QSPR study of BP and MP
  - *Etomica Demonstrations*
  - *Materials Studio*
    - » Diffusion in Polymers
    - » Polymer Miscibility

Group Contribution Project Quantum Mechanics

- Arguably the greatest discovery of the 20th century
  - *Extended classical ideas into the behavior of subatomic, atomic and molecular species*
- Forms the foundation of ALL chemistry
- Has only become computationally tractable in the last 10 years
  - *J. A. Pople Nobel Prize in 1998*

Macroscopic Approximations to Quantum Behavior

- Scientists realized early on that most properties of larger molecules can be considered as being made up of additive contributions from individual atoms or bonds in the molecules
  - *Group Contribution Approximations-the addition of empirically-derived quantities, each characteristic of a chemical sub-unit of the compound in question to arrive at a property*
    - » **The fundamental assumption is additivity of these contributions.**

Group Contribution Methods

- These methods are largely empirical although in some cases, theoretical knowledge about the inter-dependence of material properties may be used as a guide in their development
  - *Physical basis is that the forces between atoms are very short ranged, e.g., 1-5 A^o*
- The definition of constituent "groups" is very subjective
  - *At one extreme only the atoms need to be specified*
  - *At the other, every molecule is a group*

Group Contribution Classes

- Class I: Fundamental (Independent of temperature and time)
  - *Molar Mass*
  - *Number of Backbone or Carbon Atoms*
  - *van der Waals volume*
- Class II: Phase Transition
  - *Transition temperatures (T_c, T_b, T_g, T_m)*
  - *Transition volumes (V_c, V_m)*
  - *Transition pressures (p_c, p_t)*
  - *Parachor (Surface tension)*

Group Contribution Classes

- Class III: Internal Energy/Stability
  - *Heat Capacity (ideal gas and condensed)*
  - *Heat of Formation*
  - *Standard Entropy*
- Class IV: Electromagnetic
  - *Molar Polarization*
  - *Optical Refraction*
  - *Magnetic Susceptibility*

Hierarchy of Additivity Laws

- Zeroth Order -- Additivity of atom properties (exact for molecular weight)
- First Order -- Bond Properties
- Second Order -- Group Properties
  - *corrections for non-bonded interactions*
  - *gauche interactions (anything bigger than H)*
  - *ring corrections (proximity of distant groups)*

Lyderson/Joback Method Phase Transition Properties

- Critical Temperature (K)

$$T_c = T_b \left[0.584 + 0.965 \sum \Delta_T - \left(\sum \Delta_T \right)^2 \right]^{-1}$$

- Critical Pressure (bar)

$$P_c = \left(0.113 + 0.0032 N_A - \sum \Delta_p \right)^{-2}$$

- Critical Volume (cm^3/mol)

$$V_c = 17.5 + \sum \Delta_v$$

Lyderson/Joback Method Phase Transition Properties

- Normal boiling point (K)

$$T_b = 198 + \sum \Delta_b$$

- Freezing Point (K)

$$T_f = 122 + \sum \Delta_f$$

- The Δ_x are sums of group contributions

Lyderson/Joback Method Phase Transition Properties

- The boiling and freezing temperatures are difficult because they have a large entropic component

$$T_{trans} = \frac{\Delta H_{trans}}{\Delta S_{trans}}$$

  - *S is a function of molecular symmetry, i.e., depends on the number of orientations that the molecule can have in the liquid*

Example

- Estimate the transition properties of 2-methylhexane
  - *Tb =371.6, Tf=154.9, Tc=540.2, Pc=27.36, Vc=421*
  - *Groups: -CH3 (3), >CH2 (3), >CH- (1)*
  - *Sums:*

    Tc: 3(0.0141)+3(0.0189)+0.0164 =0.1154

    Pc: 3(0.0012)+3(0.0000)+0.0020 =0.0056

    Vc: 3(65)+3(56)+41 =404

    Tb: 3(23.58)+3(22.88)+21.74 =161.12

    Tf: 3(-5.10)+3(11.27)+12.64 =18.81

Example (cont'd)

- Results:
  - *Tc: 544.8 K vs. 540.2*
  - *Pc: 30.52 bar vs. 27.36*
  - *Vc: 417.5 cm^3/mol vs. 421*
  - *Tb: 359.1 K vs. 371.6*
  - *Tf: 140.8 K vs. 154.8*
- OK but not perfect. Can we do better?
  - *Probably -- QSAR/QSPR*

QSAR and QSPR: Applications in Chemical Engineering

QSAR $\Rightarrow$ Quantitative Structure-Activity Relationships

QSPR $\Rightarrow$ Quantitative Structure-Property Relationships

Goal: *Correlation and prediction of activity (biological, pharmaceutical, etc.) or properties (physicochemical) with molecular characteristics*.

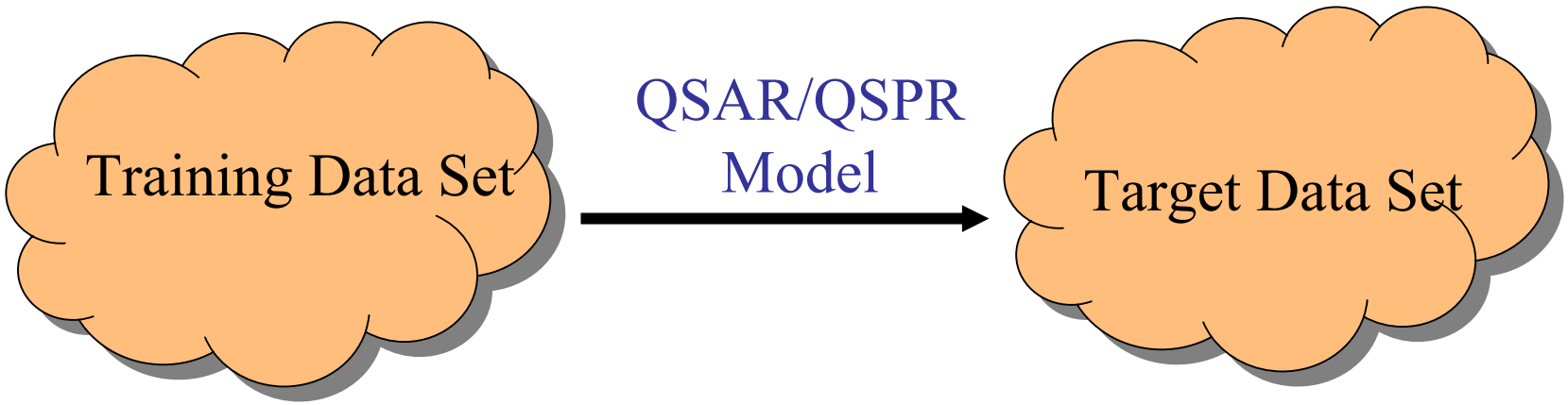
Common Families of Molecular Characteristics (Descriptors)

- Spatial
  - *Molecular Area*
  - *Molecular Volume*
  - *Density*
  - *Radius of Gyration*
- Structural
  - *Rotatable Bonds*
  - *Molecular Weight*
  - *Symmetry Number*
  - *Hydrogen Bond Donor / Acceptor*

Descriptor Families (cont'd.)

- Thermodynamic
  - *Heat of Formation*
  - *Partition Coefficient (octanol/water)*
  - *Molecular Refractivity*
- Electronic (Can be estimated from Spartan!)
  - *Dipole Moment*
  - *Charge Separation (Q^+/Q^-)*
  - *LUMO*
  - *HOMO*

Predictive Relationships



Activity $\Leftarrow\Rightarrow$ Structure

Properties $\Leftarrow\Rightarrow$ Structure

Developing Predictive Correlations

- Simple linear

$$Y = a_1 X_1 + a_2 X_2 + \ldots + a_n X_n$$

- Linear with cross-terms

$$Y = a_1 X_1 + a_2 X_2 X_3 + \ldots + a_n X_m X_n$$

- Linear with general terms

$$Y = a_1 f_1\left(\{X_j\}\right) + a_2 f_2\left(\{X_j\}\right) + \ldots + a_n f_n\left(\{X_j\}\right)$$

Developing Predictive Correlations

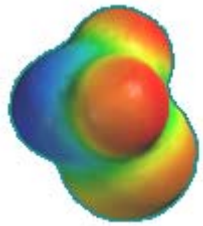
- Determine model parameters using the training set data
  - *Choice of training set and model is an open ended problem*
  - *Stepwise regression techniques can be used but in many cases Excel will suffice*
- Predict properties of other materials not included in the training set
  - *If predictions are not satisfactory, refine the model and/or training set*

Example: Predicting Boiling and Melting Points of Chlorofluorocarbons

- Question: What molecular properties should be significant in a QSPR model for boiling and/or melting points?
- Answer: Phase transition properties clearly depend upon intermolecular forces (and maybe other things!)
  - ***Dipole/Dipole Interactions***
  - ***Dipole/Induced Dipole Interactions***
  - ***Dispersion Forces***
  - ***Hydrogen Bonding***

Force - Property Relationships

Force	Property
Dipole/Dipole	Dipole Moment
Induced Dipole /Dipole	Polarizability
Dispersion	Molecular Area or Volume
Hydrogen Bonding	Hydrogen and Electronegative Atoms



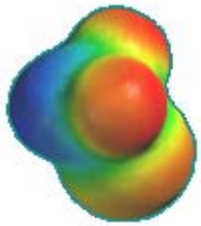
Boiling and melting points: ***J. Wei, Ind. Eng. Chem. Res. 2000, 39***

Polarizability: Cerius^2, Molecular Simulations, Inc.

Dipole, Area: Spartan, Wavefunction, Inc. (ab-initio, 3-21G)

molecule	Apol	Dipole	Area	BP (K)	MP (K)
CH4	654	0	50.937	111.6	90.7
CH3Cl	2150	1.628	69.596	249.1	175.4
CH3F	718	1.632	55.891	194.7	131.3
CH2ClF	2210	1.861	75.34	264	140.1
CF4	911	0	69.537	145.1	89.6
CCl4	6620	0	121.798	349.9	250.1
CH2F2	782	1.892	61.179	221.5	137.1
CHF3	847	1.691	65.771	191	118
CHClF2	2270	1.623	78.655	232.4	115.4
CClF3	2340	0.291	83.225	191.7	92.1
CH2Cl2	3640	1.691	85.957	313.1	178
CHCl2F	3700	1.499	89.354	282	138.1
CCl2F2	3770	0.345	96.198	243.3	115.1
CHCl3	5130	1.316	103.083	334.2	209.5
CCl3F	5190	0.322	107.266	296.8	162

QSAR BP Correlation



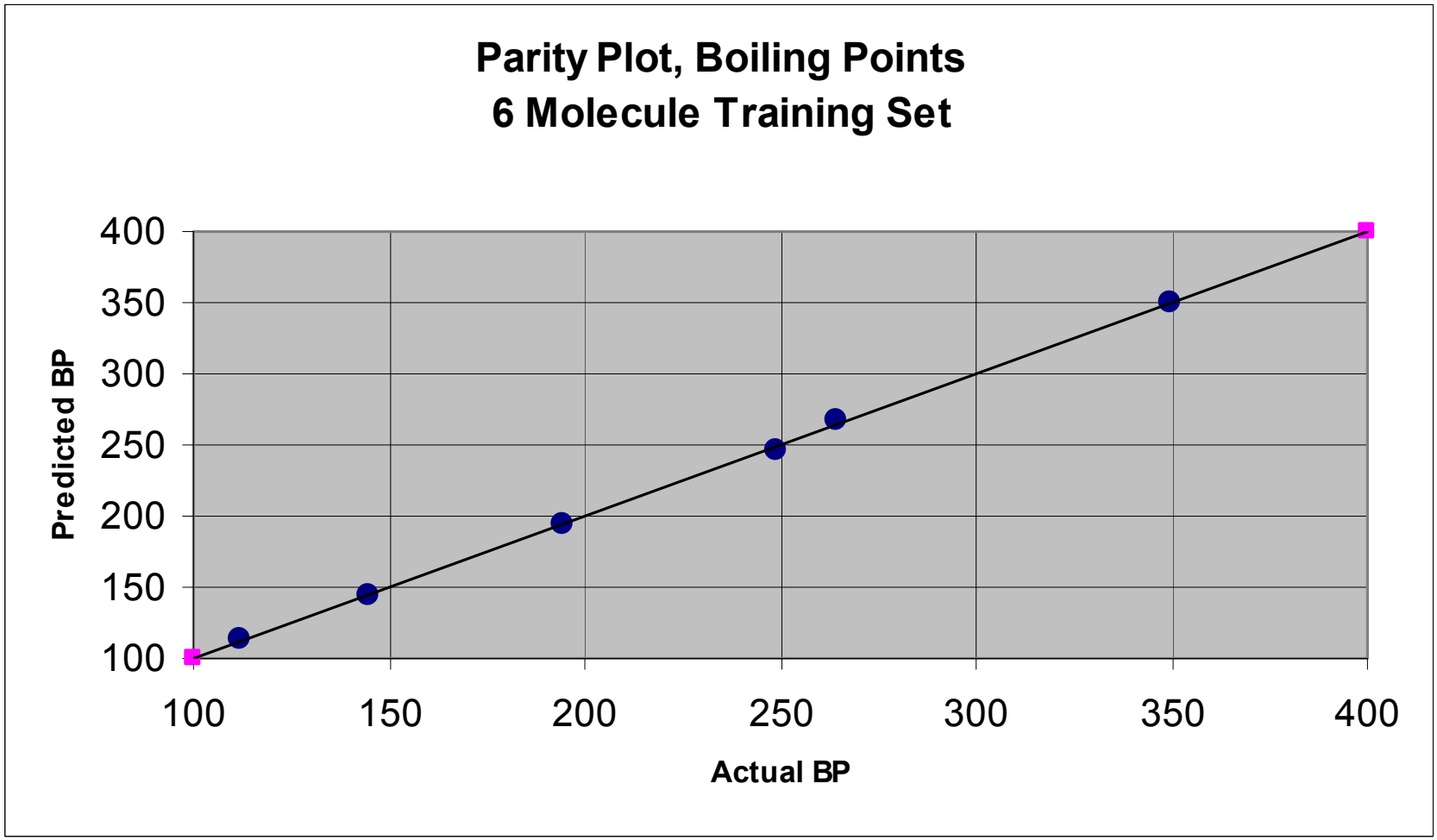
- Mathematical Relation

$$BP = a + b(polarizability) + c(dipole\,moment) + d(area)$$

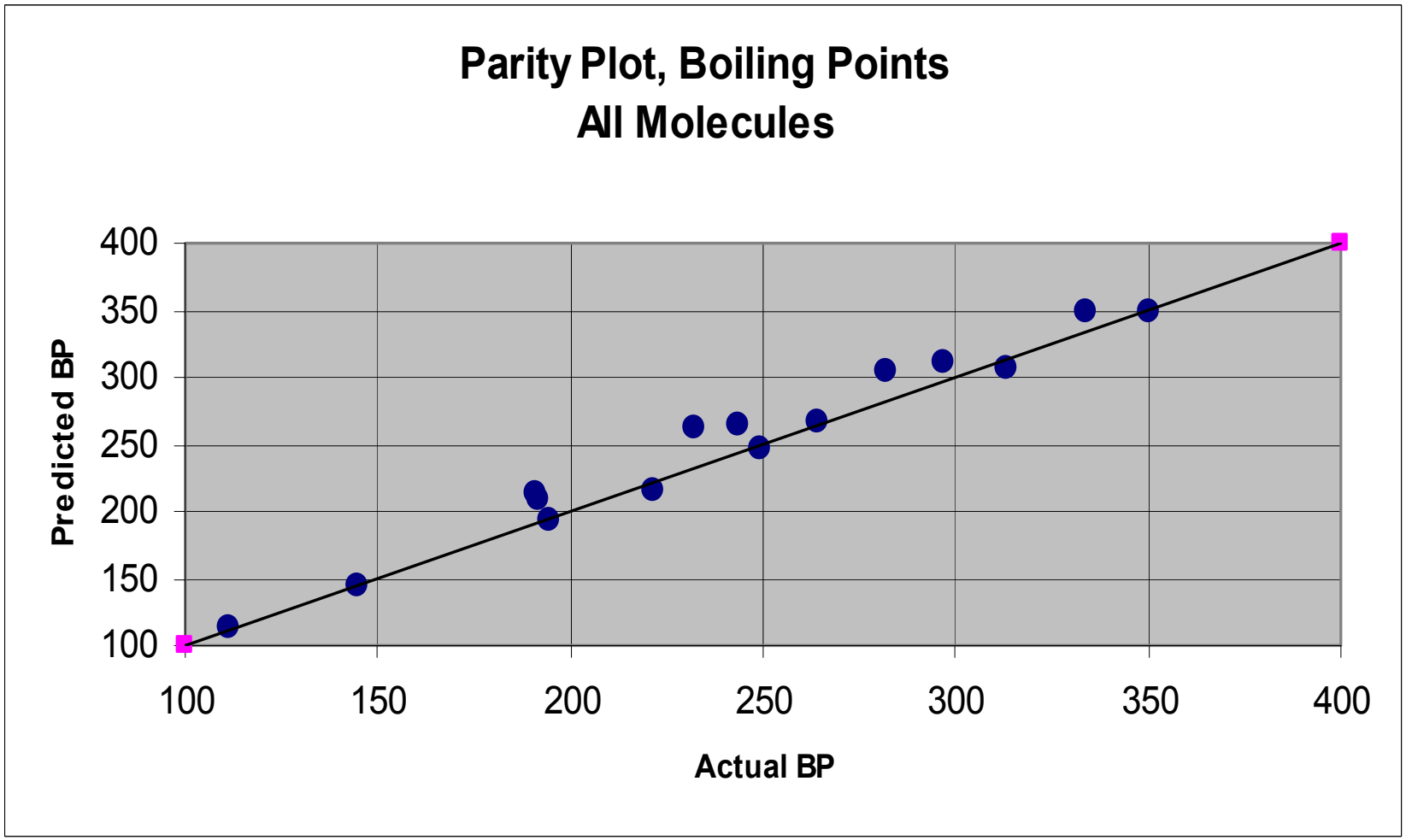
- Training set

CH_4, CH_3Cl, CH_3F, CH_2ClF, CF_4, CCl_4

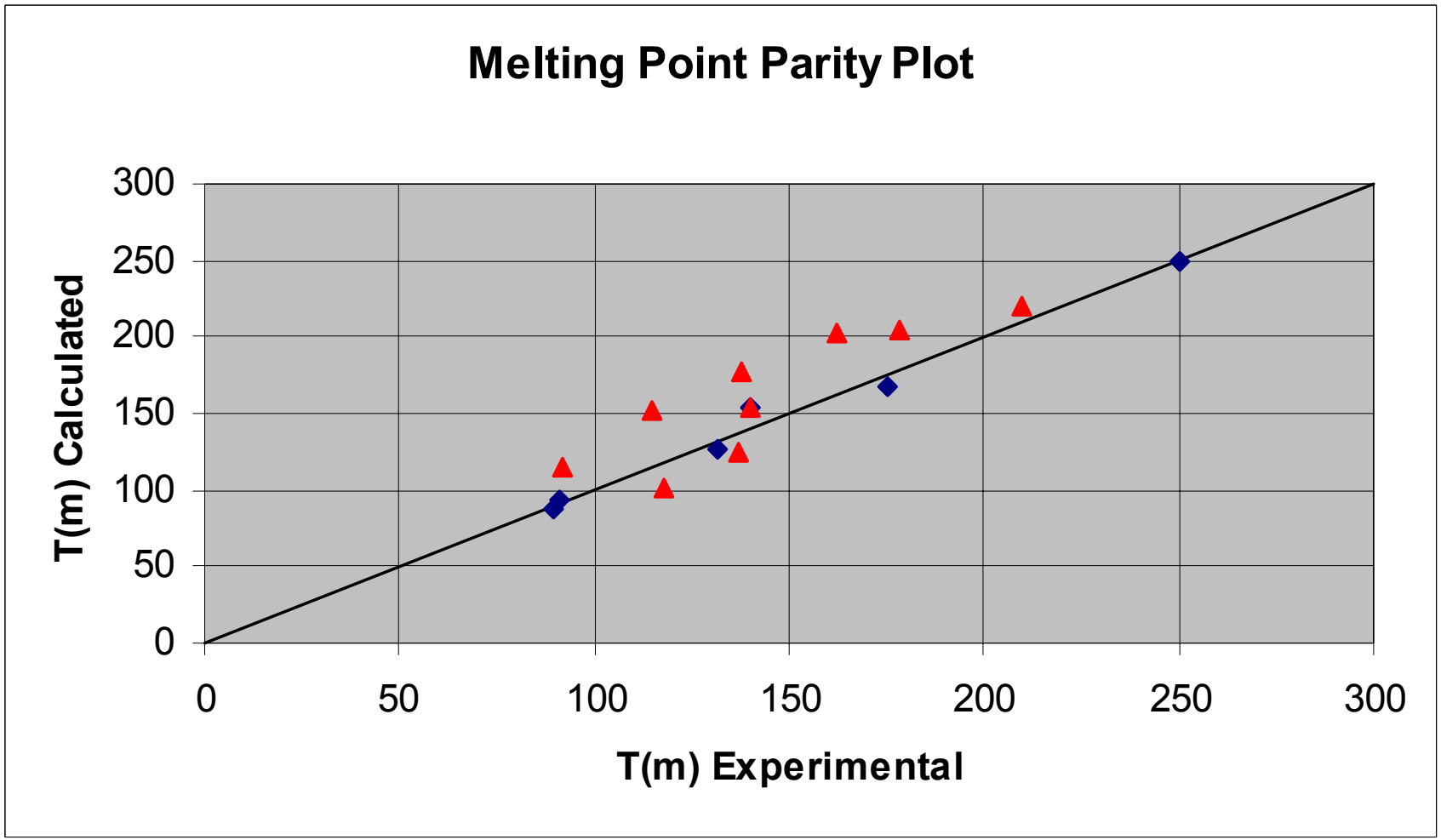
Boiling Point Training Set



Boiling Point Results



Melting Point



Etomica Project

- Download software from Dave Kofke's web-site

  www.ccr.buffalo.edu/etomica

  - *Run samples*
  - *Build Simple 2-d MD Simulation*
  - *Build Simple 2-d MC Simulation*

Materials Studio

- Simulation of Polymers
  - *Polymer Miscibility and Cohesive Energy Density*
  - *Diffusion of Gases in Polymers*