Statistical Mechanics & Simulations

Jean-Christophe SOETENS

Department of Chemistry
University of Bordeaux

jean-christophe.soetens@u-bordeaux.fr

Academic year 2023-2024

Organization

« Statistical Mechanics & Simulations »

- *Where ?* Here… A10 – Lavoisier

- *What ?* Course, without overdoing things…
  Practice : a lot...

- *Examination* :

Final exam (3h)	x 0.7	(and 2d chance)
Intermediate	x 0.3 (0.15 + 0.15)	(definitive)

Course materials : *http://theo.ism.u-bordeaux.fr/J-C.Soetens/teach.html*

Aims of this course

The purpose of this course : ***« Statistical Mechanics & Simulations »***

- To become familiar with the basics concepts underlying molecular simulations
- To understand the possibilities and limitations of (present day) simulations
- To understand the machinary of the two main methods :

  Molecular Dynamics simulations

  Monte Carlo simulations

- To be able to understand a scientific article dealing with molecular simulation and to discuss with experts in the field
- To give you some ideas on how to start your own studies
- To become familiar with a numerical scientific environment
  (software, programming, graphics, etc)

Summary

PART I : « Statistical Mechanics & Simulations »

I. Overview of Statistical Mechanics & Molecular Simulations

II. Molecular Dynamics Simulations

III. Monte Carlo methods

IV. « Outputs » : extracting properties from simulations

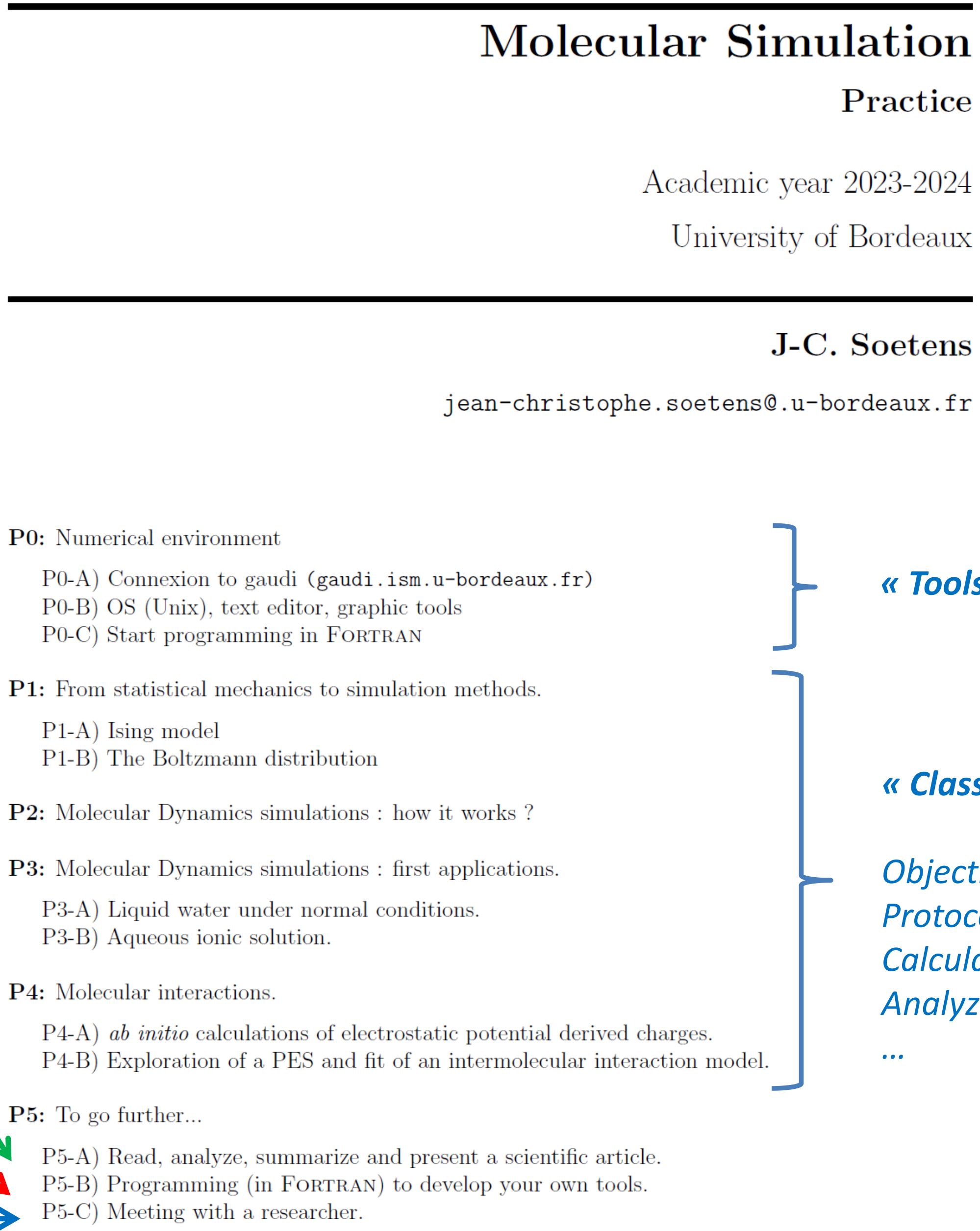
V. Initiation to statistical thermodynamics

Summary

PART II :« Force Fields & Molecular Interactions »

Summary

PRACTICE



Molecular Simulation

Practice

Academic year 2023-2024

University of Bordeaux

J-C. Soetens

`jean-christophe.soetens@u-bordeaux.fr`

P0: Numerical environment

- P0-A) Connexion to gaudi (`gaudi.ism.u-bordeaux.fr`)
- P0-B) OS (Unix), text editor, graphic tools
- P0-C) Start programming in FORTRAN

P1: From statistical mechanics to simulation methods.

- P1-A) Ising model
- P1-B) The Boltzmann distribution

P2: Molecular Dynamics simulations : how it works ?

P3: Molecular Dynamics simulations : first applications.

- P3-A) Liquid water under normal conditions.
- P3-B) Aqueous ionic solution.

P4: Molecular interactions.

- P4-A) *ab initio* calculations of electrostatic potential derived charges.
- P4-B) Exploration of a PES and fit of an intermolecular interaction model.

P5: To go further...

- P5-A) Read, analyze, summarize and present a scientific article.
- P5-B) Programming (in FORTRAN) to develop your own tools.
- P5-C) Meeting with a researcher.

« Tools »

« Classic part »

Objectives
Protocols
Calculations
Analyzes
…

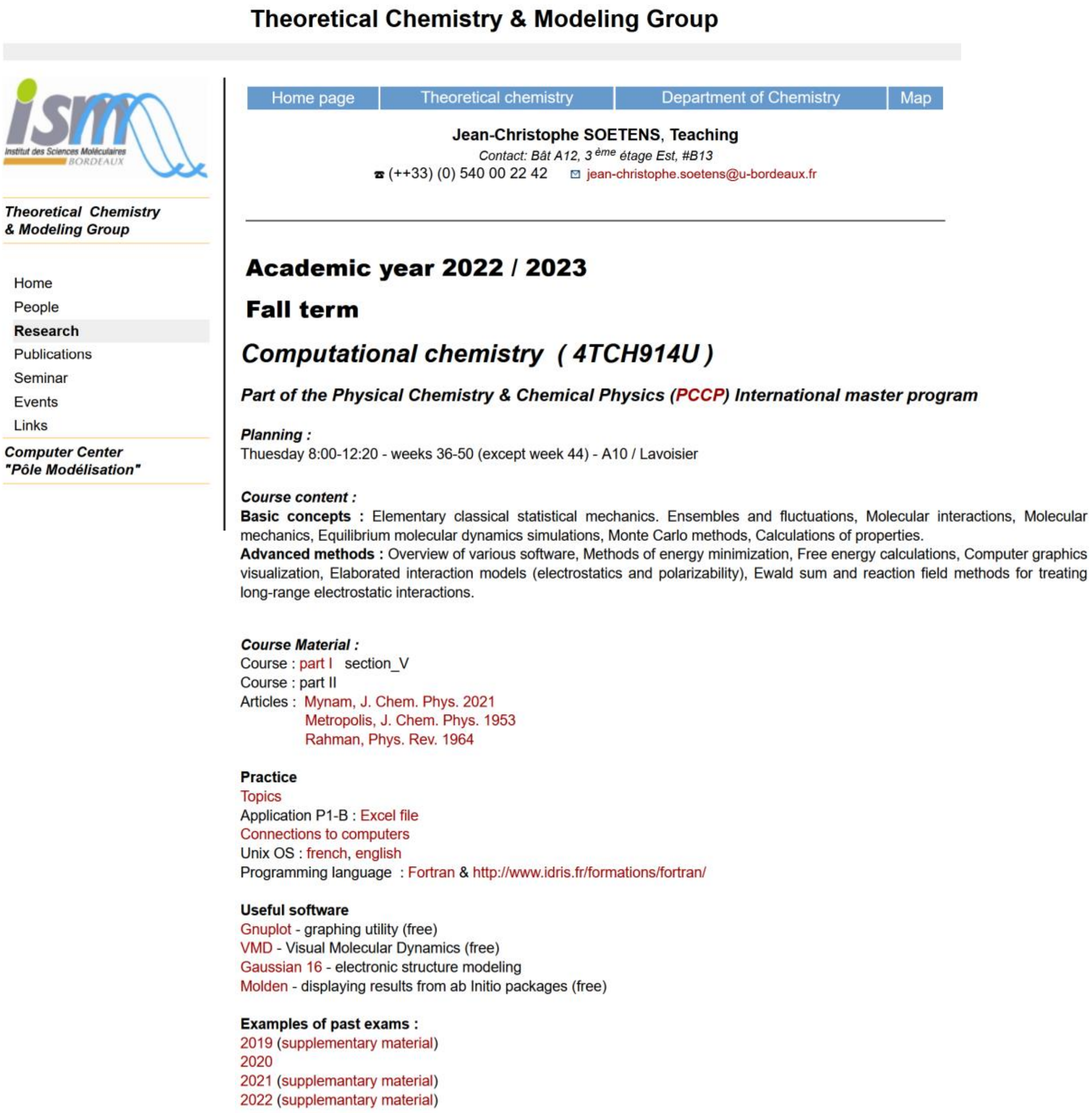
« Personal part »
From scratch

P5-A) Oral presentation
(intermediate x 0.15)

P5-B) Unix, Fortran langage
Gnuplot, VMD, etc.
(intermediate x 0.15)

P5-C) Around November

Course materials : *http://theo.ism.u-bordeaux.fr/J-C.Soetens/teach.html*



Theoretical Chemistry & Modeling Group

Jean-Christophe SOETENS, Teaching

Contact: Bât A12, 3 ème étage Est, #B13

☎ (++33) (0) 540 00 22 42 ✉ jean-christophe.soetens@u-bordeaux.fr

---

Academic year 2022 / 2023

Fall term

Computational chemistry (4TCH914U)

Part of the Physical Chemistry & Chemical Physics (PCCP) International master program

Planning :
Thuesday 8:00-12:20 - weeks 36-50 (except week 44) - A10 / Lavoisier

Course content :
Basic concepts : Elementary classical statistical mechanics. Ensembles and fluctuations, Molecular interactions, Molecular mechanics, Equilibrium molecular dynamics simulations, Monte Carlo methods, Calculations of properties.
Advanced methods : Overview of various software, Methods of energy minimization, Free energy calculations, Computer graphics visualization, Elaborated interaction models (electrostatics and polarizability), Ewald sum and reaction field methods for treating long-range electrostatic interactions.

Course Material :
Course : part I section_V
Course : part II
Articles : Mynam, J. Chem. Phys. 2021
Metropolis, J. Chem. Phys. 1953
Rahman, Phys. Rev. 1964

Practice
Topics
Application P1-B : Excel file
Connections to computers
Unix OS : french, english
Programming language : Fortran & http://www.idris.fr/formations/fortran/

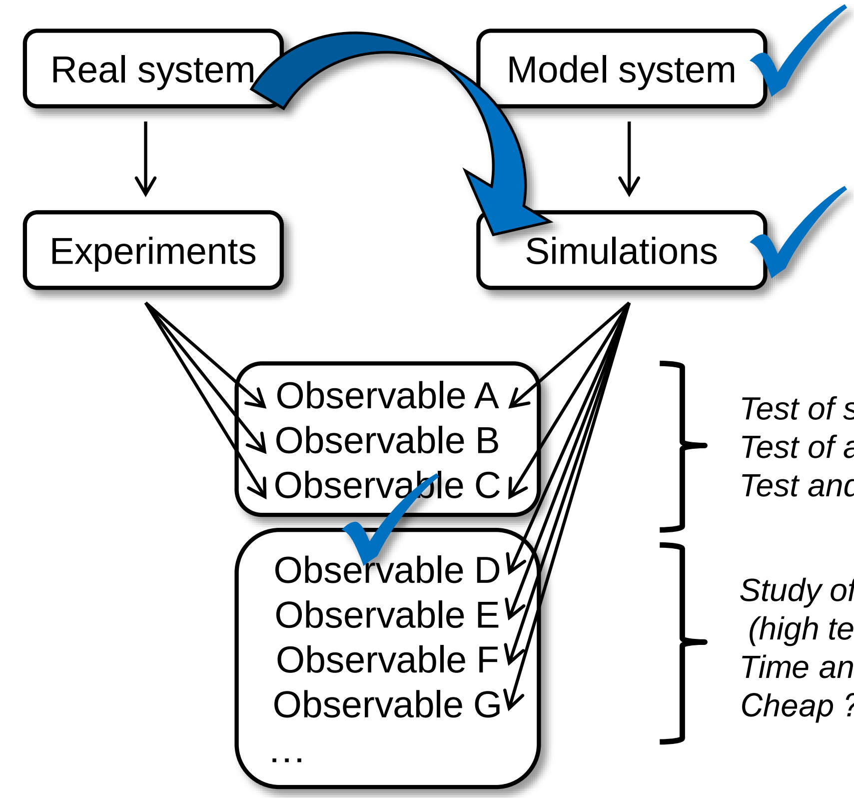
Useful software
Gnuplot - graphing utility (free)
VMD - Visual Molecular Dynamics (free)
Gaussian 16 - electronic structure modeling
Molden - displaying results from ab Initio packages (free)

Examples of past exams :
2019 (supplementary material)
2020
2021 (supplemantary material)
2022 (supplemantary material)

I. Overview of Statistical Mechanics & Molecular Simulations

$$\underbrace{\textbf{Simulations}}_{\textit{simulation approach}} \quad \textbf{of} \quad \underbrace{\textbf{Molecular Systems ?}}_{\textit{specific strategies}}$$

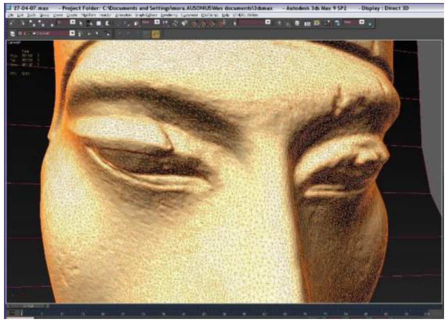
simulation approach ⟺ *specific strategies*

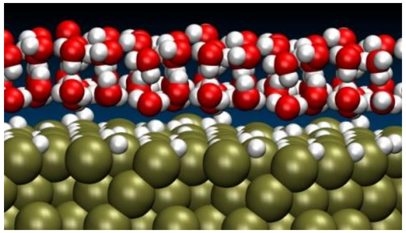


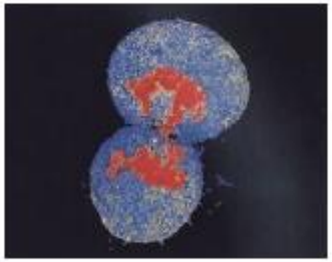
Test of simulation models
Test of approximate theories
Test and/or interpretation of experiments

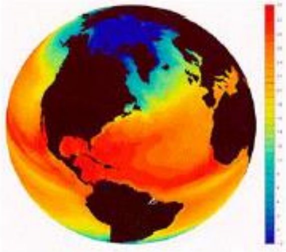
Study of systems at extreme conditions
(high temperatures or pressures)
Time and space resolution inaccessible to experiment
Cheap ? (Moor's law)

Some words about « simulations »

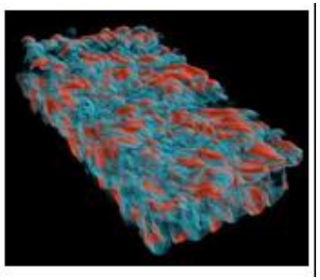


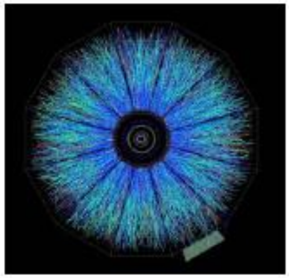


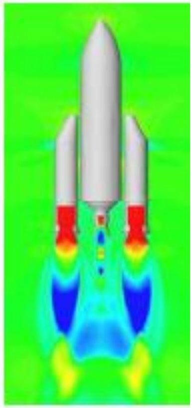


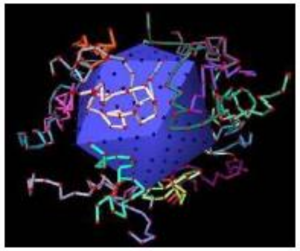


- ***Meteorology, climatology, Earth Sciences***
  Measure warming and climate change
  Anticipate weather events
- ***Astrophysics, particle physics and plasma physics***
  Understand the evolution of galaxies
  Explore quantum theories (QCD, ITER)
- ***Materials science, chemistry, nanoscience***
  Development of smart materials
  Better understanding of the properties
- ***Life Sciences, human sciences***
  Exploring biological systems, brain
  Develop targeted drug molecules
- ***Engineering***
- ***Etc…***

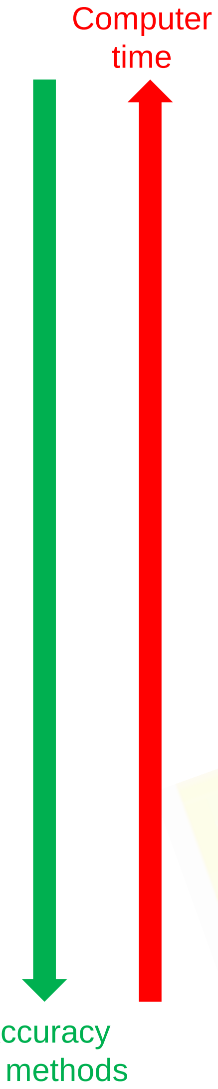




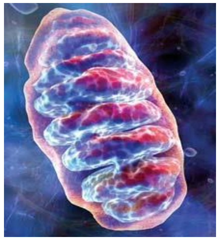




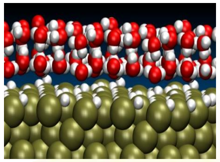
Time and length scales



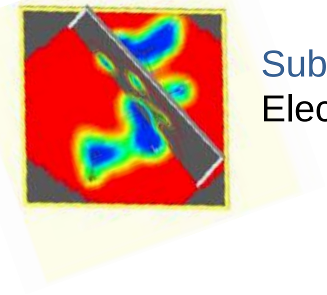
Macroscale
Time > s
Lengthscale > 1 µm



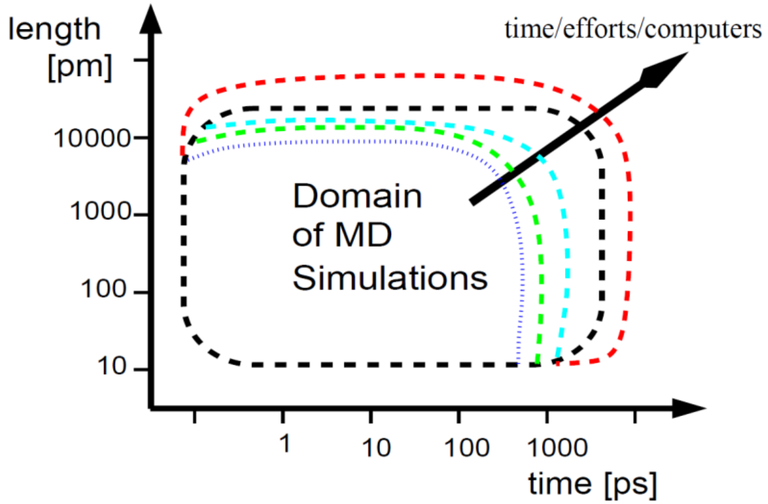
Mesoscale
Time : ns – ms
Lenghtscale :10-10000 nm



Nanoscale
Time : fs – ns
Lengthscale : 0.1 – 10 nm



Subatomic scale
Electronic structure



Phase space
Statistics
Free energy, entropy
Molecular Dynamics, Monte Carlo

ab initio
Electrons
Potentiel Energy Surface

The purpose of this course : ***« Statistical Mechanics & Simulations »***

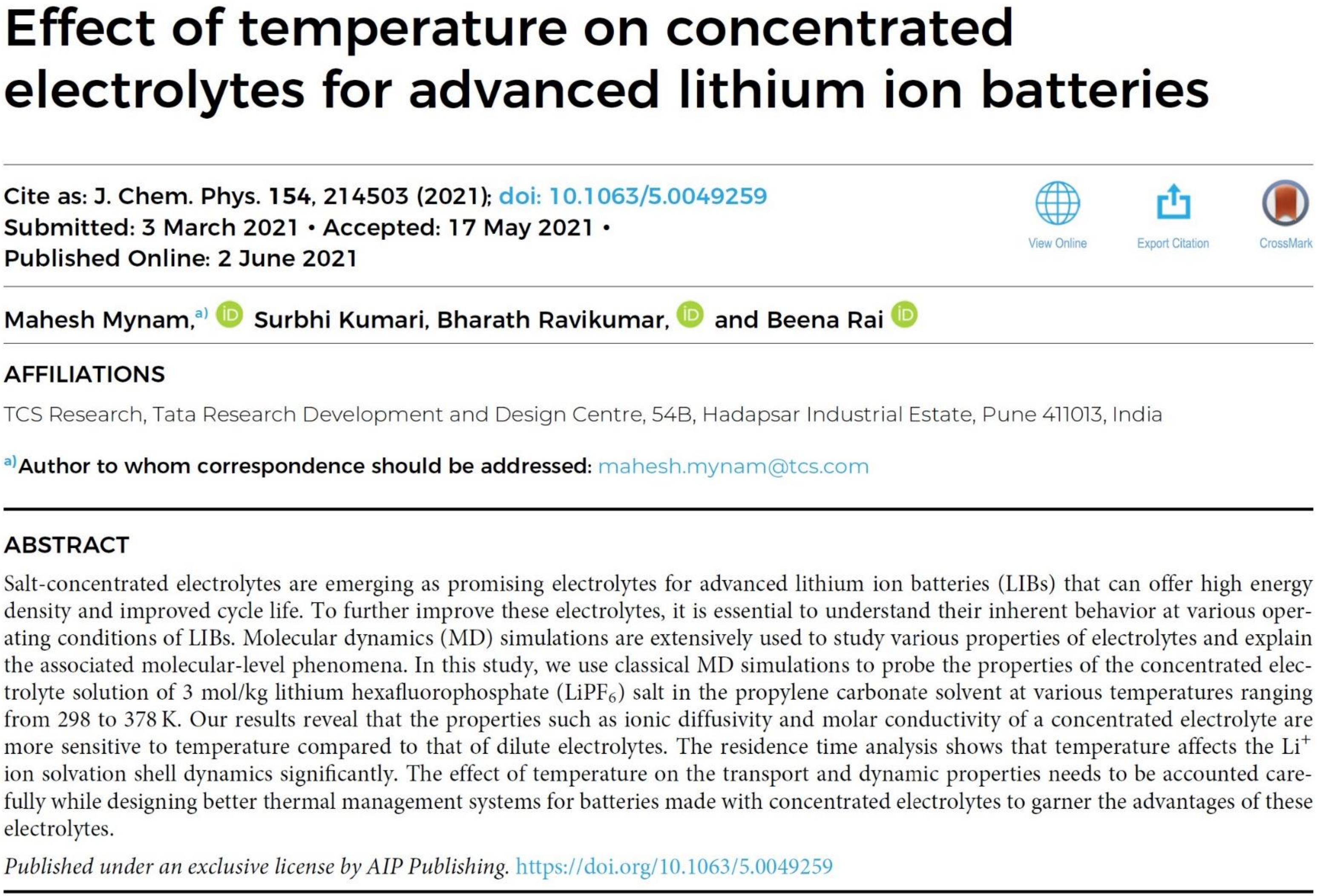
- To become familiar with the basics concepts underlying molecular simulations
- To understand the possibilities and limitations of (present day) simulations
- To understand the machinary of the two main methods :

  Molecular Dynamics simulations

  Monte Carlo simulations

- **To be able to understand a scientific article dealing with molecular simulation and to discuss with experts in the field**
- To give you some ideas on how to start your own studies
- To become familiar with a numerical scientific environment
  (software, programming, graphics, etc)

Exercise *: let's look at a typical molecular simulation article in physical chemistry to see what we need to learn before understanding such an article !?*



Effect of temperature on concentrated electrolytes for advanced lithium ion batteries

Cite as: J. Chem. Phys. **154**, 214503 (2021); doi: 10.1063/5.0049259
Submitted: 3 March 2021 • Accepted: 17 May 2021 •
Published Online: 2 June 2021

Mahesh Mynam,[a] Surbhi Kumari, Bharath Ravikumar, and Beena Rai

AFFILIATIONS

TCS Research, Tata Research Development and Design Centre, 54B, Hadapsar Industrial Estate, Pune 411013, India

[a] **Author to whom correspondence should be addressed:** mahesh.mynam@tcs.com

ABSTRACT

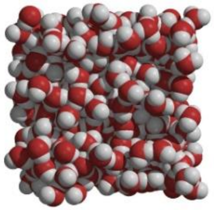
Salt-concentrated electrolytes are emerging as promising electrolytes for advanced lithium ion batteries (LIBs) that can offer high energy density and improved cycle life. To further improve these electrolytes, it is essential to understand their inherent behavior at various operating conditions of LIBs. Molecular dynamics (MD) simulations are extensively used to study various properties of electrolytes and explain the associated molecular-level phenomena. In this study, we use classical MD simulations to probe the properties of the concentrated electrolyte solution of 3 mol/kg lithium hexafluorophosphate ($LiPF_6$) salt in the propylene carbonate solvent at various temperatures ranging from 298 to 378 K. Our results reveal that the properties such as ionic diffusivity and molar conductivity of a concentrated electrolyte are more sensitive to temperature compared to that of dilute electrolytes. The residence time analysis shows that temperature affects the Li^+ ion solvation shell dynamics significantly. The effect of temperature on the transport and dynamic properties needs to be accounted carefully while designing better thermal management systems for batteries made with concentrated electrolytes to garner the advantages of these electrolytes.

« Simulations » in chemical physics

Macroscopic point of view

Real system
N_A *molecules or particles*
described by few variables
such as N, V, p, T…

Microscopic point of view



Model system
classical → microstates
$\mathcal{P}_i(\vec{r}_i, \vec{v}_i\ )$

One <u>Macro</u>scopic state ⬄ Many compatible <u>micro</u>scopic states

⇒ for a classical system of N particles at time t : $\mathcal{P}_i(\vec{r}_{i1}, \vec{r}_{i2}, \vec{r}_{i3}, \ldots, \vec{v}_{i1}, \vec{v}_{i2}, \vec{v}_{i3}, \ldots\ )$

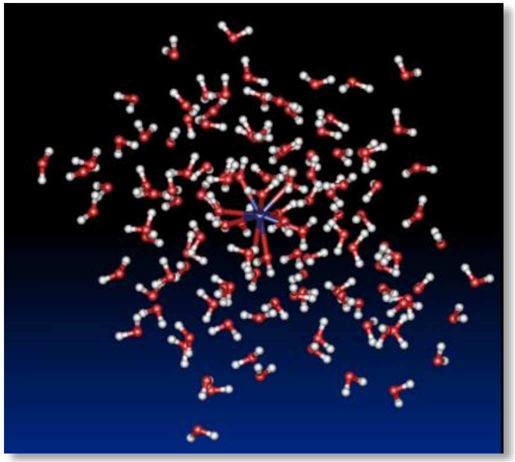
⇒ to compute the thermodynamic average of a quantity $\mathcal{A}$

$$< \mathcal{A} >= \frac{1}{Z} \int\int \mathcal{A}(\vec{r}_i, \vec{v}_i)\mathcal{W}(\vec{r}_i, \vec{v}_i)drdp \quad \text{with} \quad Z = \int\int \mathcal{W}(\vec{r}_i, \vec{v}_i)drdp$$

we thus have to know for each point of the phase space :

$\mathcal{A}(\vec{r}_i, \vec{v}_i)$ value of the property in the microstate $\mathcal{P}_i(\vec{r}_i, \vec{v}_i\ )$
$\mathcal{W}(\vec{r}_i, \vec{v}_i)$ weight (probability) of the microstate

Microscopic point of view : how many compatible microscopic states ?



Application:
system of 36 water molecules ↔ 108 atoms
7 spatial coordinates considered for each atom :
- original position (x_o, y_o, z_o)
- 2 moves in each of the 3 directions $(x_o \pm \delta, y_o \pm \delta, z_o \pm \delta)$

Work to do :
remember : $< \mathcal{A} >= \frac{1}{Z} \int \int \mathcal{A}(\vec{r}_i, \vec{v}_i) \mathcal{W}(\vec{r}_i, \vec{v}_i) drdp$

⇒ Generate all corresponding configurations C_i
⇒ Compute each $A(C_i)$ and $W(C_i)$
(potential energy function for example, 1 ns on a laptop)
⇒ Result : <A>

Question : computer time necessary to do this work ?

Answer :
⇒ Generate all corresponding configurations C_i → Nb of configurations = ?

⇒ Compute each $A(C_i)$ and $W(C_i)$ → Total computer time = ?
⇒ Result : <A>

~~Laptop~~ Supercomputer (1.5 million processors) → ?

Quizz : 1 day on this supercomputer ↔ how many H_2O ? → ? molecules

Summary up to there

➢We cannot use, for practical reasons, the formulae that we find in statistical mechanics textbooks in a straightforward manner to compute averages.

➢What we need / what we can do :

1- to construct a list of microstates compatible with the macrostate we are interesting in.

2- to compute the properties for each element of the list.

3- to take the averages over the ensemble of elements.

➢Warning : this list will always be finite !

Take home message…

Real condensed phase ?

- phase space ⇔ $\mathcal{P}_i(\vec{r}_i, \vec{v}_i\ )$
- large number of molecules
- complex potential energy functions

⇒

“Brute force” approach is infeasible
and
Extremely inefficient

we know only at the end which microstates contribute to the integral

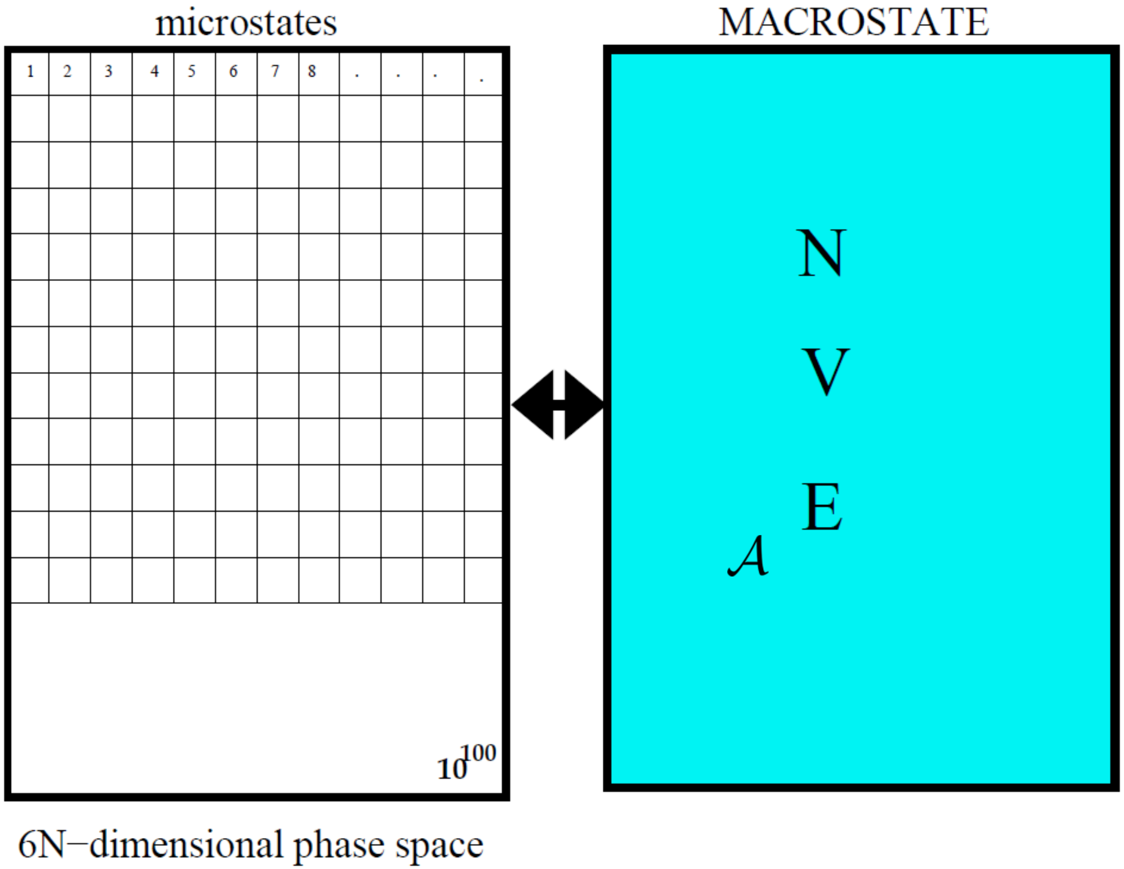
How to generate the microstates efficiently ?

SOLUTION = SIMULATION

Simulations = *Methods to explore the dominant regions of the phase space*

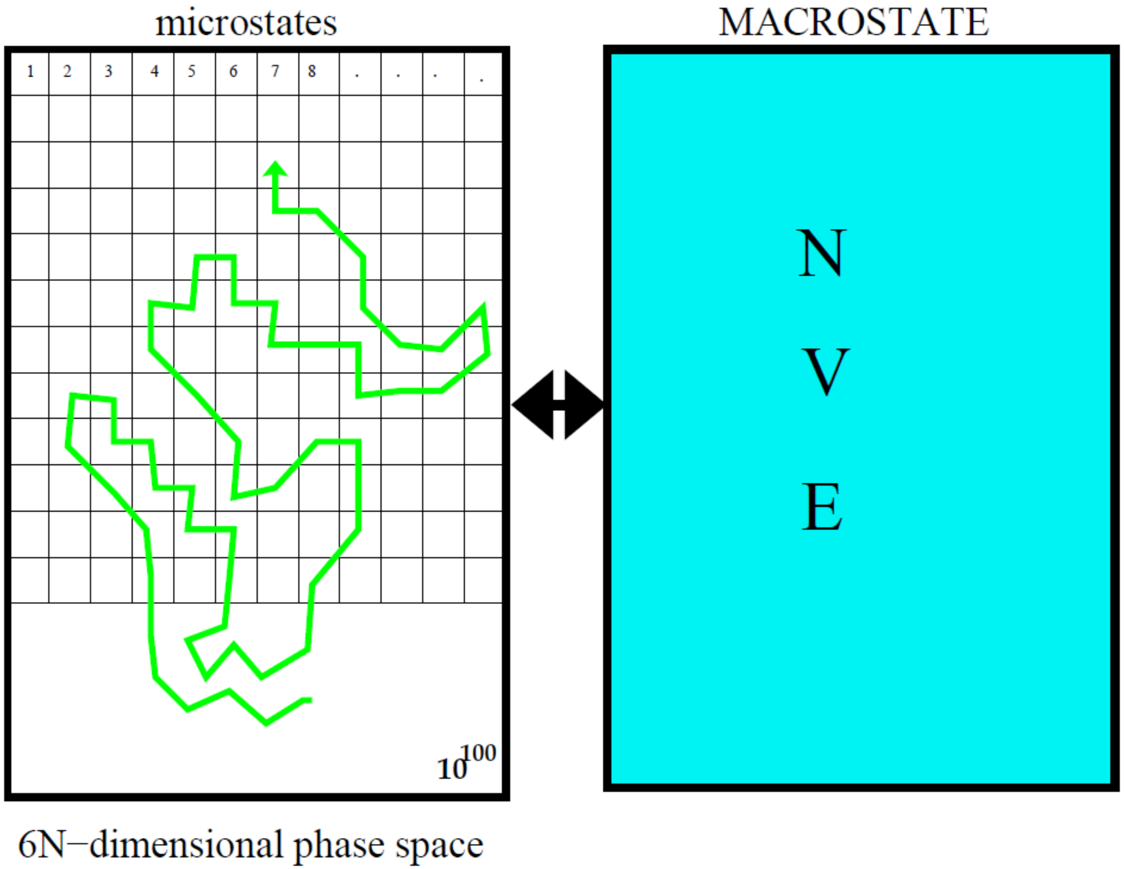
= *Methods to generate the dominant contributions to the integral*

Phase space :

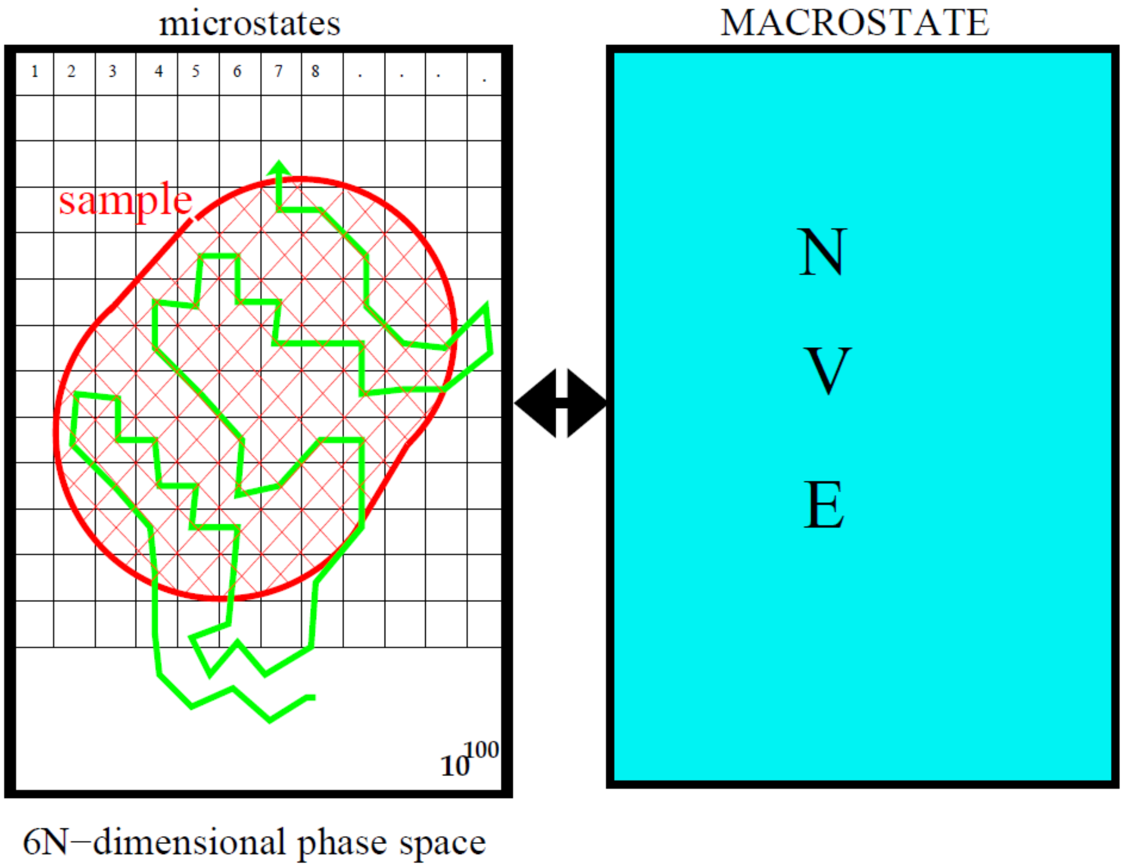


Classical system of N particles at time t : $\mathcal{P}_i\left(\vec{r}_{i1}, \vec{r}_{i2}, \vec{r}_{i3}, \ldots, \vec{v}_{i1}, \vec{v}_{i2}, \vec{v}_{i3}, \ldots\right)$

Exploration of the dominant regions of the phase space :

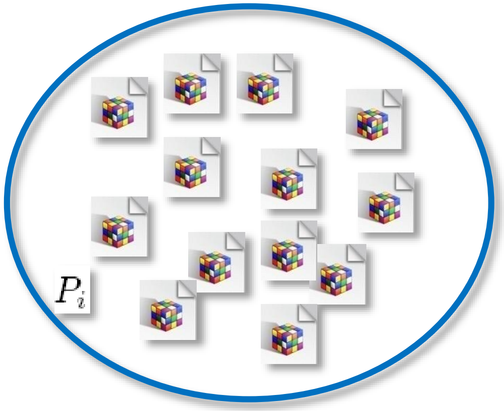


Finite sample of microstates:



Simulation approach :

⇒ Construct a finite list of microscopic states which are representative and compatible with some set of macroscopic variables (T, P, ρ…)



Ensemble of N microstates $\mathcal{P}_i$

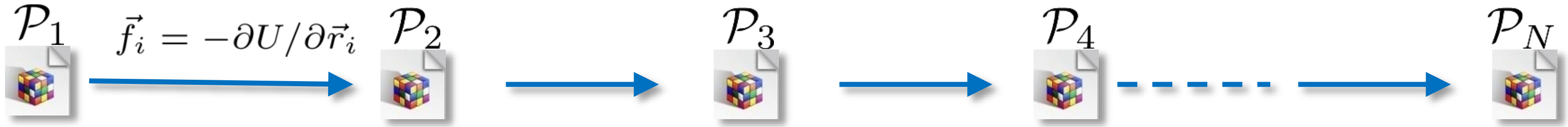
$$< \mathcal{A}(\mathcal{P}_i) > = \frac{1}{N}\sum_{i=1}^{N} \mathcal{A}(\mathcal{P}_i)$$

Representative ?

$$< \mathcal{A}(\mathcal{P}_i) >_N \approx < \mathcal{A}(\mathcal{P}_i) >_{N\to\infty}$$

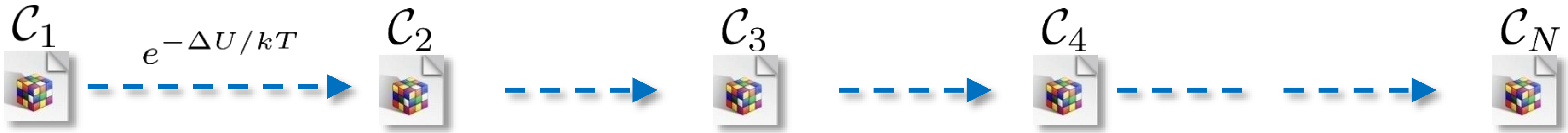
Molecular Dynamics

⇒ list of P_i in Phase space (**r**,**v**) and ordered in time



Monte Carlo

⇒ list of C_i in Configuration space (**r**)



Molecular Dynamics : MD

Generates Ps characterized by :

- *Constant number of particules, N*
- *Constant volume, V*
- *Constant total energy, E*

This ensemble is known as the NVE ensemble
*(**microcanonic ensemble** of statistical mechanics)*

Added values

MD generates the Ps in a meaningful order
The order parameter is the time, t

Consequences

MD generates a trajectory

MD allow to compute dynamical properties

Monte Carlo : MC

Generates only the position part of the Ps (Cs) characterized by :

- *Constant number of particules, N*
- *Constant volume, V*
- *Constant temperature, T*

This ensemble is known as the NVT ensemble
*(**canonic ensemble** of statistical mechanics)*

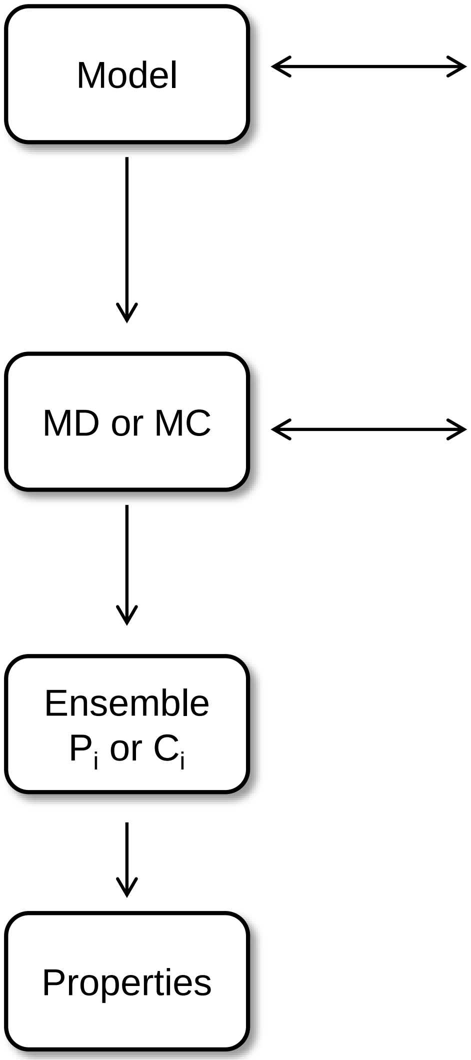
MC generates the Cs in a random order

Consequences

No trajectories, just a package of Cs

Dynamical properties inaccessible

The machinery :



Model

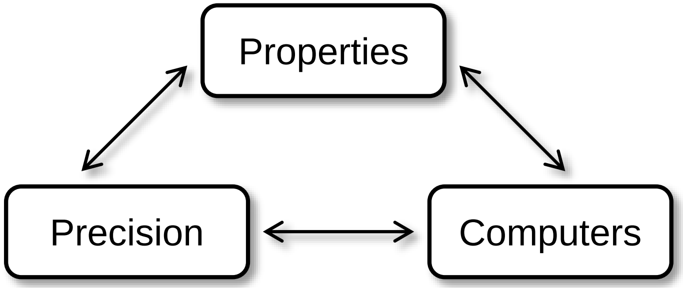
MD or MC

Ensemble
P_i or C_i

Properties

Definition of the system
Thermodynamics conditions
Interaction model

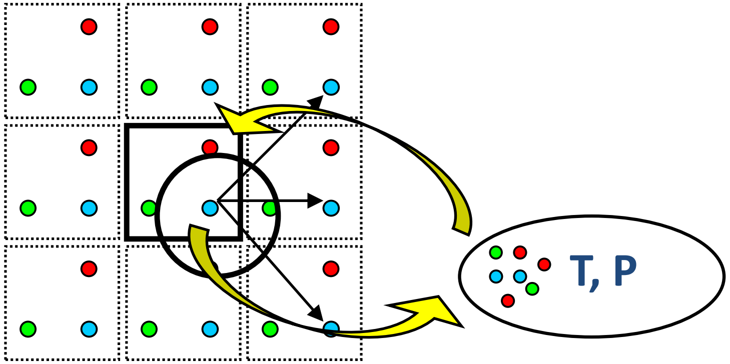
Many technical details :
Simulation box
Periodic boundary conditions
Minimim image convention
Long range interactions
Thermodynamics ensemble
etc.

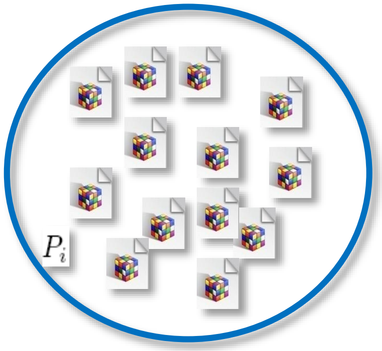


Properties

Precision

Computers





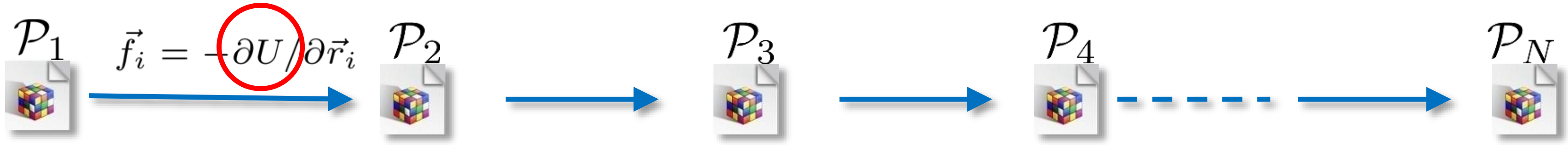
The model is the crucial input of a simulation

« Model » is a complex concept… it includes the technical details of the simulation and the **interaction potential**.

But, when everything is done, every success or failure of the simulation is ascribable to a success or a failure of the model.

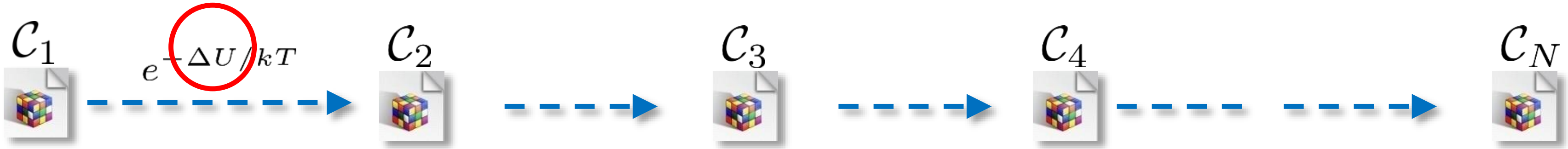
Molecular Dynamics

⇒ list of P_i in Phase space (**r**,**v**) and ordered in time



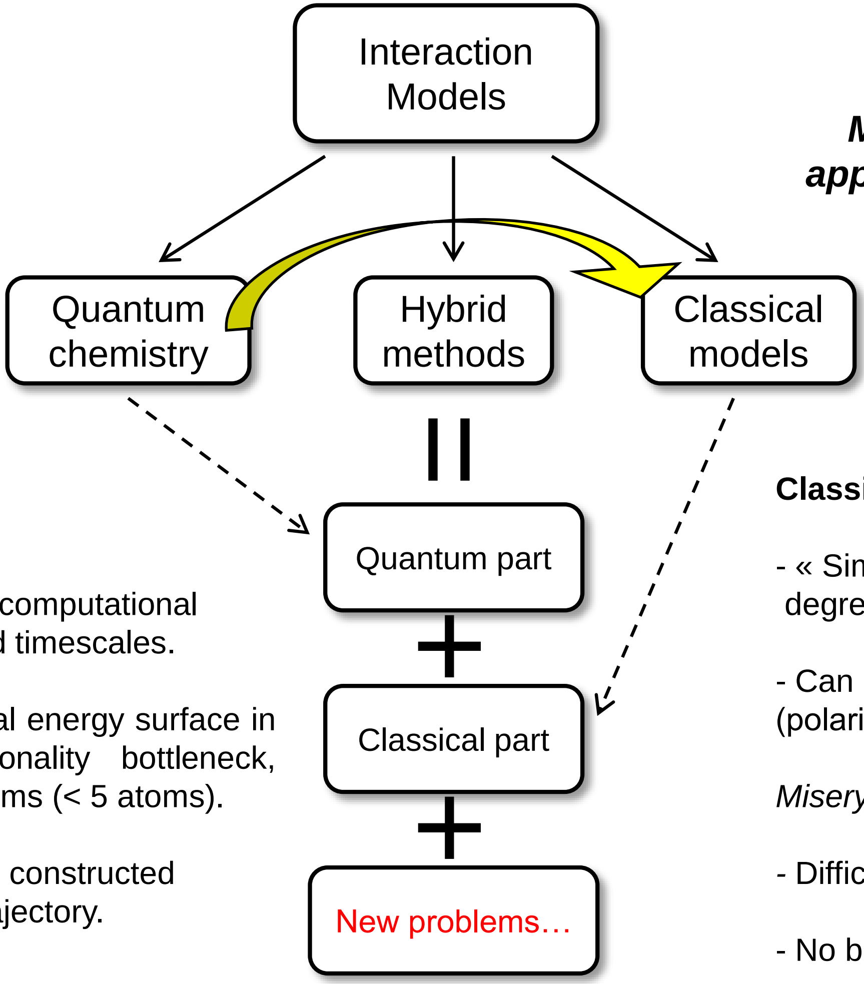
Monte Carlo

⇒ list of C_i in Configuration space (**r**)



I. Overview of Statistical Mechanics & Molecular Simulations

Necessary condition to perform simulation : ***Interaction potential***



Interaction Models

Quantum chemistry

Hybrid methods

Classical models

Quantum part

Classical part

New problems…

Many simplifications and approximations are necessary

Potentials and forces from quantum chemistry

- Price of dramatically increased computational costs: much smaller systems and timescales.

- Constructing the whole potential energy surface in advance: exponential dimensionality bottleneck, possibly only for very small systems (< 5 atoms).

- Alternative: on-the-fly potentials constructed along the molecular dynamics trajectory.

Programs for AIMD :
CPMD, CP2K, VASP, NWChem, CASTEP, CP-PAW, fhi98md,…

Classical models

- « Simple » and cheap : large systems, many degrees of freedom, large timescales.

- Can be refined by including additional terms (polarization, cross intramolecular terms, …).

Misery of empirical force fields:

- Difficult to choose an available model.

- No bond making/breaking – no chemistry !

- Difficult to improve in a systematic way.

- The development of original models is a science in itself.

Classical models

A multidimensional function $\hat{V}(\{R_I^{nuclei}\})$ is still too complex.

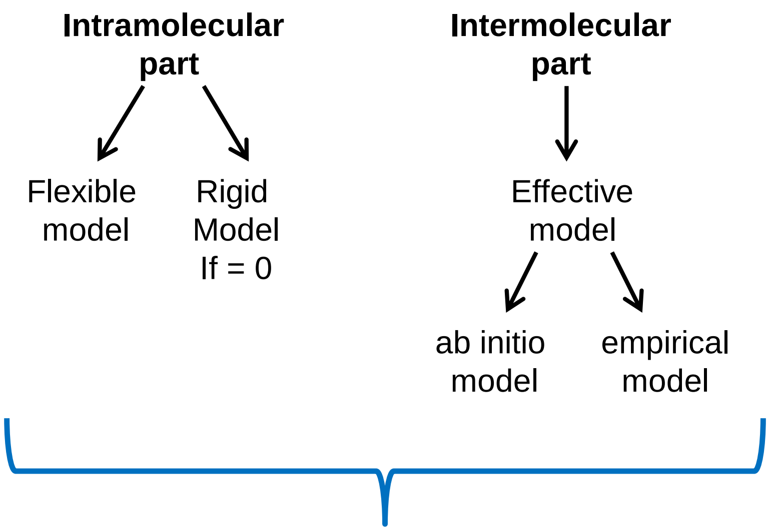
The pair potential approximation consists to developp this function over monomers and pair of molecules :

Beyond the pair potential approximation ?

$$\mathcal{V}(\mathbf{r}) = \sum_{i=1}^{N} \mathcal{V}(\mathbf{r}_i) + \sum_{i=1}^{N-1} \sum_{j>i} \mathcal{V}(\mathbf{r}_{ij})$$

$$+ \sum_{i=1}^{N-1} \sum_{j>i} \sum_{k>j} \mathcal{V}(\mathbf{r}_{ijk}) + \dots$$

Three body term



Intramolecular part

Flexible model

Rigid Model
If = 0

Intermolecular part

Effective model

ab initio model

empirical model

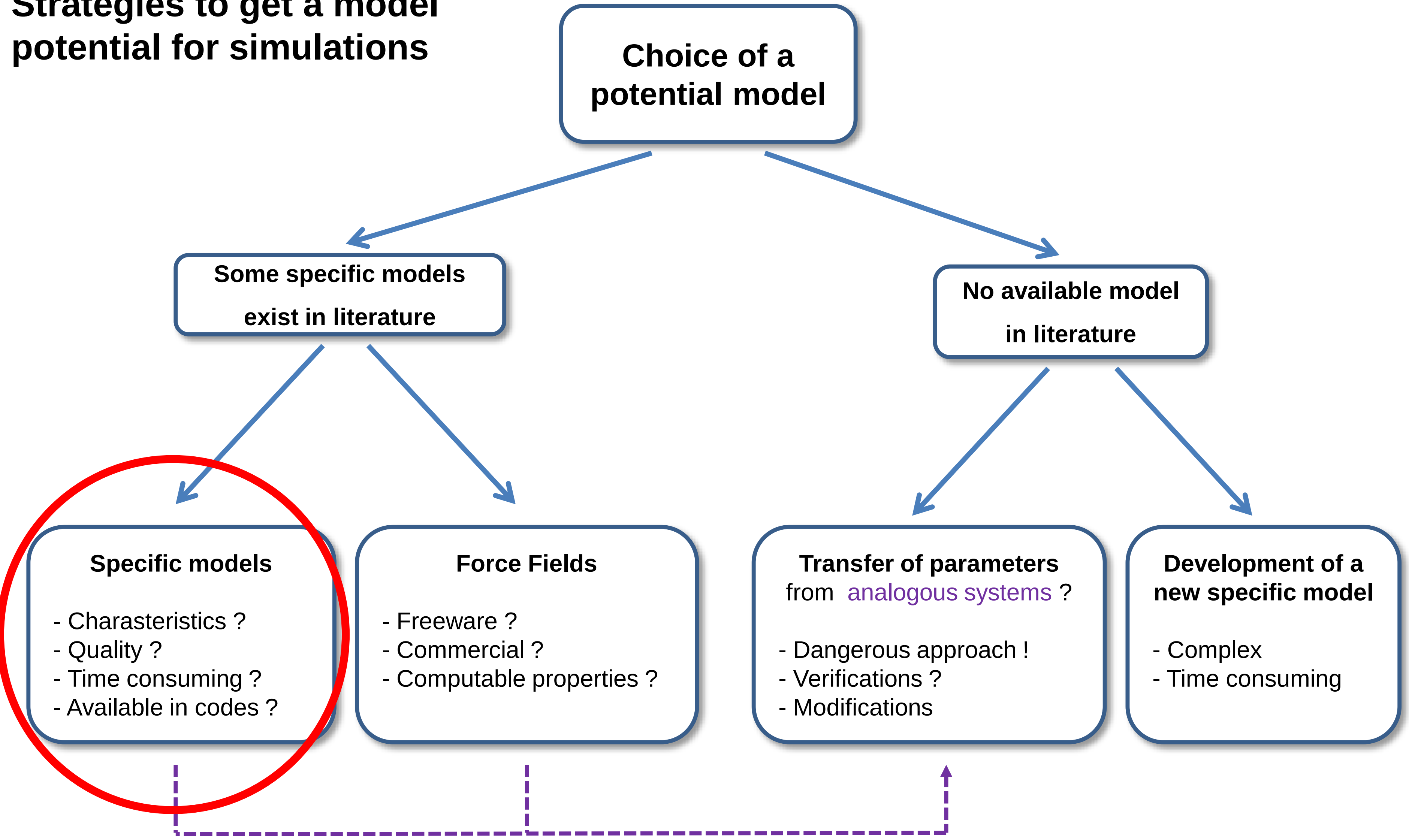
and / or

$$+ \mathcal{V}_{induction}$$

N - body term
Explicit polarizabilities

The pair potential approximation is a very drastic ssumption

Strategies to get a model potential for simulations



Choice of a potential model

Some specific models exist in literature

No available model in literature

Specific models

- Charasteristics ?
- Quality ?
- Time consuming ?
- Available in codes ?

Force Fields

- Freeware ?
- Commercial ?
- Computable properties ?

Transfer of parameters from analogous systems ?

- Dangerous approach !
- Verifications ?
- Modifications

Development of a new specific model

- Complex
- Time consuming

Example of liquid WATER

Just for water (H2O), there are probably about 50 reasonably good, but quite different models in the literature. They are known by acronyms like BNS, ST2, MCY, TIPS, TIP3P, TIP4P, CF2, CF3, BJH, SPC, SPC/E, etc. etc. etc

Why so many?

Are you lucky ?

Rigid models

Flexible models

Molecular dipoles : changing or not (flexible, polarizable)

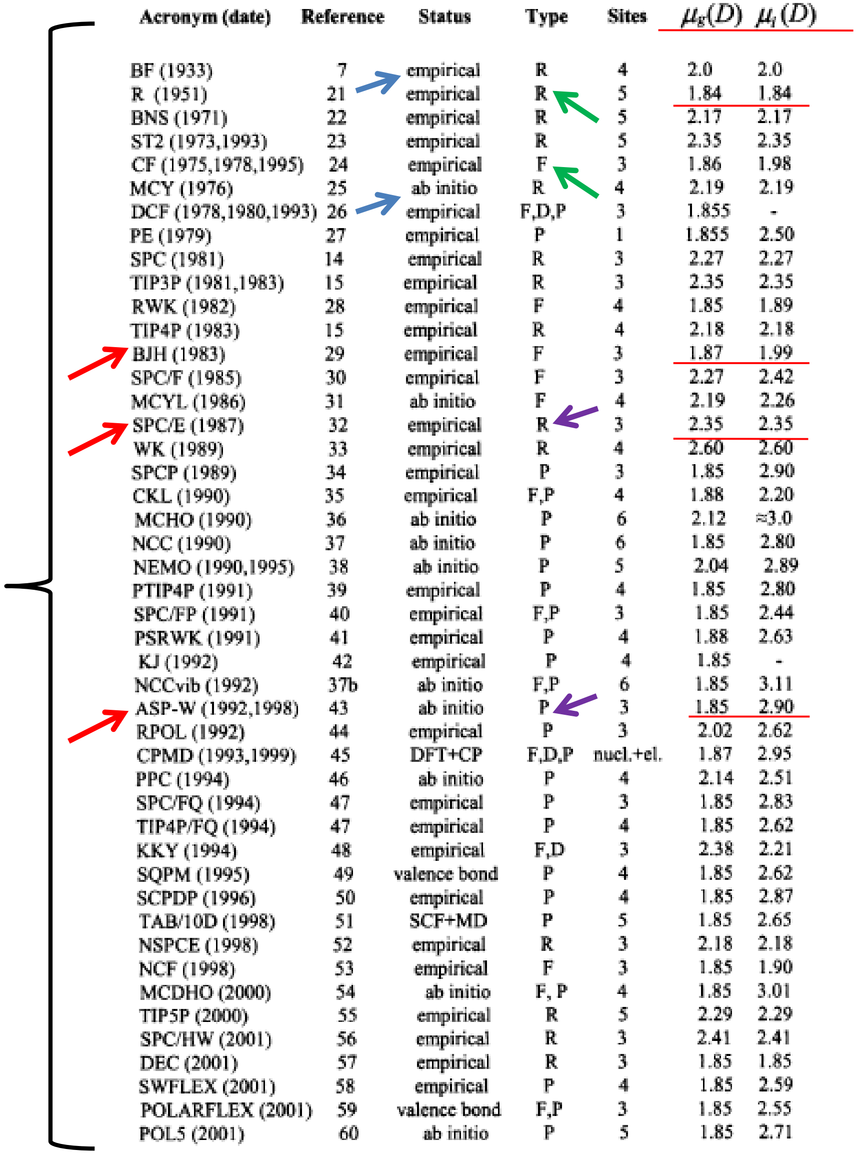
Polarizable models

Effective models

Empirical models : tuned to get certain properties right

ab initio models : fitted from quantum chemistry

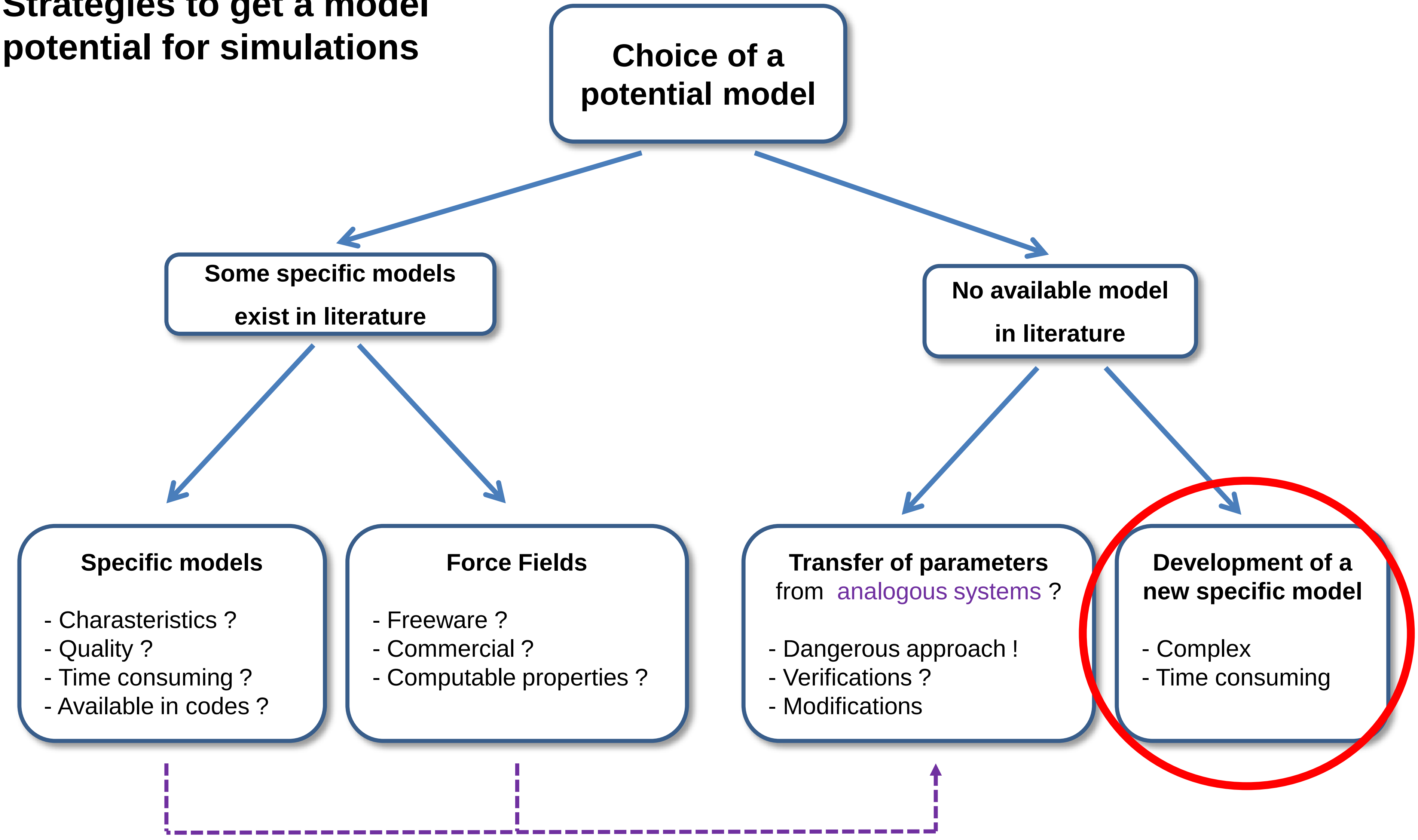
H_2O



Acronym (date)	Reference	Status	Type	Sites	$\mu_g(D)$	$\mu_l(D)$
BF (1933)	7	empirical	R	4	2.0	2.0
R (1951)	21	empirical	R	5	1.84	1.84
BNS (1971)	22	empirical	R	5	2.17	2.17
ST2 (1973,1993)	23	empirical	R	5	2.35	2.35
CF (1975,1978,1995)	24	empirical	F	3	1.86	1.98
MCY (1976)	25	ab initio	R	4	2.19	2.19
DCF (1978,1980,1993)	26	empirical	F,D,P	3	1.855	-
PE (1979)	27	empirical	P	1	1.855	2.50
SPC (1981)	14	empirical	R	3	2.27	2.27
TIP3P (1981,1983)	15	empirical	R	3	2.35	2.35
RWK (1982)	28	empirical	F	4	1.85	1.89
TIP4P (1983)	15	empirical	R	4	2.18	2.18
BJH (1983)	29	empirical	F	3	1.87	1.99
SPC/F (1985)	30	empirical	F	3	2.27	2.42
MCYL (1986)	31	ab initio	F	4	2.19	2.26
SPC/E (1987)	32	empirical	R	3	2.35	2.35
WK (1989)	33	empirical	R	4	2.60	2.60
SPCP (1989)	34	empirical	P	3	1.85	2.90
CKL (1990)	35	empirical	F,P	4	1.88	2.20
MCHO (1990)	36	ab initio	P	6	2.12	≈3.0
NCC (1990)	37	ab initio	P	6	1.85	2.80
NEMO (1990,1995)	38	ab initio	P	5	2.04	2.89
PTIP4P (1991)	39	empirical	P	4	1.85	2.80
SPC/FP (1991)	40	empirical	F,P	3	1.85	2.44
PSRWK (1991)	41	empirical	P	4	1.88	2.63
KJ (1992)	42	empirical	P	4	1.85	-
NCCvib (1992)	37b	ab initio	F,P	6	1.85	3.11
ASP-W (1992,1998)	43	ab initio	P	3	1.85	2.90
RPOL (1992)	44	empirical	P	3	2.02	2.62
CPMD (1993,1999)	45	DFT+CP	F,D,P	nucl.+el.	1.87	2.95
PPC (1994)	46	ab initio	P	4	2.14	2.51
SPC/FQ (1994)	47	empirical	P	3	1.85	2.83
TIP4P/FQ (1994)	47	empirical	P	4	1.85	2.62
KKY (1994)	48	empirical	F,D	3	2.38	2.21
SQPM (1995)	49	valence bond	P	4	1.85	2.62
SCPDP (1996)	50	empirical	P	4	1.85	2.87
TAB/10D (1998)	51	SCF+MD	P	5	1.85	2.65
NSPCE (1998)	52	empirical	R	3	2.18	2.18
NCF (1998)	53	empirical	F	3	1.85	1.90
MCDHO (2000)	54	ab initio	F, P	4	1.85	3.01
TIP5P (2000)	55	empirical	R	5	2.29	2.29
SPC/HW (2001)	56	empirical	R	3	2.41	2.41
DEC (2001)	57	empirical	R	3	1.85	1.85
SWFLEX (2001)	58	empirical	P	4	1.85	2.59
POLARFLEX (2001)	59	valence bond	F,P	3	1.85	2.55
POL5 (2001)	60	ab initio	P	5	1.85	2.71

B. Guillot, Journal of Molecular Liquids ***101,*** *219 (2002)*

Strategies to get a model potential for simulations



Choice of a potential model

Some specific models exist in literature

No available model in literature

Specific models

- Charasteristics ?
- Quality ?
- Time consuming ?
- Available in codes ?

Force Fields

- Freeware ?
- Commercial ?
- Computable properties ?

Transfer of parameters from analogous systems ?

- Dangerous approach !
- Verifications ?
- Modifications

Development of a new specific model

- Complex
- Time consuming

You want to study a system by the use of molecular simulations but :

- *you cannot find a (good) model potential.*
- *you are not confident in the use of parameters coming from analogous systems.*
- *you are interested in the challenging task that consists in the development of an original model.*

Whatever the number of reasons, the question is : how to make a model ?

1. select sites, decide if you need a rigid or a flexible model.

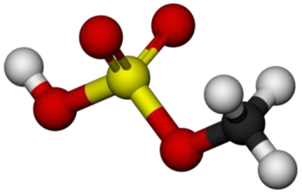
2. be as smart as you can and invent a good (whatever that means) functional form for V with free parameters.

3. determine the free parameters by fitting the V-functions.

   - *do quantum mechanical calculations on a small number of molecules (2, maybe 3 or 4) for as many relative geometries as possible.*
   - *empirically using simulations*

4. test, improve, work hard (sometimes back to step 1 or 2 or 3…), ..., try & error procedures…

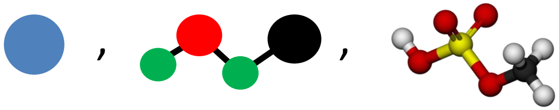
Conclusion: *the development of models is a science in itself.*



Development of a new potential model

Nature of the model

All atoms, unified atoms,
Flexible or rigid model



The development of models is a science in itself !

Mathematical expression for the potential energy function

"Use the right level of description to catch the phenomena of interest…"

Empirical model

Fit the parameters through simulations (MD or MC) of some experimental properties

Ab initiol model

Fit the parameters through QM calculations

- Geometry
- Frequencies
- PES of dimers, trimers…

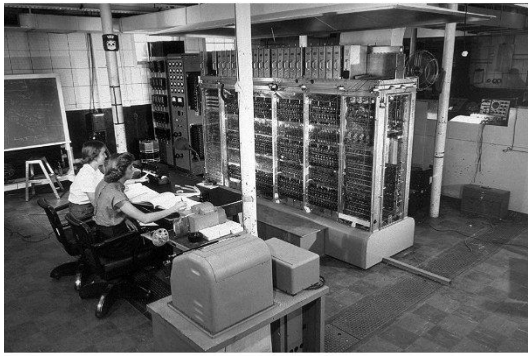
Electrostatic model

-Model derived from electrostatic properties of single molecule.

Fit of the rest (non electrostatic) of the potential energy

Historical perspectives on molecular simulations…

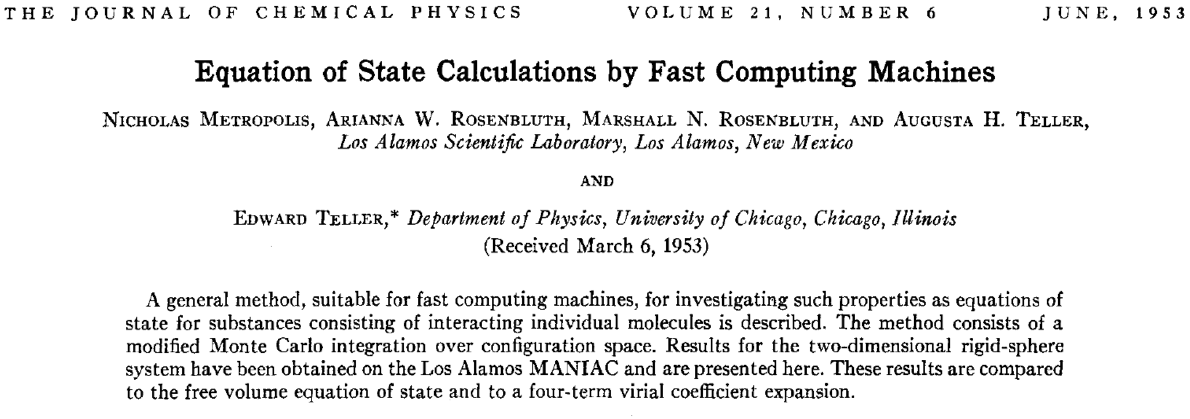
60 years ago : first Monte Carlo simulation (1953)



MANIAC
Mathematical Analyzer,
Numerical Integrator,
and Computer

Mars 1952
Los Alamos

1 KOPS



THE JOURNAL OF CHEMICAL PHYSICS VOLUME 21, NUMBER 6 JUNE, 1953

Equation of State Calculations by Fast Computing Machines

NICHOLAS METROPOLIS, ARIANNA W. ROSENBLUTH, MARSHALL N. ROSENBLUTH, AND AUGUSTA H. TELLER,
Los Alamos Scientific Laboratory, Los Alamos, New Mexico

AND

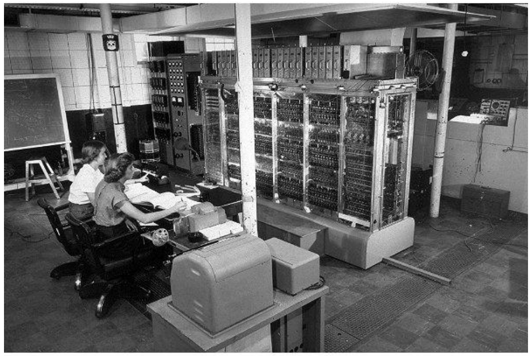
EDWARD TELLER,* *Department of Physics, University of Chicago, Chicago, Illinois*

(Received March 6, 1953)

A general method, suitable for fast computing machines, for investigating such properties as equations of state for substances consisting of interacting individual molecules is described. The method consists of a modified Monte Carlo integration over configuration space. Results for the two-dimensional rigid-sphere system have been obtained on the Los Alamos MANIAC and are presented here. These results are compared to the free volume equation of state and to a four-term virial coefficient expansion.

Historical perspectives on molecular simulations…

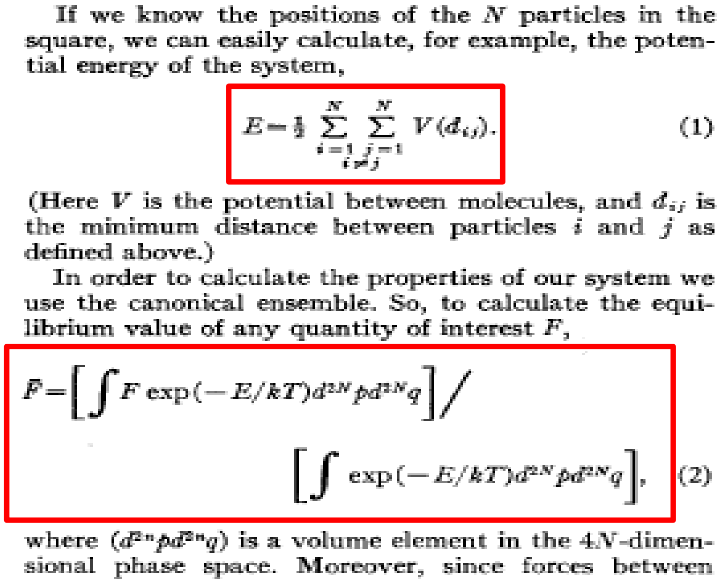
First MC simulation (1953)



MANIAC
Mathematical Analyzer,
Numerical Integrator,
and Computer

Mars 1952
Los Alamos

1 KOPS



If we know the positions of the N particles in the square, we can easily calculate, for example, the potential energy of the system,

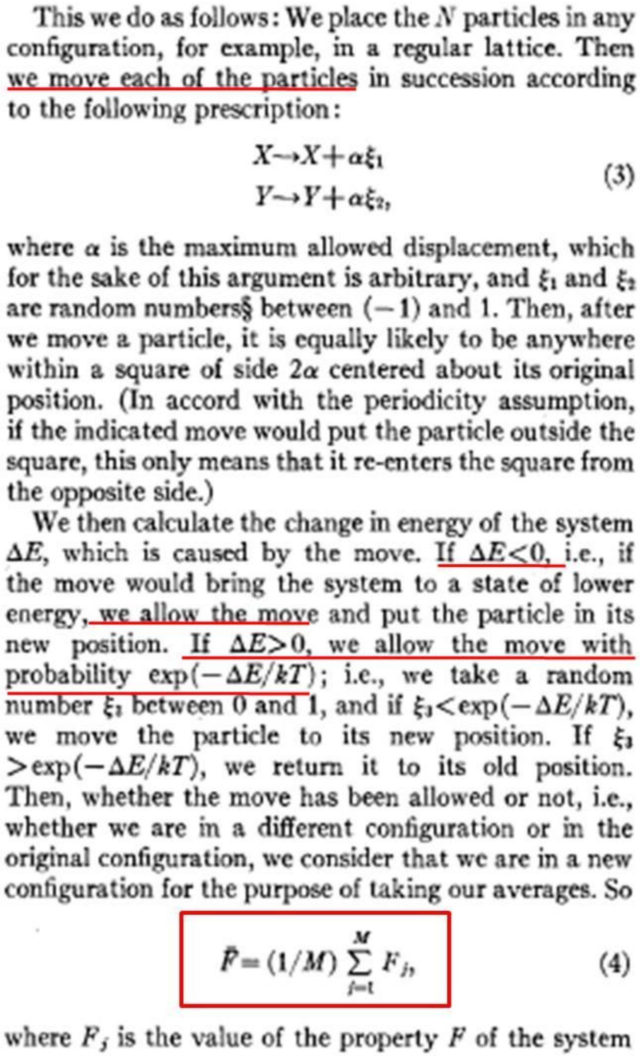
$$E=\tfrac{1}{2}\sum_{i=1}^{N}\sum_{\substack{j=1\\ i\neq j}}^{N} V(d_{ij}). \tag{1}$$

(Here V is the potential between molecules, and d_{ij} is the minimum distance between particles i and j as defined above.)

In order to calculate the properties of our system we use the canonical ensemble. So, to calculate the equilibrium value of any quantity of interest F,

$$\bar{F}=\left[\int F\exp(-E/kT)d^{2N}pd^{2N}q\right]\Big/\left[\int \exp(-E/kT)d^{2N}pd^{2N}q\right], \tag{2}$$

where $(d^{2n}pd^{2n}q)$ is a volume element in the $4N$-dimensional phase space. Moreover, since forces between



This we do as follows: We place the N particles in any configuration, for example, in a regular lattice. Then we move each of the particles in succession according to the following prescription:

$$\begin{aligned} X&\rightarrow X+\alpha\xi_1\\ Y&\rightarrow Y+\alpha\xi_2, \end{aligned} \tag{3}$$

where α is the maximum allowed displacement, which for the sake of this argument is arbitrary, and ξ_1 and ξ_2 are random numbers§ between (-1) and 1. Then, after we move a particle, it is equally likely to be anywhere within a square of side 2α centered about its original position. (In accord with the periodicity assumption, if the indicated move would put the particle outside the square, this only means that it re-enters the square from the opposite side.)

We then calculate the change in energy of the system ΔE, which is caused by the move. If $\Delta E<0$, i.e., if the move would bring the system to a state of lower energy, we allow the move and put the particle in its new position. If $\Delta E>0$, we allow the move with probability $\exp(-\Delta E/kT)$; i.e., we take a random number ξ_3 between 0 and 1, and if $\xi_3<\exp(-\Delta E/kT)$, we move the particle to its new position. If $\xi_3>\exp(-\Delta E/kT)$, we return it to its old position. Then, whether the move has been allowed or not, i.e., whether we are in a different configuration or in the original configuration, we consider that we are in a new configuration for the purpose of taking our averages. So

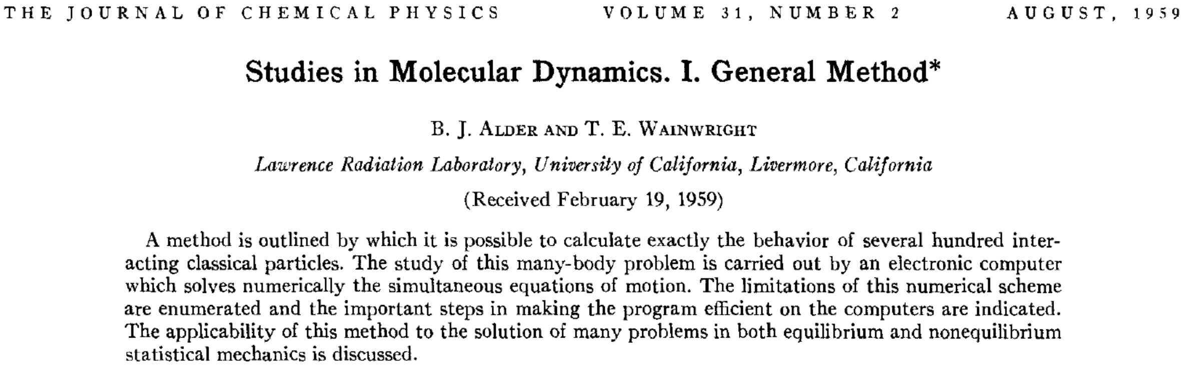
$$\bar{F}=(1/M)\sum_{j=1}^{M}F_j, \tag{4}$$

where F_j is the value of the property F of the system

Historical perspectives on molecular simulations…

First Molecular Dynamics simulation (1959)

IBM-704



THE JOURNAL OF CHEMICAL PHYSICS VOLUME 31, NUMBER 2 AUGUST, 1959

Studies in Molecular Dynamics. I. General Method*

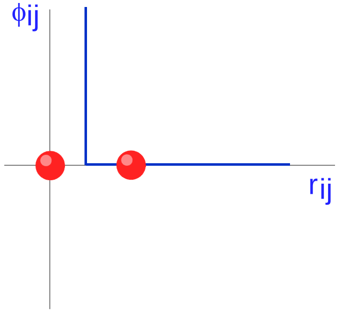
B. J. Alder and T. E. Wainwright

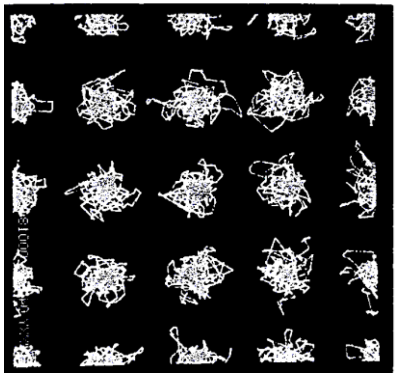
Lawrence Radiation Laboratory, University of California, Livermore, California

(Received February 19, 1959)

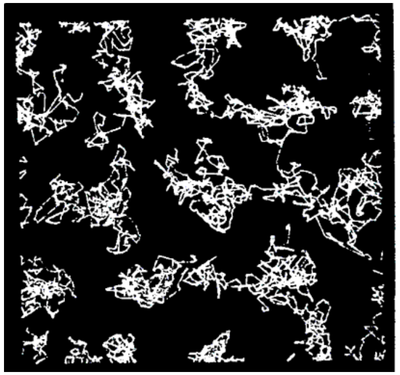
A method is outlined by which it is possible to calculate exactly the behavior of several hundred interacting classical particles. The study of this many-body problem is carried out by an electronic computer which solves numerically the simultaneous equations of motion. The limitations of this numerical scheme are enumerated and the important steps in making the program efficient on the computers are indicated. The applicability of this method to the solution of many problems in both equilibrium and nonequilibrium statistical mechanics is discussed.

Hard disks and spheres

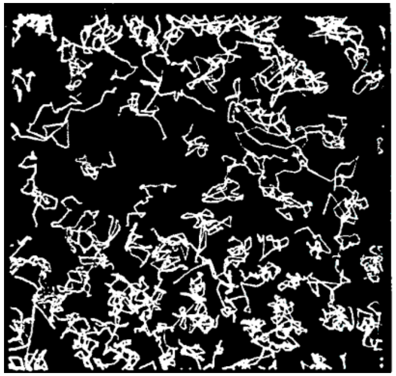




solid phase



liquid phase



liquid-vapour-phase

32 particules 7000 collisions / hcpu

500 particules 500 collisions / hcpu

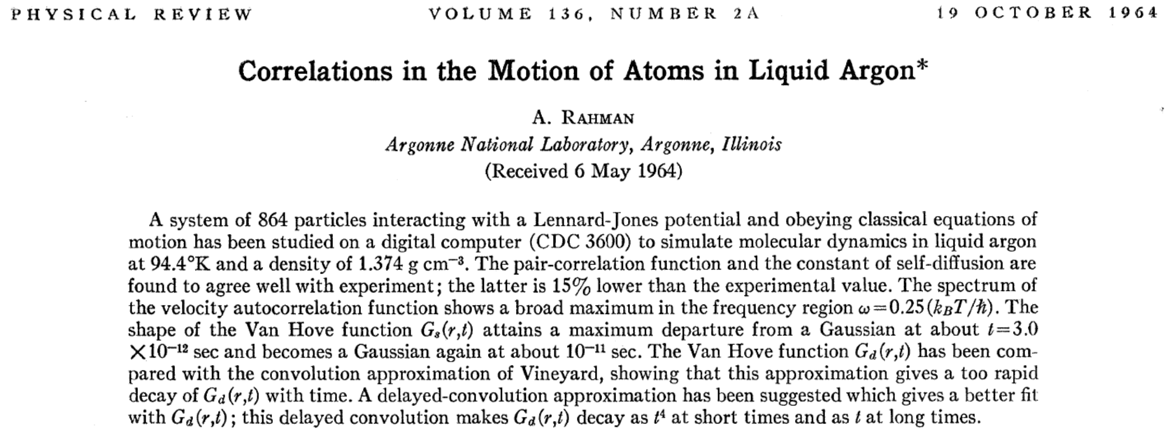
Production run ~20000 steps

$$\begin{cases} N=32 \rightarrow 6.5\times10^5 \text{ coll.} \rightarrow 4 \text{ days} \\ N=500 \rightarrow 10^7 \text{ coll.} \rightarrow 2.3 \text{ years} \end{cases}$$

Historical perspectives on molecular simulations…

First MD simulation with continuous potential (1964)

CDC 3600



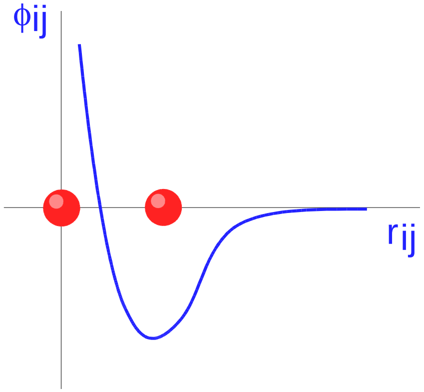
PHYSICAL REVIEW VOLUME 136, NUMBER 2A 19 OCTOBER 1964

Correlations in the Motion of Atoms in Liquid Argon*

A. Rahman
Argonne National Laboratory, Argonne, Illinois
(Received 6 May 1964)

A system of 864 particles interacting with a Lennard-Jones potential and obeying classical equations of motion has been studied on a digital computer (CDC 3600) to simulate molecular dynamics in liquid argon at 94.4°K and a density of 1.374 g cm^{-3}. The pair-correlation function and the constant of self-diffusion are found to agree well with experiment; the latter is 15% lower than the experimental value. The spectrum of the velocity autocorrelation function shows a broad maximum in the frequency region $\omega = 0.25(k_BT/\hbar)$. The shape of the Van Hove function $G_s(r,t)$ attains a maximum departure from a Gaussian at about $t = 3.0 \times 10^{-12}$ sec and becomes a Gaussian again at about 10^{-11} sec. The Van Hove function $G_d(r,t)$ has been compared with the convolution approximation of Vineyard, showing that this approximation gives a too rapid decay of $G_d(r,t)$ with time. A delayed-convolution approximation has been suggested which gives a better fit with $G_d(r,t)$; this delayed convolution makes $G_d(r,t)$ decay as t^4 at short times and as t at long times.

Lennard-Jones potential



Thermodynamics

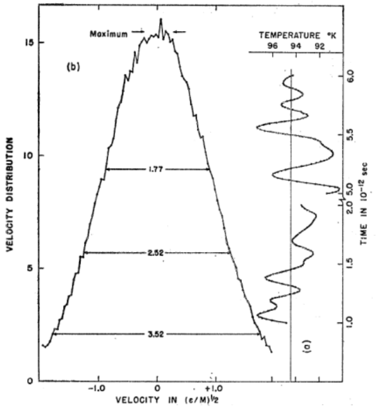


Fig. 1. Fluctuations of temperature with time in two sample regions (curve a); distribution of velocities is shown as curve b; widths of the distribution are shown at $e^{-1/2}$, e^{-1}, and e^{-2} of maximum.

Structure

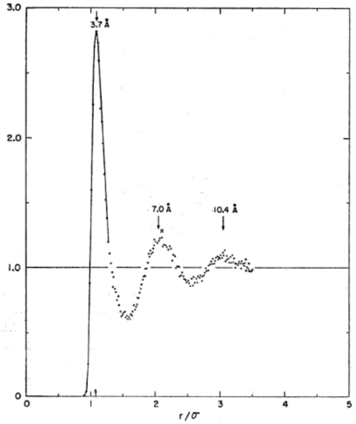


Fig. 2. Pair-correlation function obtained in this calculation at 94.4°K and 1.374 gcm^{-3}. The Fourier transform of this function has peaks at $\kappa\sigma = 6.8$, 12.5, 18.5, 24.8.

Diffusion

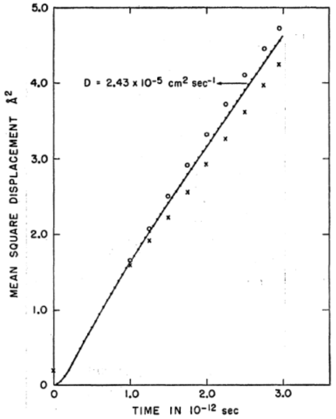
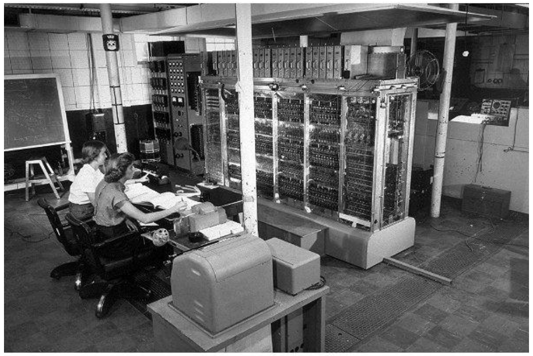


Fig. 3. Mean-square displacement of particles. The continuous curve is the mean of a set of 64 curves; the two members of the set which have *maximum* departures from the mean are shown as circles and as crosses. The asymptotic form of the continuous curve is $6Dt+C$, with D as shown on the figure and $C = 0.2$ Å^2.

864 particles
Time step ~ 10 fs

~20000 MD steps, 200 ps (45 s / step) ≅ 10 days
(standard PC today : few minutes…)

I. Overview of Statistical Mechanics & Molecular Simulations



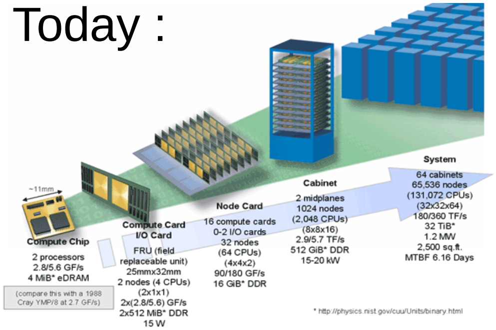
Past :

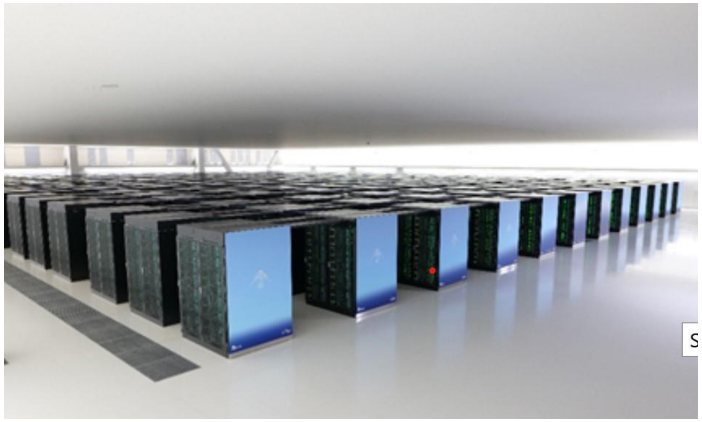
MANIAC
Mathematical Analyzer,
Numerical Integrator,
and Computer

Mars 1952
Los Alamos

1000 OPS

Today :

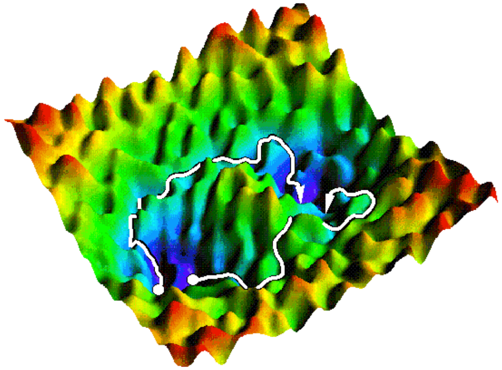




Supercomputer Fugaku (Japan)
Cores : 7,630,848
Memory : 5,087,232 GB
537,212.0 Tflop/s = **5,37 10^{17} Flop/s**

See: https://www.top500.org/lists/top500/

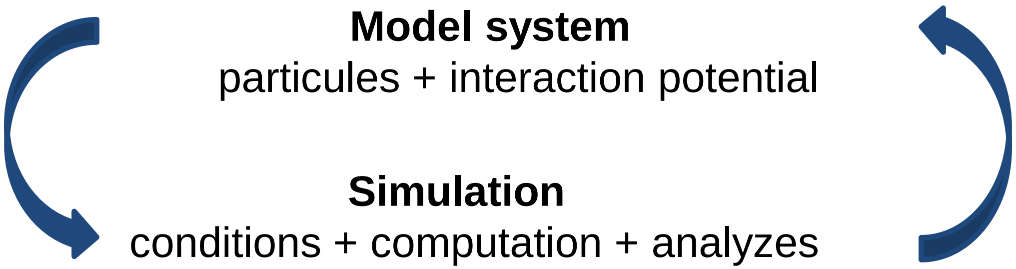
MCIA – Bordeaux (2015)



Simulations…?

- Real condensed phase : phases space → infinity
- It is impossible to know exhaustively the PES (Potential Energy Surface)
- Solution = exploration of the « interesting part » of the PES during the simulation
- By the way… collection and average over the generated states : $< \mathcal{A}(\mathcal{P}_i) > = \frac{1}{N}\sum_{i=1}^{N}\mathcal{A}(\mathcal{P}_i)$

In practical



Model system
particules + interaction potential

Simulation
conditions + computation + analyzes

Summary

« Statistical Mechanics & Simulations »

I. Overview of Statistical Mechanics & Molecular Simulations

II. Molecular Dynamics Simulations

III. Monte Carlo methods

IV. « Outputs » : extracting properties from simulations

V. Initiation to statistical thermodynamics

Molecular Dynamics

The goal

Making the sample : model the motion of some group of particules

The result is a trajectory : collection of phase point P(t) which contain positions and velocities of all particules of the system.

The method

- Let us assume that we know one phase space point P compatible with the macrostate we are interested in.

- Let us assume that our particules are heavy enough and/or the temperature high enough, so that the particles De Broglie wavelength is small compared with the distances between particules.

- We can replace the time dependant Schrodinger equation for the particle motions by Newton’s equation.
- We know the Hamiltonian for the whole things, it contains the kinetic energy and a potential energy based on a chosen model.

II. Molecular Dynamics Simulations

- Let us consider a system of N particles restricted in a volume V
  i.e. P(t0) which correspond to some macrostate

Question : can we characterize this macrostate ?

Answer : yes we can !
we know how to compute the total energy E of P(t0) : $E = E^{total} = T + V$

P(t0) contains all velocities so we can compute the kinetic energy T
P(t0) contains all positions so we can compute the potential energy V

P(t0) is thus characterized by the values N,V and E

The N Newton's equations for all particules in the configuration P(t0) are :

$$m_i\,.\ddot{\vec{R}}_i = -\vec{grad}_i V = \vec{F}_i \qquad i = 1, 2\ldots N$$

If we can solve this N 2d order differential equations, then we could get after some time δt a new configuration P(t0 + δt).

II. Molecular Dynamics Simulations

So we have now :

P(t0) characterized by
N
V
E

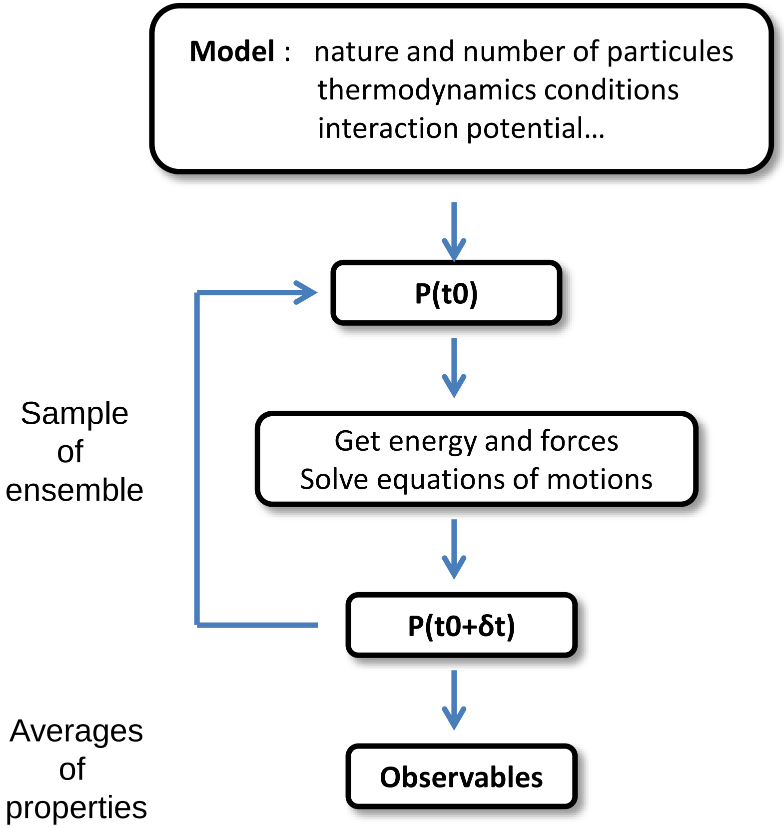
P(t0 + δt) characterized by
N, because Newton's equation do not change the particle numbers
V, because the particules are restricted to this volume
E, because the total energy is constant for an isolated system

Thus :

If P(t0) is a microstate belonging to a certain macrostate

P(t0 + δt) is also a microstate belonging to <u>the same macrostate</u>

Work flow of an MD simulation



Model : nature and number of particules
thermodynamics conditions
interaction potential…

P(t0)

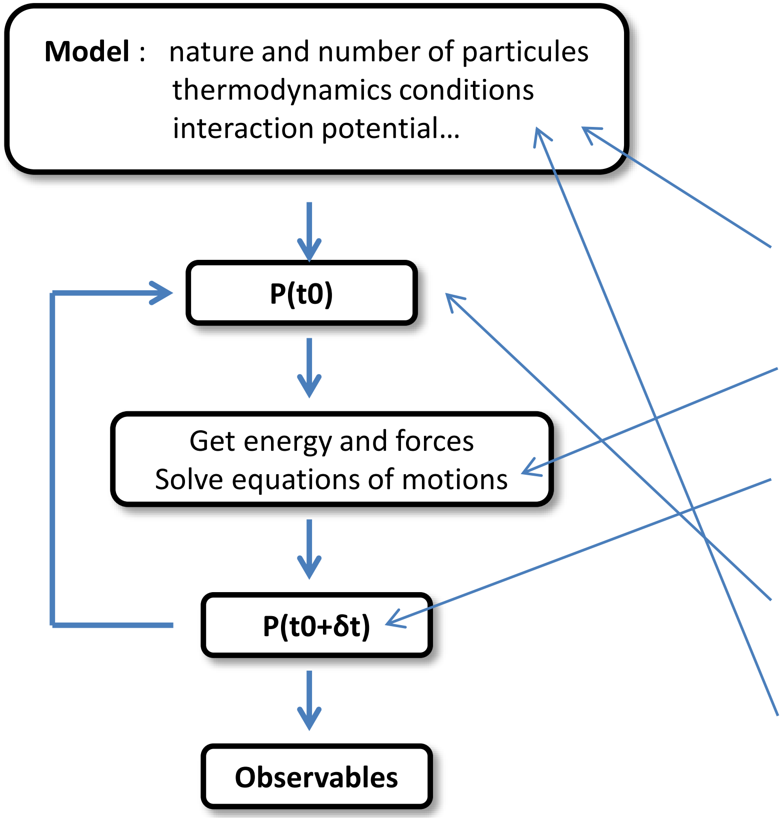
Get energy and forces
Solve equations of motions

P(t0+δt)

Observables

Sample of ensemble

Averages of properties



Model : nature and number of particules
thermodynamics conditions
interaction potential…

P(t0)

Get energy and forces
Solve equations of motions

P(t0+δt)

Observables

Many questions arise :

(a) How to deal with volume restriction

(b) How to deal with the integration of Newton’s equation

(c) How to chose the time step δt

(d) How to find P(t0)

(e) How to control a thermodynamic ensemble

etc.

II. Molecular Dynamics Simulations

Question (a) : how to make sure that the particules are always restricted to the volume

Answer : use of Periodic Boundary Conditions (PBC)
combined with the Minimum Distance Convention (MDC)

Some general comments :

Our spatial scales are limited !

Solutions :

No boundaries
ex. droplet of liquid, molecules in vacuum

Fixed boundaries
make use of a sacrificial region
not really physical, not really efficient

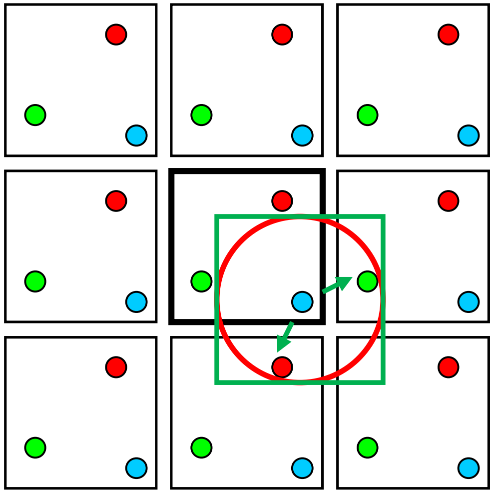
Periodic boundaries
This is the most widely used solution

Basic box

Periodic boxes = mirror of basic box

Minimum Distance Convention

Cut-off sphere



II. Molecular Dynamics Simulations

__Question (a)__ : how to make sure that the particules are always restricted to the volume

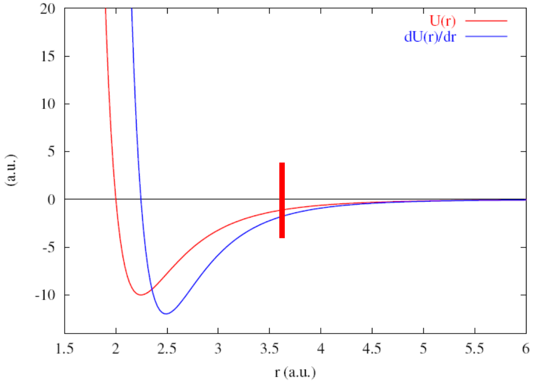
__Answer__ : use of Periodic Boundary Conditions (PBC)
combined with the Minimum Distance Convention (MDC)

It allows to simulate an « infinite » system (or « continous » System) with a finite number of particules.

There is no surface at the limit of the basic box.

Particules crossing a boundary of the simulation box emerge back from the opposite side.
There will always be N particules in the basic box.

Cut-off : no direct interactions beyond a certain distance



Unphysical.
Discontinuities.
Artefacts…

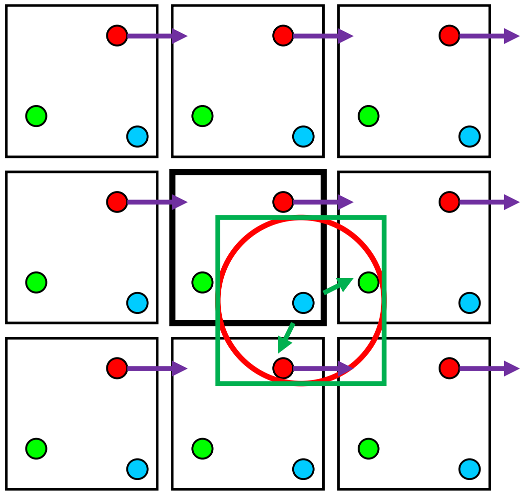
Some methods exists to avoid this problem.

Basic box

Periodic boxes = mirror of basic box

Minimum Distance Convention

Cut-off sphere



Question (b) : how to deal with the integration of Newton's equation

Recall :

$$m_i.\ddot{\vec{R}}_i = -\vec{grad}_i V = \vec{F}_i \qquad i = 1, 2 \ldots N$$

But these N 2d order differential equations are coupled :

$$m_i.\ddot{\vec{R}}_i = -\vec{grad}_i V(\vec{R}_1, \vec{R}_2, \vec{R}_3, \ldots \vec{R}_N)$$

There is no analytic solution for this problem.

Answer : do a numerical solution, i.e. a numerical integration.

Example of the Verlet algoritm :

$$\vec{R}_i(t+\delta t) = \vec{R}_i(t) + \vec{V}_i(t)\delta t + \tfrac{1}{2}\delta t^2 \gamma_i$$

$$\vec{R}_i(t-\delta t) = \vec{R}_i(t) - \vec{V}_i(t)\delta t + \tfrac{1}{2}\delta t^2 \gamma_i$$

$$\vec{R}_i(t+\delta t) = 2\vec{R}_i(t) - \vec{R}_i(t-\delta t) + \delta t^2 \gamma_i$$

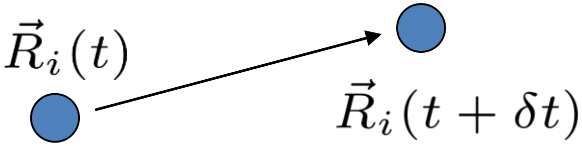
$$\vec{V}_i(t) = \tfrac{1}{2\delta t}(\vec{R}_i(t+\delta t) - \vec{R}_i(t-\delta t))$$

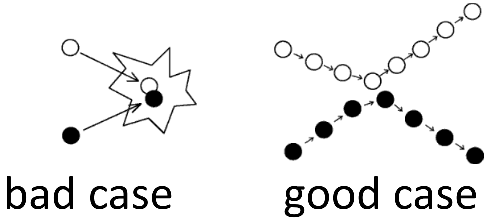
II. Molecular Dynamics Simulations

Question (c) : how to chose the time step δt

Inside the Verlet algoritm, we see : $\vec{R}_i(t+\delta t) = 2\vec{R}_i(t) - \vec{R}_i(t-\delta t) + \delta t^2 \gamma_i$

Such a numerical integration consider that the forces acting on particules are constant during the move :



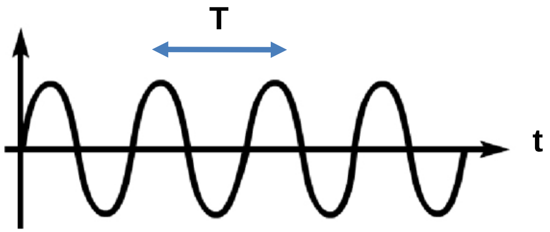


Answer : this condition is only valid for a small step in space with respect to the intermolecular distances.

What about the step in time : δt ?

It depends on the time scale of the motions in the system of interest !
⇒ mass, temperature, interaction model…

Application : determination of the time step δt in a system containing vibrating OH group



$$\bar{\nu}_{OH} = 4000\ cm^{-1}$$

T= ??? δt = T / 30 ≈ ???

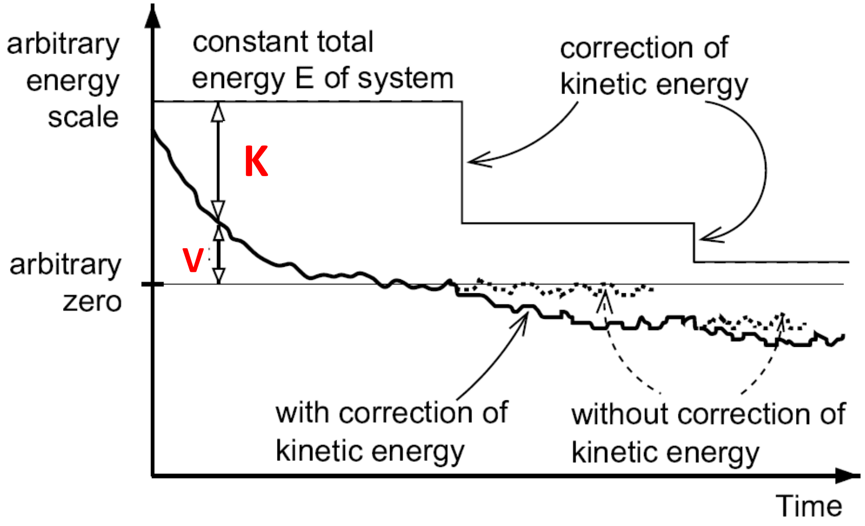
II. Molecular Dynamics Simulations

Question (d) : how to find P(t0) ?

- our macrostates are characterized by N,V,E being constant.
- however, we usually do not know E for the system we want to study.

Answer : this figure shows how to proceed to find a P(t0)

$$E = E^{total} = K + V$$



1) We "guess" a preliminary $\mathcal{P}(t_0)$ (educated guesses preferred!).
2) We start the simulation (numeric integration).
3) The total energy E will be constant if we integrate correctly.
4) However, almost always, our guessed $\mathcal{P}(t_0)$ will have too much potential energy (interactions) compared to its kinetic energy.
5) The kinetic energy ($\propto$ temperature) will thus increase, the potential energy decrease, first quickly, then more and more slowly.
6) If our system is *well-behaved*, it will approach a steady state. (only fluctuations between kinetic and potential energy, no drift.)
7) Trouble! The MACROSTATE that we have is not at all the one we wanted to study (usually much too much kinetic energy!).
8) We now remove kinetic energy from the system (which is easy: scale the velocities) and use the $\mathcal{P}$ that we have now reached as a new, better guess for $\mathcal{P}(t_0)$.
9) We restart the simulation form the new $\mathcal{P}(t_0)$ (GOTO 2) until this $\mathcal{P}(t_0)$ is really the one that we want In this case, we forget this whole "equilibration procedure" and start the **real simulation**.

Question (e) : How to control a thermodynamic ensemble

Thermodynamic ensemble

Microcanonical: NVE

$$\mathcal{W}^{NVE}(\mathbf{r},\mathbf{p}) \;=\; \delta(\mathscr{H}(\mathbf{r},\mathbf{p}) \;-\; E)$$

Canonical NVT

$$\mathcal{W}^{NVT}(\mathbf{r},\mathbf{p}) \;=\; e^{-\mathscr{H}(\mathbf{r},\mathbf{p})/k_B T}$$

Berendsen, J. Chem. Phys. ***81****, 3684 (1984)*
Nosé-Hoover, Phys. Rev. A ***31****, 1695 (1985)*

$$\ddot{\mathbf{r}}_i(t) \;=\; \frac{\mathbf{f}_i(t)}{m_i} \;-\; \xi(t)\dot{\mathbf{r}}_i(t)$$

$$\dot{\xi}(t) \;=\; \nu_T^2 \;\left(\; \frac{2E_c}{f\; k_B\; T_{ext}} \;-1\;\right)$$

Isotherm-isobar : NPT

$$\mathcal{W}^{NPT}(\mathbf{r},\mathbf{p}) \;=\; e^{-(\mathscr{H}(\mathbf{r},\mathbf{p})+PV)/k_B T}$$

Nosé-Andersen, Mol. Phys. ***52****, 255(1984)*

$$\begin{aligned} L \;=\; & \sum_i \frac{m_i}{2} s^2 V^{2/3} \dot{\mathbf{x}}_i^2 \;-\; E(V^{1/3}\mathbf{x}) \\ & + \frac{Q}{2}\dot{s}^2 \;+\; \frac{W}{2}\dot{V}^2 \;-\; (f+1)k_B T_{ext} ln(s) \\ & - P_{ext}V \end{aligned}$$

$$\ddot{\mathbf{r}}_i(t) \;=\; \frac{\mathbf{f}_i(t)}{m_i} \;-\; \frac{\mathbf{p}_i(t)}{m_i}\left(\frac{\dot{s}}{s}\right) \;+\; \frac{\mathbf{r}_i(t)}{3V}\left(\ddot{V} - \frac{2}{3}\frac{\dot{V}^2}{V}\right)$$

$$\ddot{s} \;=\; s\nu_T^2\left(\frac{T(t)}{T_{ext}} - 1\right) \;+\; \frac{\dot{s}^2}{s}$$

$$\ddot{V} \;=\; s^2\left(\frac{P(t)-P_{ext}}{W}\right) \;+\; \frac{\dot{s}}{s}\dot{V}$$

$$\mathbf{p}_i(t) \;=\; m_i\left(\dot{\mathbf{r}}_i(t) - \frac{\dot{V}}{3V}\mathbf{r}_i(t)\right)$$

Summary

« Statistical Mechanics & Simulations »

I. Overview of Statistical Mechanics & Molecular Simulations

II. Molecular Dynamics Simulations

III. Monte Carlo methods

IV. « Outputs » : extracting properties from simulations

V. Initiation to statistical thermodynamics

Monte Carlo

The goal

Making the sample : create a collection of configurations of some group of particules

The result is a randomly ordered collection of configurations, which contain only positions of all particules of the system.

The method

- We put in the computer N particules in a volume V.

- Let us assume that we know one configuration compatible with the macrostate we are interested in.

- Let us assume that we know a model.

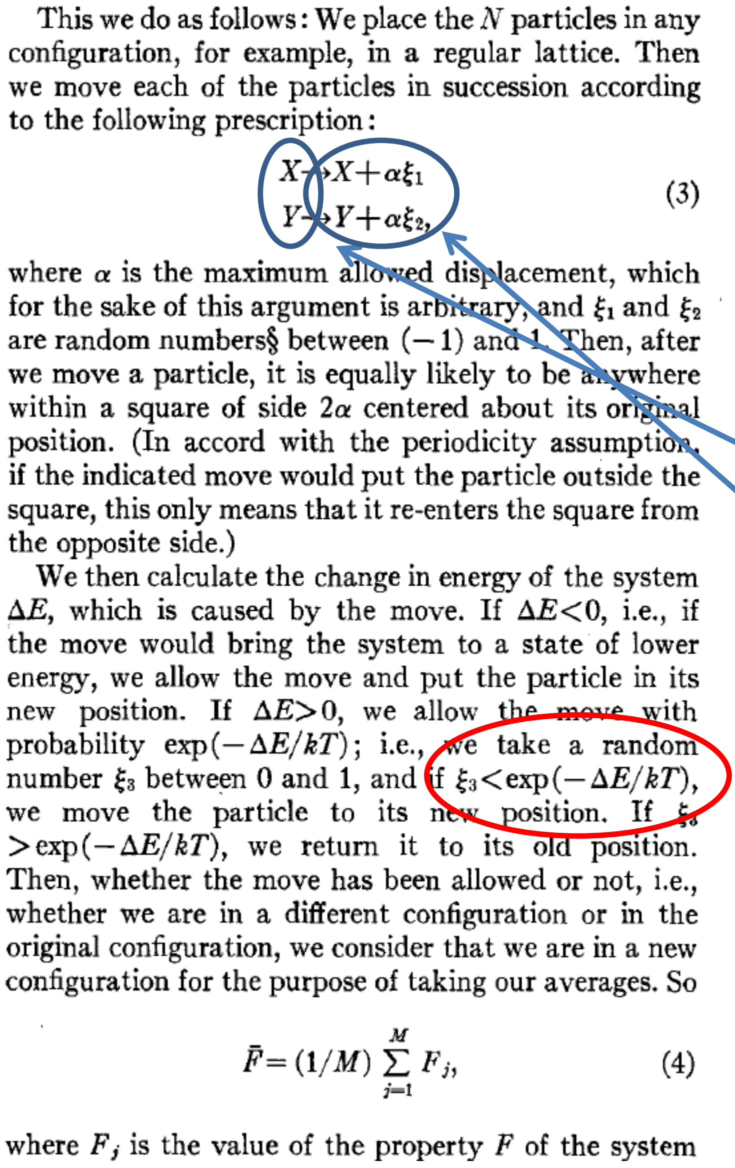
- From the interactions, we can compute the total energy of the system, and thus the energy difference between two configurations (old and new).

- The Metropolis algorithm tell us if the ‘new’ configuration is acceptable.

III. Monte Carlo method

Some details…

First MC simulation (1953)



This we do as follows: We place the N particles in any configuration, for example, in a regular lattice. Then we move each of the particles in succession according to the following prescription:

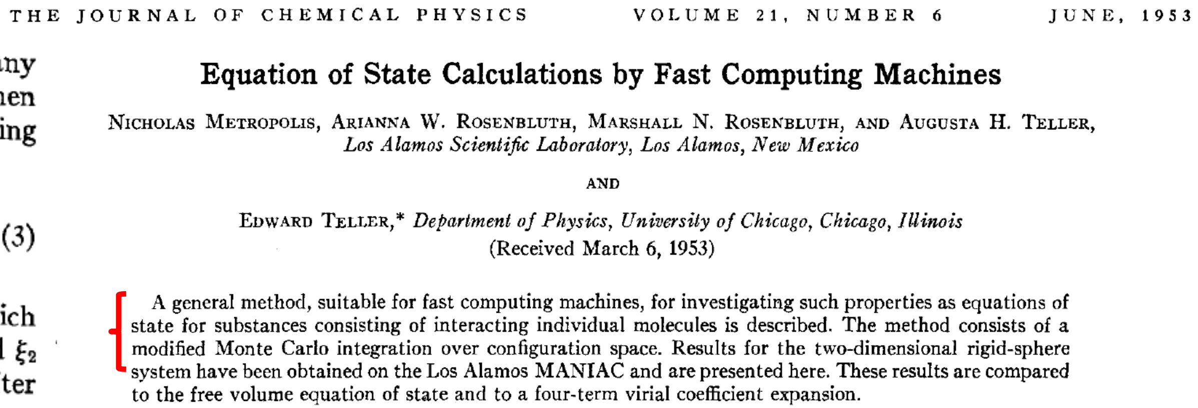
$$X \to X+\alpha\xi_1$$
$$Y \to Y+\alpha\xi_2, \tag{3}$$

where α is the maximum allowed displacement, which for the sake of this argument is arbitrary, and ξ_1 and ξ_2 are random numbers§ between (-1) and 1. Then, after we move a particle, it is equally likely to be anywhere within a square of side 2α centered about its original position. (In accord with the periodicity assumption, if the indicated move would put the particle outside the square, this only means that it re-enters the square from the opposite side.)

We then calculate the change in energy of the system ΔE, which is caused by the move. If $\Delta E<0$, i.e., if the move would bring the system to a state of lower energy, we allow the move and put the particle in its new position. If $\Delta E>0$, we allow the move with probability $\exp(-\Delta E/kT)$; i.e., we take a random number ξ_3 between 0 and 1, and if $\xi_3<\exp(-\Delta E/kT)$, we move the particle to its new position. If $\xi_3>\exp(-\Delta E/kT)$, we return it to its old position. Then, whether the move has been allowed or not, i.e., whether we are in a different configuration or in the original configuration, we consider that we are in a new configuration for the purpose of taking our averages. So

$$\bar{F}=(1/M)\sum_{j=1}^{M} F_j, \tag{4}$$

where F_j is the value of the property F of the system



THE JOURNAL OF CHEMICAL PHYSICS VOLUME 21, NUMBER 6 JUNE, 1953

Equation of State Calculations by Fast Computing Machines

Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, and Augusta H. Teller,
Los Alamos Scientific Laboratory, Los Alamos, New Mexico

and

Edward Teller,* *Department of Physics, University of Chicago, Chicago, Illinois*

(Received March 6, 1953)

A general method, suitable for fast computing machines, for investigating such properties as equations of state for substances consisting of interacting individual molecules is described. The method consists of a modified Monte Carlo integration over configuration space. Results for the two-dimensional rigid-sphere system have been obtained on the Los Alamos MANIAC and are presented here. These results are compared to the free volume equation of state and to a four-term virial coefficient expansion.

1) We know one configuration, call C_1, its energy E_1.

2) We create randomly a new configuration, call C_2.

3) We compute the energy E_2 of C_2 and $\Delta E = E_2 - E_1$.

4) We test the sign of ΔE

$\Delta E \leq 0$, we accept the new configuration :
C_2 become C_1, go back to 2)

$\Delta E > 0$, we accept the new configuration according to a probability based on **Boltzmann factor** :
- if accepted, C_2 become C_1, go back to step 2)
- if rejected, $C_2 = C_1$, go back to step 2)

This kind of 'random walk' through the Cs (Markov chain) generate a representation of the canonic ensemble.

III. Monte Carlo method

1) We know one configuration, call C_1, its energy E_1.

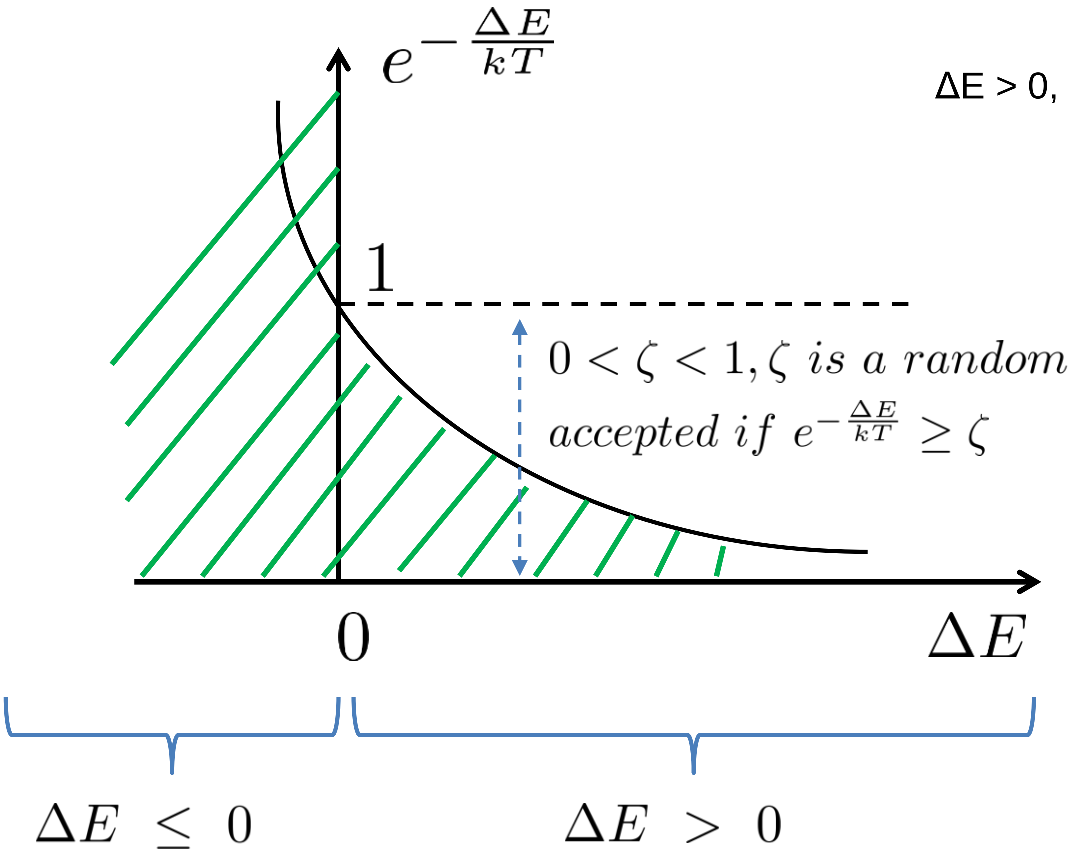
2) We create randomly a new configuration, call C_2.

3) We compute the energy E_2 of C_2 and $\Delta E = E_2 - E_1$.

4) We test the sign of ΔE

$\Delta E \leq 0$, we accept the new configuration :
C_2 become C_1, go back to 2)

$\Delta E > 0$, we accept the new configuration according to a probability based on **Boltzmann factor** :
- if accepted, C_2 become C_1, go back to 2)
- if rejected, $C_2 = C_1$, go back to 2)



$0 < \zeta < 1, \zeta\ is\ a\ random\ number$

$accepted\ if\ e^{-\frac{\Delta E}{kT}} \geq \zeta$

New configuration always accepted

New configuration accepted according to the so-called "asymmetric rule"

MC makes use of random numbers
- to choose a particule to move
- to control the displacement
- for the « asymmetric rule »

The parameter α should be adjusted empirically in such a way that in step 4 approximately one out of two attempted steps is accepted.

III. Monte Carlo method

Additional explanations :

The MC algorithm generate a stationary **Markovian stochastic** sequence which converge to the equilibrium configuration.

Stochastic means random process between two configurations i -> j
Markovian means that the probability of the transition i -> j depend only of the configuration i

Evolution equation :

If P(i,s) is the probability for the system to be in configuration i at « step » s, then :

$$P(i, s+1) = P(i, s) + \sum_j (W(j \to i)P(j, s) - W(i \to j)P(i, s))$$

W being the conditional probability of a transition.

This equation of evolution will lead to an equilibrium if $\sum_j W(j \to i)N_j \;=\; N_i \sum_j W(i \to j)$

A solution, call *micro reversibility*, correspond to $W(j \to i)N_j \;=\; W(i \to j)N_i$

III. Monte Carlo method

$$W(j \rightarrow i)N_j \; = \; W(i \rightarrow j)N_i$$

$$\frac{W(i \rightarrow j)}{W(j \rightarrow i)} = \frac{N_j}{N_i} = exp(-\beta(U_j - U_i))$$

as $\frac{N_i}{N} = \frac{exp(-\beta U_i)}{Z}$

***Consequence** : the unknown parameters W do not depend of Z but only of the Boltzmann factor !*

Proposition of Metropolis et al :

$$W(i \rightarrow j) = \alpha(i \rightarrow j)\pi(i \rightarrow j)$$

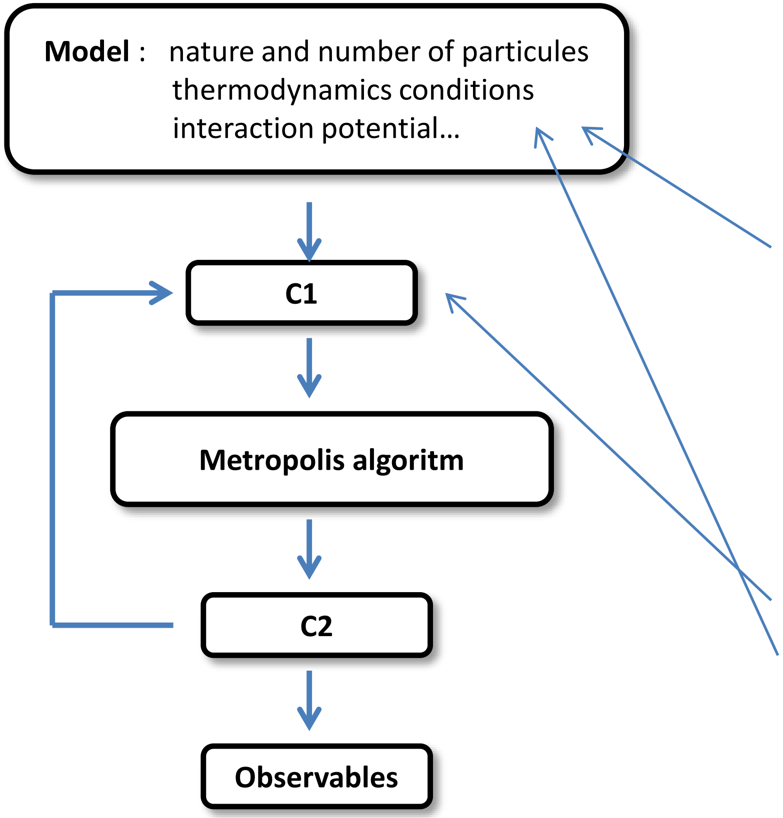
- from a configuration i we choose randomly a configuration j with a probability $\alpha(i \rightarrow j)$
- this new configuration j is accepted with a probability $\pi(i \rightarrow j)$

$$\frac{W(i \rightarrow j)}{W(j \rightarrow i)} = \frac{\pi(i \rightarrow j)}{\pi(j \rightarrow i)} = exp(-\beta(U_j - U_i))$$

Solution of Metropolis
=
Metropolis algorithm

$$\begin{cases} \pi(i \rightarrow j) = exp(-\beta(U_j - U_i)) & if\ U_j > U_i \\ \pi(i \rightarrow j) = 1 & if\ U_j < U_i \end{cases}$$

III. Monte Carlo method



Model : nature and number of particules
thermodynamics conditions
interaction potential…

C1

Metropolis algoritm

C2

Observables

Not so many questions arise :

(a) How to deal with volume restriction

⇒ Same as MD : PBC, MIC…

~~(b) How to deal with the integration of Newton's equation~~

~~(c) How to chose the time step δt~~

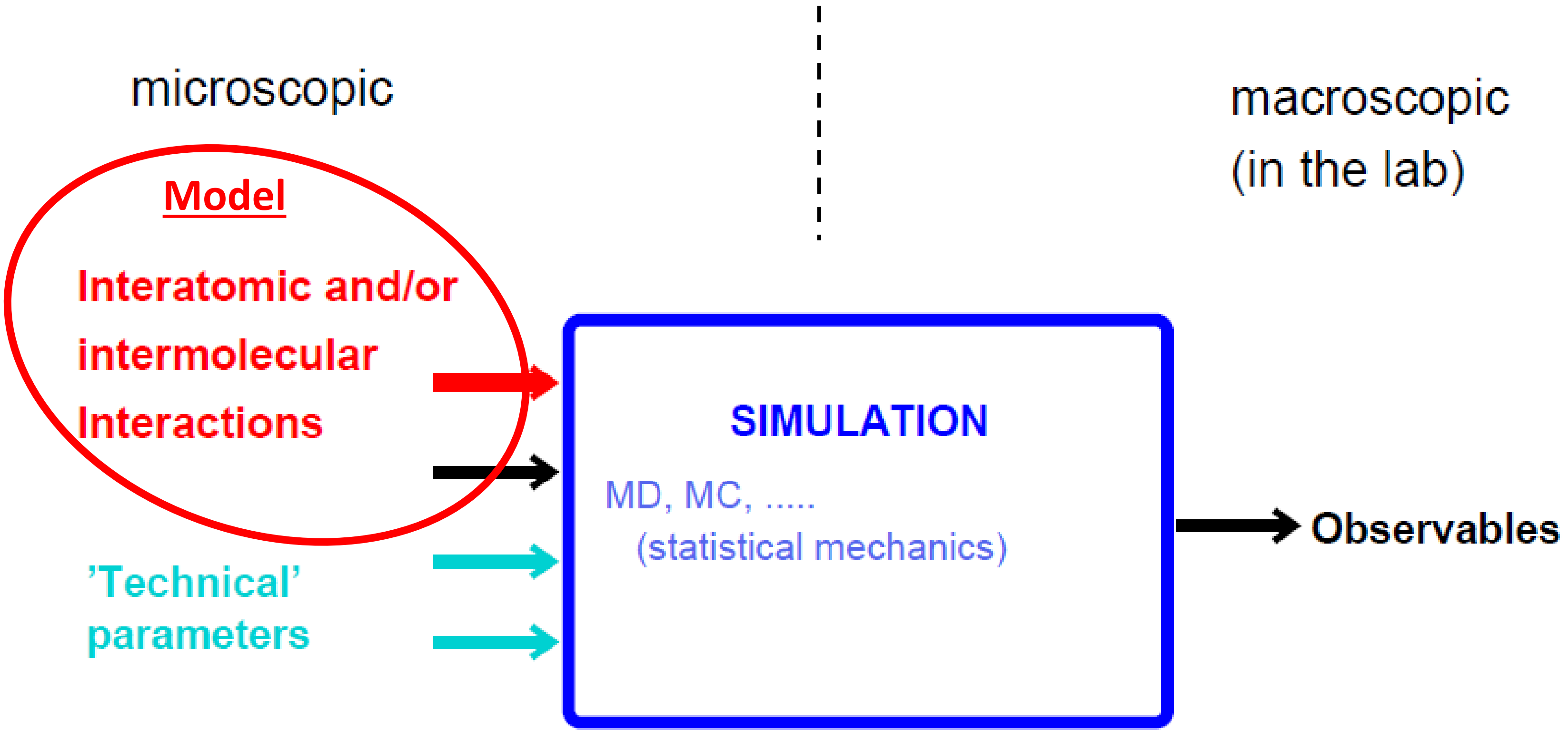
(d) How to find C1

(e) How to control a thermodynamic ensemble

⇒ Nearly the same as MD

etc.

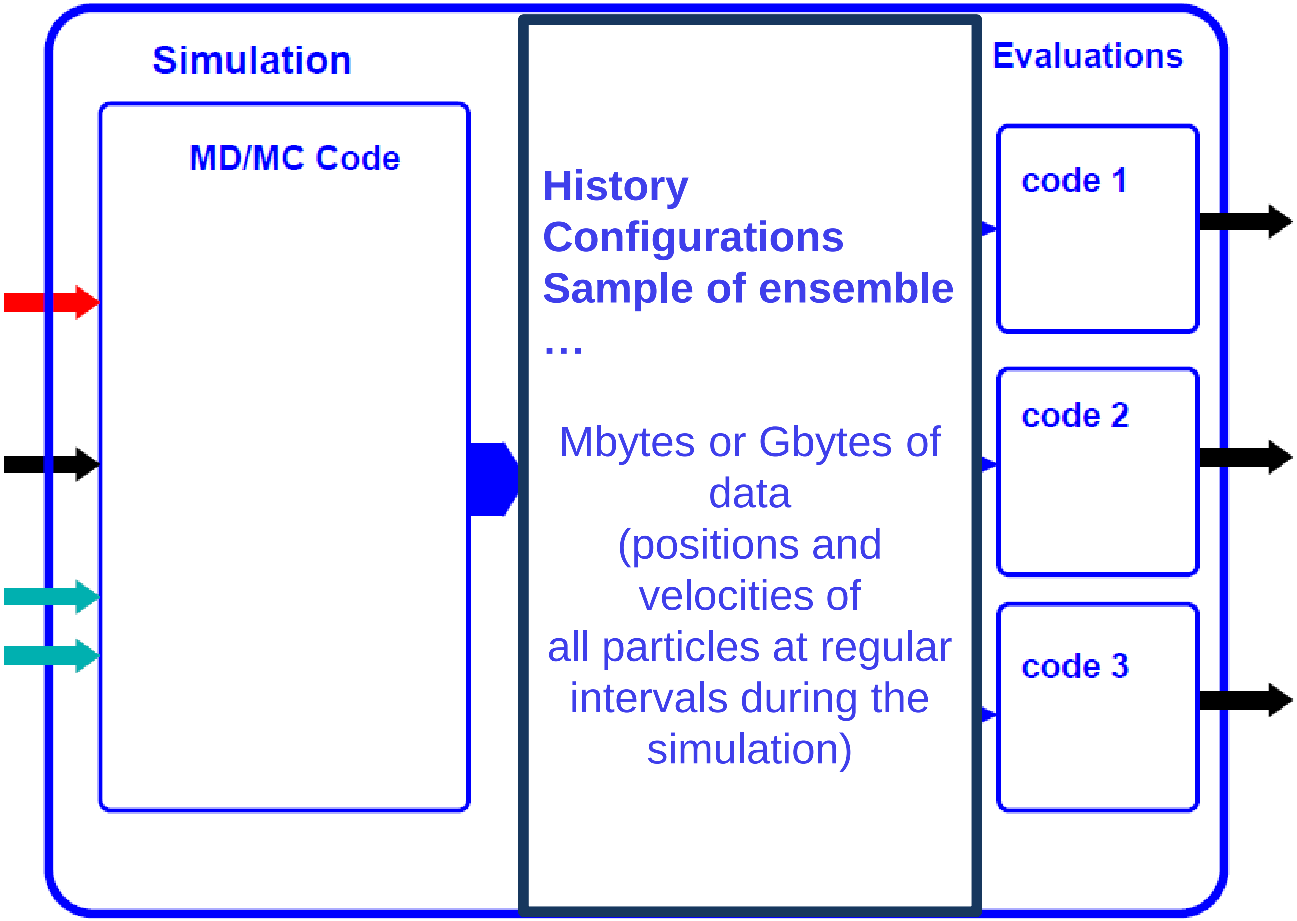
Principle of a simulation (MD or MC) :



In a « good » simulation : the **Model** should be the only input that matters.

i.e. not the technical details; number of particles (size of te box), how to compute the (long range) interactions, how to integrate the equations of motion, etc.

Principle of a simulation (MD or MC) :



Simulation

MD/MC Code

History
Configurations
Sample of ensemble
…

Mbytes or Gbytes of data
(positions and velocities of
all particles at regular intervals during the simulation)

Evaluations

code 1

code 2

code 3

Summary

« Statistical Mechanics & Simulations »

I. Overview of Statistical Mechanics & Molecular Simulations

II. Molecular Dynamics Simulations

III. Monte Carlo methods

IV. « Outputs » : extracting properties from simulations

V. Initiation to statistical thermodynamics

Evaluation of simulations experiments

We now have our sample : $\mathcal{P}_i \ (MD\ case)\ or\ \mathcal{C}_i\ (MC\ case)$

$$i = 1, M$$

M usually equal 100000 up to several millions…

For every quantity A than can be computed from positions (and velocities in MD) of the particules, <u>we can calculate the thermodynamic average as</u> :

$$< \mathcal{A}(\mathcal{P}_i) > = \frac{1}{M} \sum_{i=1}^{M} \mathcal{A}(\mathcal{P}_i)$$ *over the M microstates*

This is the most important formula of this whole simulation business !

A can be a basic thermodynamic observables (such as temperature, internal energy, pressure), a property describing the structure of the system,… **<u>or any other property you might imagine</u>** .

IV. « Outputs » : extracting properties from simulations

Some examples of quantities A :

Internal energy U:

$$U \; = < \mathcal{H}(\mathbf{r},\mathbf{p}) > = \; K \; + \; V \; = < \mathcal{K}(\mathbf{p}) > \; + \; < \mathcal{V}(\mathbf{r}) >$$

Kinetic energy: $< \mathcal{K}(\mathbf{p}) > = < \sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{2m_i} >$

Potential energy : $< \mathcal{V}(\mathbf{r}) >$

Temperature :

$$T \; = < \mathcal{T}(\mathbf{p}) > = \; \frac{2}{f k_B} \; < \mathcal{K}(\mathbf{p}) > = \; \frac{2}{f k_B} < \sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{2m_i} >$$

f *: number of degrees of freedom of the system*

Pressure :

$$P \; = \; \frac{N k_B}{V} T \; - \; \frac{1}{3V} < \mathcal{W}(\mathbf{p}) >$$

Ideal part Non ideal part ↔ interactions

Virial : $< \mathcal{W}(\mathbf{p}) > = \; - \; < \sum_{i=1}^{N-1} \; \sum_{j>i}^{N} \; \mathbf{R}_{ij} \; \mathbf{F}_{ij} >$

IV. « Outputs » : extracting properties from simulations

Structure : *Pair correlation functions, the way to characterize the structure of a disordered systems.*

An average of the quantity

$$a(r;\mathcal{P}_M) = \sum_i \delta(r_i - r)$$

represents the probability densitiy of some particle being situated near r position in each configuration.
The average is thus the mean fluid density at r position :

$$\langle a(r)\rangle = \rho(r) = \langle \sum_i \delta(r_i - r)\rangle$$

As the density is a constant in a fluid, we can write the « pair correlation function » (PCF) :

$$g(r) = \frac{V}{4\pi r^2 N^2}\langle \sum_i \sum_{j\neq i} \delta(r - r_{ij})\rangle$$

It is the conditional probability density of finding a particle at r, given that there is a particle at the coordinate origin.

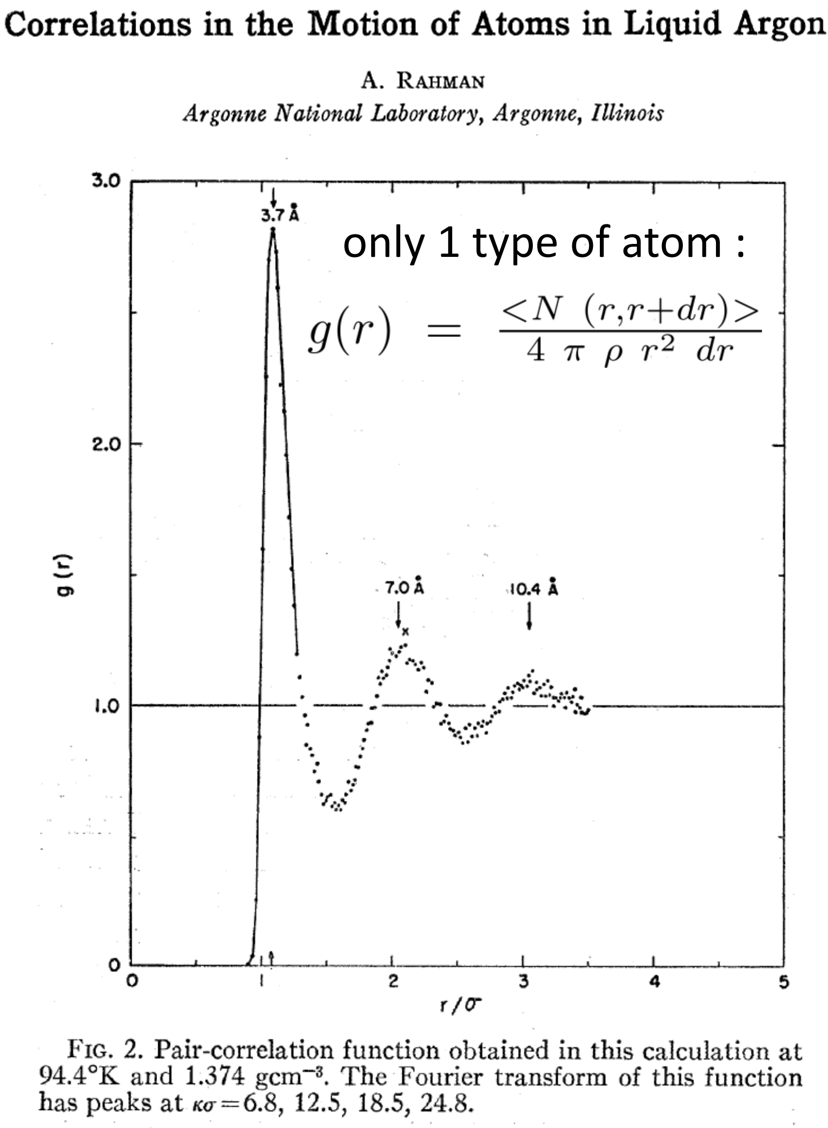
In a system containing different type of atoms, there are as many PCF as there are types of pairs :

$$\boxed{g_{\alpha\beta}(r) = \frac{<N_\beta \;(r,r+dr)>}{4\ \pi\ \rho_\beta\ r^2\ dr}}$$

where N is the number of atoms of type β at a distance between (r) and (r+dr) from an atom of type α
and ρ_β the number density of atom β

IV. « Outputs » : extracting properties from simulations

Example of pair correlation function :



Correlations in the Motion of Atoms in Liquid Argon*

A. Rahman

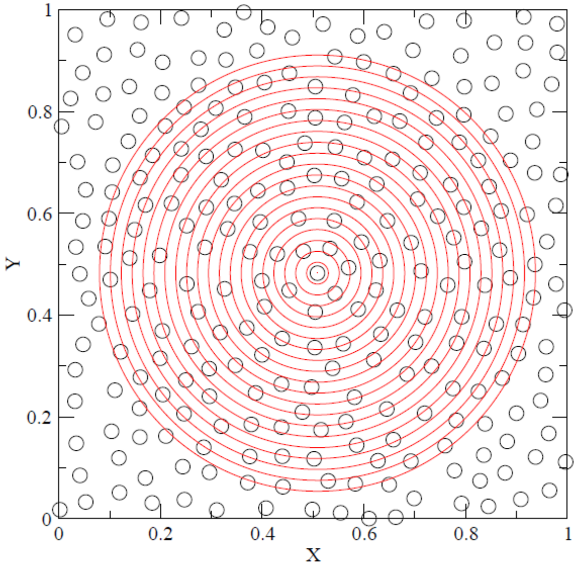
Argonne National Laboratory, Argonne, Illinois

only 1 type of atom :

$$g(r) = \frac{<N\ (r,r+dr)>}{4\ \pi\ \rho\ r^2\ dr}$$

Fig. 2. Pair-correlation function obtained in this calculation at 94.4°K and 1.374 gcm^{-3}. The Fourier transform of this function has peaks at $\kappa\sigma = 6.8$, 12.5, 18.5, 24.8.

Calculation principle :



Questions :

(a) What mean these oscillations ?

$\Rightarrow$ *a deviation from the mean density*

(b) What is the responsible of these oscillations ?

$\Rightarrow$ *the interactions between particules*

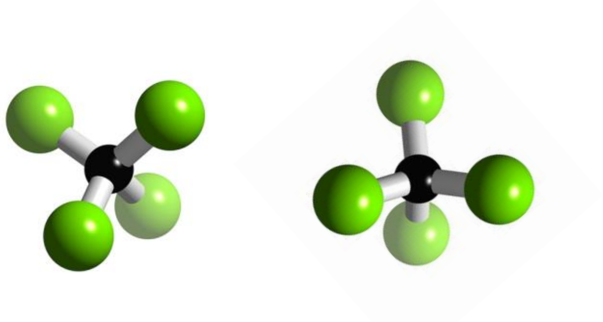
(c) Why do it converge to 1 at high value of r ?

$\Rightarrow$ *the normalization (ρ) in the formulae*

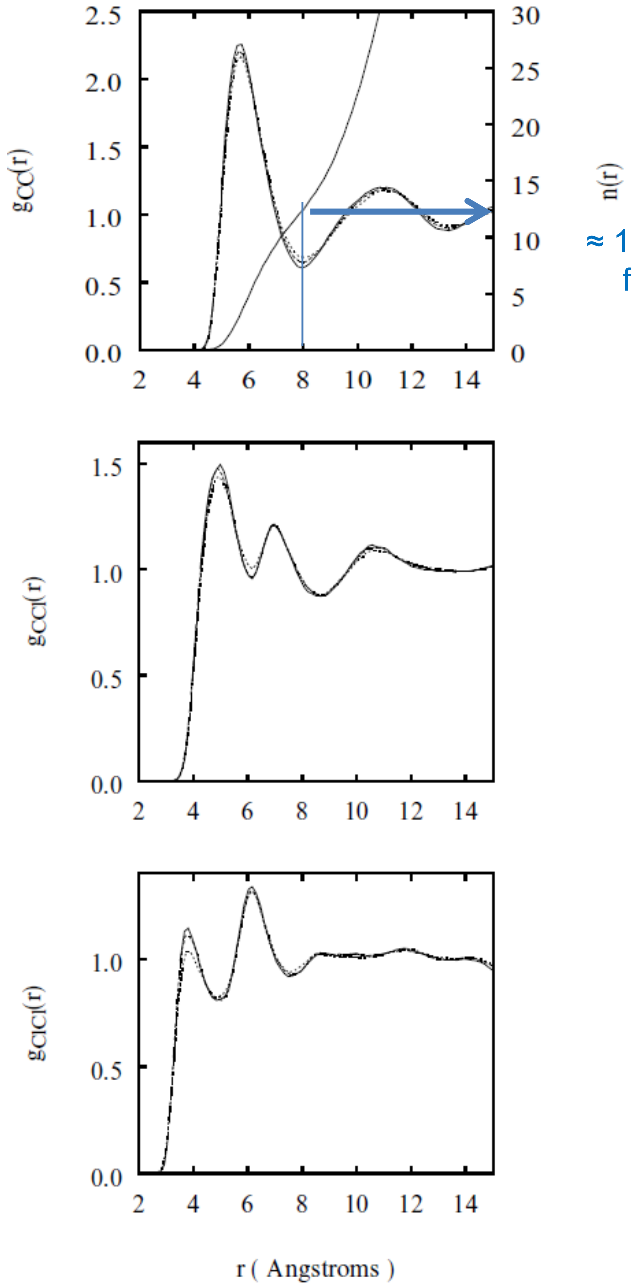
IV. « Outputs » : extracting properties from simulations

Example of pair correlation functions :

Liquid carbon tetrachloride : CCl_4



for a dimer : 1 $g_{CC}(r)$
8 $g_{CCl}(r)$
16 $g_{ClCl}(r)$



≈ 13 molecules in the first shell.

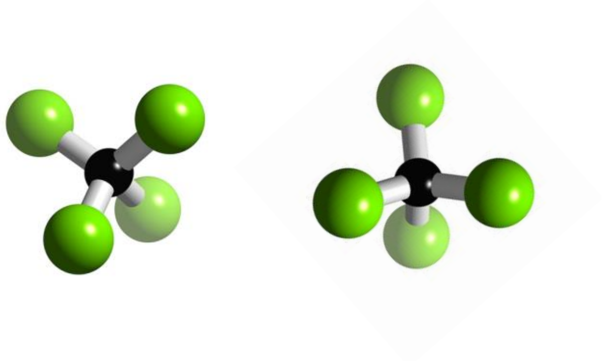
Running coordination number :

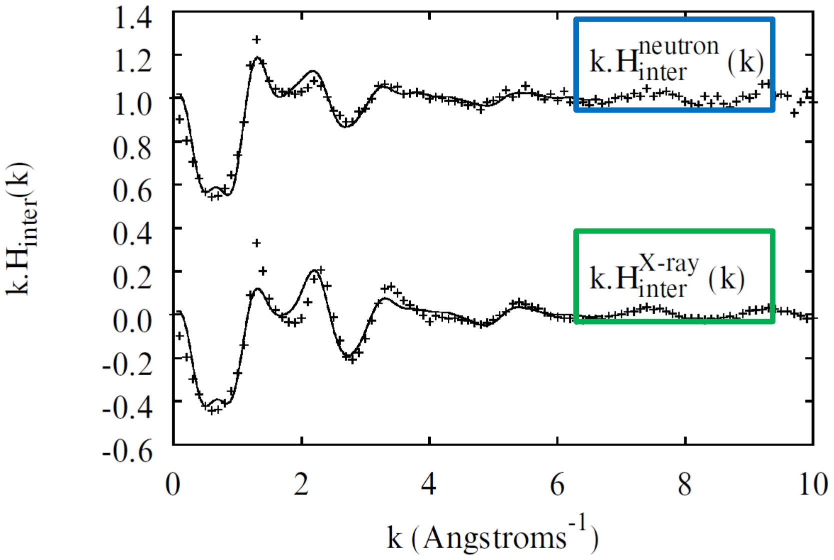
$$N_\beta(R) = \int_0^R 4\ \pi\ \rho_\beta\ g_{\alpha\beta}(r)\ r^2\ dr$$

IV. « Outputs » : extracting properties from simulations

Comparison with experiments : neutron and X-ray diffraction experiments

Liquid carbon tetrachloride : CCl_4





The intermolecular total structure factor $H_{\text{inter}}(k)$ is obtained from

$$H_{\text{inter}}(k) = 4\pi\rho \int_0^{\infty} r^2 \frac{\sin(kr)}{kr} (G_{\text{inter}}(r) - 1)\, \mathrm{d}r, \quad (8)$$

where ρ is the molecular number density and $G_{\text{inter}}(r)$ is a weighted sum of the partial intermolecular radial distribution functions according to

$$G_{\text{inter}}(r) = \sum_{\alpha,\beta} \frac{b_\alpha b_\beta g_{\alpha\beta}(r)}{\left(\sum_\alpha b_\alpha\right)^2} \quad (9)$$

and b_α are the coherent scattering lengths of atoms ($b_{\text{C}} = 6.65$ fm, $b_{\text{Cl}} = 9.58$ fm for neutron diffraction). This leads to

$$G_{\text{inter}}^{\text{neutron}}(r) = 0.02 g_{\text{CC}}(r) + 0.25 g_{\text{CCl}}(r) + 0.73 g_{\text{ClCl}}(r), \quad (10)$$

$$G_{\text{inter}}^{\text{X-ray}}(r) = 0.01 g_{\text{CC}}(r) + 0.15 g_{\text{CCl}}(r) + 0.84 g_{\text{ClCl}}(r), \quad (11)$$

IV. « Outputs » : extracting properties from simulations

Question : when we use the formula

$$< \mathcal{A}(\mathcal{P}_i) > = \frac{1}{M}\sum_{i=1}^{M}\mathcal{A}(\mathcal{P}_i)$$

over the M microstates obtained in a MD simulation

or

$$< \mathcal{A}(\mathcal{C}_i) > = \frac{1}{M}\sum_{i=1}^{M}\mathcal{A}(\mathcal{C}_i)$$

over the M microstates obtained in a MC simulation

What did we miss !?!?

Answer : the fact that in MD, the Ps are ordered,
so the fact that they contain the time evolution of the system !

We can thus define temporal correlations functions to study the time evolution of any properties of the system we are interested in.

Of course, this concept is only valid for MD simulations !

Time auto-correlation function

$$c_{AA}(t) = \frac{1}{\mathrm{Norm}} \cdot < \ A(\mathcal{P}(\tau)) \odot A(\mathcal{P}(\tau + t)) \ >_{\tau}$$

Method :

- we compute, like before, the quantity we are interested in, A, for each P.
- we relate with each other through an operator (scalar produc, +,…) the quantities A at two times (τ) and (t+ τ).
- we look at the result as a function of the time delay (t) between both.
- in the equilibrium ensemble, we average over all “initial” times (τ) and get the auto-correlation function.

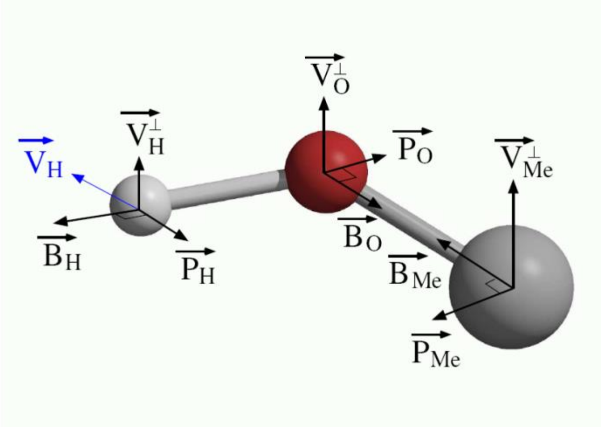
With this concept, we are here leaving the domain of statistical thermodynamics and enter the more general domain of statistical mechanics.

These functions are important objects in statistical mechanics (see e.g. D. McQuarrie's book). In particular, one shows that:

1) the integrals over these functions are related to the so-called “transport coefficients”.
   The transport coefficient related to the velocity autocorrelation function is the self-diffusion coefficient Ds.

2) the Fourier transforms of these functions are called "spectra". They are not only called spectra, they are spectra like the ones measured in Infrared, Raman, neutron inelastic scattering, ... spectroscopies.
   The Fourier transform of the velocity autocorrelation function is related to neutron scattering.

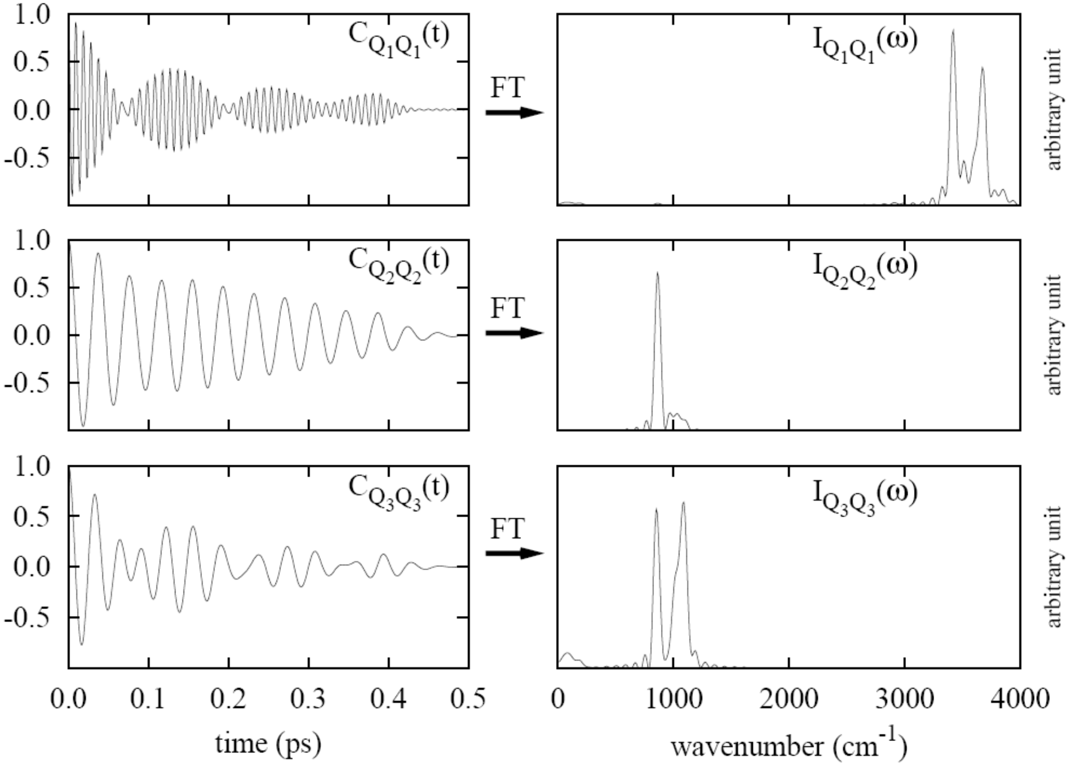
Example : atomic velocity autocorrelation functions in methanol



MD simulation of a methanol dimer in a solvent (CCl4).

Normal mode like combinations

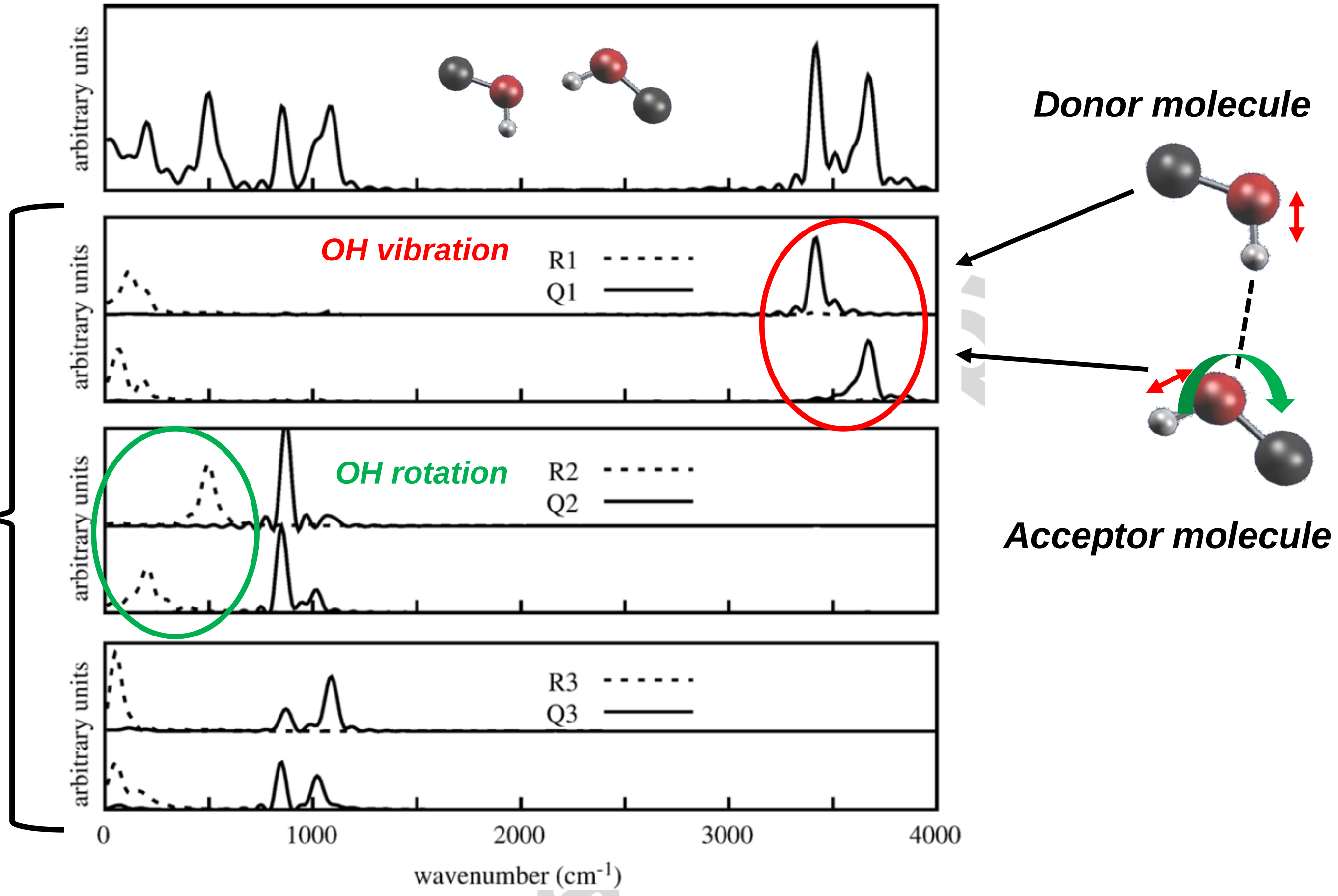
$Q_1 = B_H$	O–H stretching
$Q_2 = B_{Me} + B_O$	O–Me stretching
$Q_3 = P_H + P_{Me}$	H–O–Me bending
$R_1 = P_O + P_{Me}$	Rotation in plane
$R_2 = V_H^\perp$	Rotation of H roughly around the O–Me bond
$R_3 = V_O^\perp - V_{Me}^\perp$	Rotation roughly around an axis perpendicular to the O–Me bond



Example : atomic velocity autocorrelation functions in methanol (dimer in CCl_4)

Spectral density from the total velocities of both H atoms

Fourier transforms of the autocorrelation functions Qs and Rs



IV. « Outputs » : extracting properties from simulations

Generalisation : Time cross-correlation function

$$c_{AB}(t) = \frac{1}{\mathrm{Norm}} \cdot < \; A(\mathcal{P}(\tau)) \odot B(\mathcal{P}(\tau + t)) \; >_\tau$$

Conclusions :

We have two types of averaging procedures over our sample:

(I) We look at one configuration P(t) at a time and average over all configurations.
These averages led to the often called “static properties"

(II) We look at two conguration P(t1) and P(t2) and average over all pairs of configurations which have the same spacing in time
These averages led to the often called “dynamic properties”

Summary

« Statistical Mechanics & Simulations »

I. Overview of Statistical Mechanics & Molecular Simulations

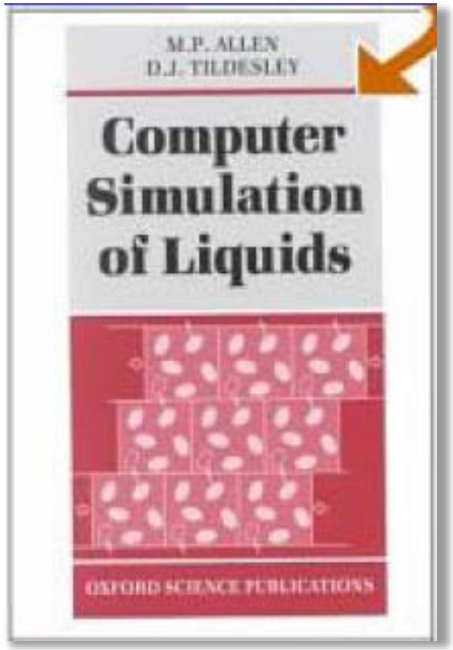
II. Molecular Dynamics Simulations

III. Monte Carlo methods

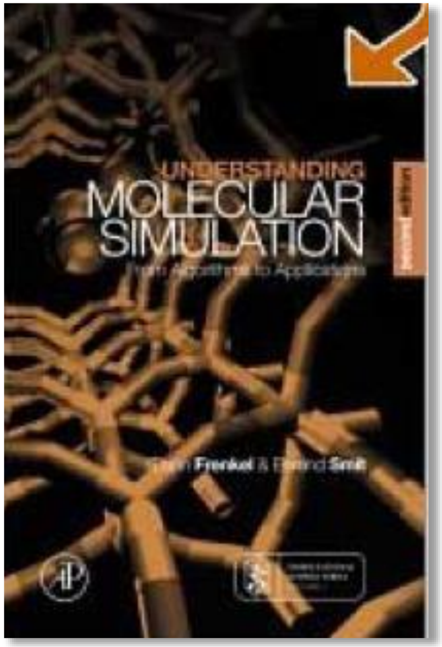
IV. « Outputs » : extracting properties from simulations

V. Initiation to statistical thermodynamics

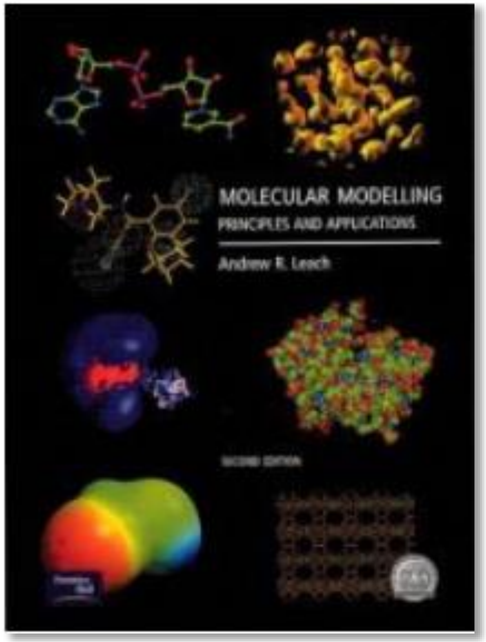
References



M. P. Allen & D. J. Tildesley
Computer simulation Of liquids
Clarendon Press,
Oxford, 2006.



D. Frenkel & B. Smit
Understanding molecular simulation
Academic Press,
New York, 1996



A.R. Leach
Molecular Modelling
Principles and Applications
AIP, Dorchester, 2001.

See also :

J. A. McCammon, S. C. Harvey (1987), Dynamics of Proteins and Nucleic Acids, Cambridge University Press

D. C. Rapaport (1996), The Art of Molecular Dynamics Simulation

J. M. Haile (2001), Molecular Dynamics Simulation: Elementary Methods, Wiley

A. Gerschel (1995), Liaisons intermoléculaires, EDP Sciences, CNRS Edition

Contact

Pr. Jean-Christophe Soetens

Institut des Sciences Moléculaires
Groupe Chimie Théorique & Modélisation
UMR CNRS - Université de Bordeaux 5255
351, Cours de la Libération
F-33405 Talence Cedex

(++33) (0) 540 00 22 42

jean-christophe.soetens@u-bordeaux.fr

http://theo.ism.u-bordeaux.fr/J-C.Soetens/