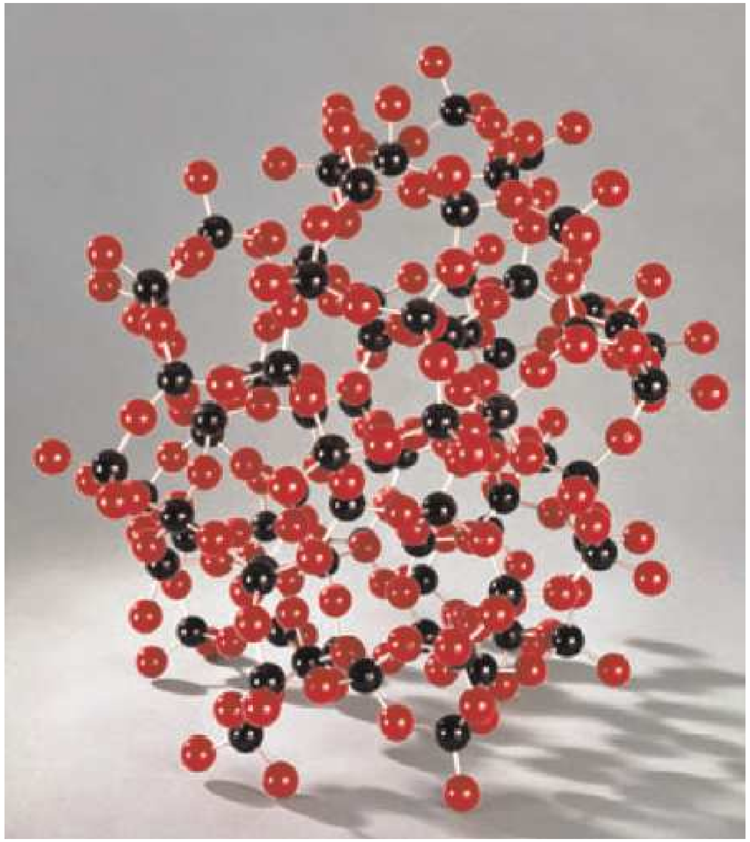
LECTURE 6 : BASICS FOR MOLECULAR SIMULATIONS

- Historical perspective
- Skimming over Statistical Mechanics
- General idea of Molecular Dynamics
- Force calculations, structure of MD, equations of motion
- Statistical Ensembles

matthieu.micoulaut@upmc.fr

What we want to avoid …

Computing the radial distribution function of vitreous SiO_2 from a ball and stick model



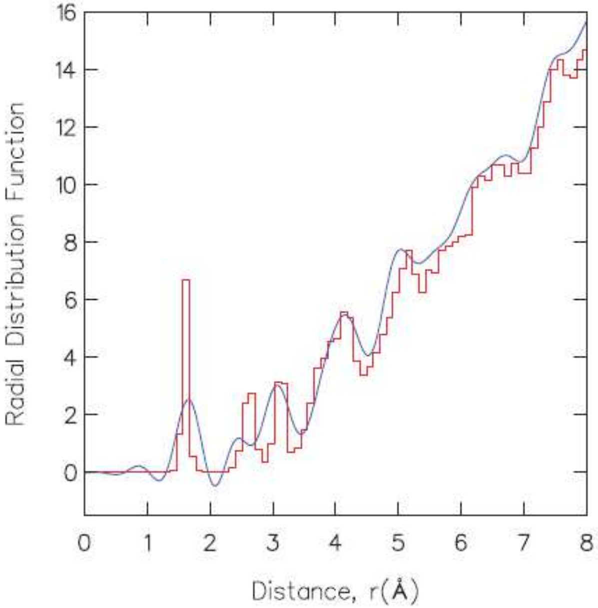


Figure 4 (online colour at: www.pss-b.com) A comparison of the relaxed [9] Bell and Dean [4] model (1980) for vitreous silica (red histogram) with X-ray data (blue curve). The histogram does not include thermal broadening, nor the experimental broadening in real space, whilst the X-ray data of Mozzi and Warren [6] have been Fourier transformed to simulate the earlier (1936) Radial Distribution Function of Warren et al. [10].

Bell and Dean, 1972

matthieu.micoulaut@upmc.fr

1972 … and the historical perspective of molecular simulations

~1900 Concept of force field in the analysis of spectroscopy

1929 Model vibrationl excitations: atomic potentials (P.M. Morse and J.E. Lennard-Jones)

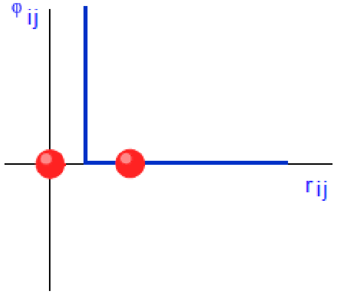
1937 London dispersion forces due to polarisation (origin Van der Waals forces),

1946 Molecular Mechanics: use of Newton's equations and force fields for the caracterization of molecular conformations

1953 Monte Carlo simulations after the Manhattan project (1943): computation of thermodynamic properties (Metropolis, Van Neuman, Teller, Fermi)

1957 Hard sphere MD simulations (Alder-Wainwright) :
Potential non-differentiable, no force calculations,
free flight during collisions, momentum balance.

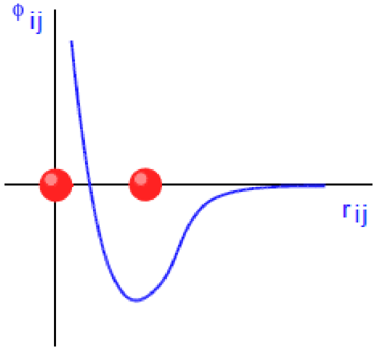
Event-driven algorithms



matthieu.micoulaut@upmc.fr

1964 LJ MD of liquid Argon (Rahman): differentiable potential, solve Newton’s equation of motion, accurate trajectories.

Contains already the main ingredients of modern simulations



1970s Simulation of liquids (water, molten salts and metals,…) Rahman, Stillinger, Catlow …

Algorthims to handle long-range Coulomb interactions (Ewald summation method,…)

Development of potentials (Born-Mayer-Huggins, BMH)

1976 First simulations of SiO2 (silica) using a BMH potential. Achieves tetrahedral coordination (Woodcock, Angell, Soules,…)

1980 Andersen constant-pressure algorithm

Rahman Parrinello constant pressure algorithm

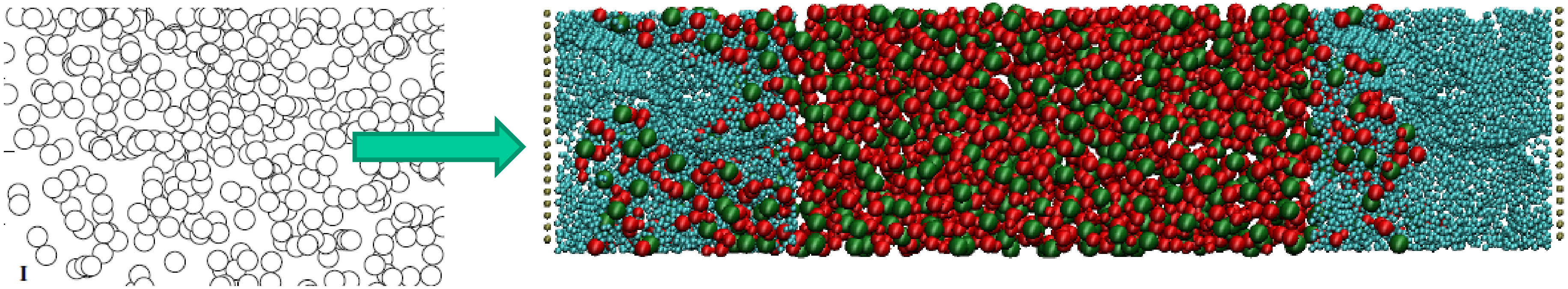
matthieu.micoulaut@upmc.fr

1985 Car-Parrinello quantum mechanical MD (CPMD) using density functional theory

1990s Improvment of interaction potentials (Stillinger-Weber, Vashishta, Finney, Ciccotti,)
3-body, mutli-body,…
Charge transfer reactive force fields (Goddard, Madden,…)

1992 Transfer of the CECAM (Centre Européen de Calcul Atomique et Moléculaire) to Lyon (Fr). Promotion and tutorials of advanced computational methods in Materials Science.

2000s Improvment and massive diffusion of CPMD techniques among the community
Need of massively parallel computing (MPI, OpenMPI,…)



Performances

Computing power of computers : units of Gflops (billion of floating point operations per second).
Scientific programs incorporating libraries which perform parallel computing (OpenMPI, MPI)

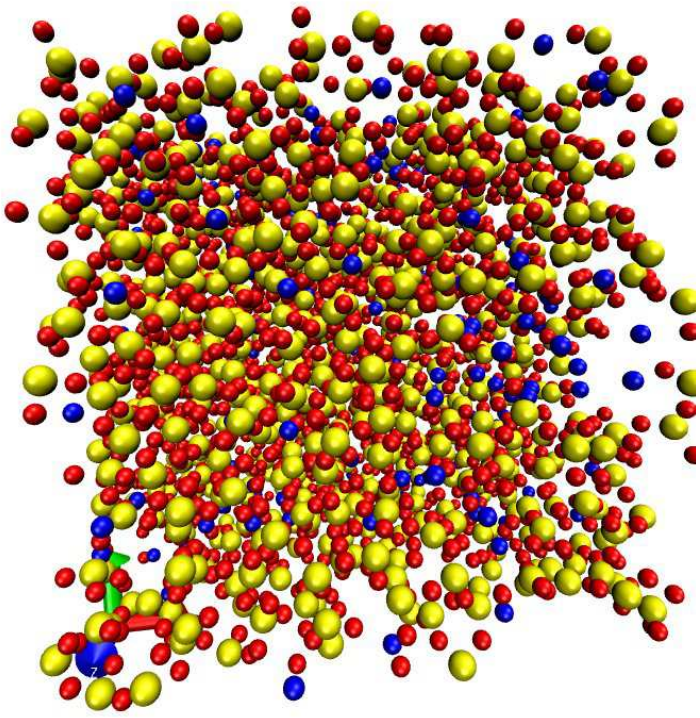
Personal PC computer (4 core processor)	**24 Gflops**
Titan (n.1, Oak Ridge Nat. Lab.)	**27×10^6 Gflops**

Pflops= exponential function of time (Moore's law)

1979	300 atoms (MD)	**2012**	10^7 atoms (MD)
		2012	500 atoms (QMD)

2013: a few years years after 1972…

3000 atom **system** of SiO_2-$2Na_2O$ glass



Thermodynamic **system** described either by macroscopic variables

V (volume),
P (pressure)
T (temperature)
N_i (nb of particles)

Or with microscopic ones

$\mathbf{v}_i$ (velocities)
$\mathbf{r}_i$ (positions)

Linking (v_i ,r_i) with (V,T,P) is the subject of statistical mechanics

matthieu.micoulaut@upmc.fr

Linking (v_i ,r_i) with e.g. (V,T,P) is the subject of statistical mechanics

- In a MD simulation of a glass or a liquid, one can measure instaneous positions and velocities at any time for all atoms.
- Hardly comparable to experimental data. There is no real experiment able to provide such detailed information.
- Rather, a typical experiment measures an average property (over time and a large number of particles 10^{23}).
- To be comparable, one must extract from the basic MD information averages.
- The computation of such averages needs the definition of **Statistical Mechanics** (Ensembles).

Basic Statistical Mechanics: Phase space

❑ At a microscopic level, a system made of N particles is caracterized by the positions $\mathbf{r}_i$ and momenta $\mathbf{p}_i=m\mathbf{v}_i$.

❑ These ($\mathbf{r}_i$,$\mathbf{p}_i$) can take any values (or **states**), constrained by the external (thermodynamic) parameters.

❑ The **phase space** is the space in which all possible states of a system are represented. Each possible state of the system corresponds to one unique point in the phase space (**microstate**). Its differential volume element is :

is

$$d\mathbf{r}^N d\mathbf{p}^N = dx_1 dy_1 \ldots dz_N dp_{x,1} dp_{y,1} \ldots dp_{z,N}.$$

It is a 6N dimensional space.

❑ For an N particle system, the volume of the total phase space is :

$$\Lambda = \int dp_{1,x}.dp_{1,y}.dp_{1,z}.dp_{2,x} \ldots dp_{N,z}.dx_1.dy_1.dz_1 \ldots dz_N$$

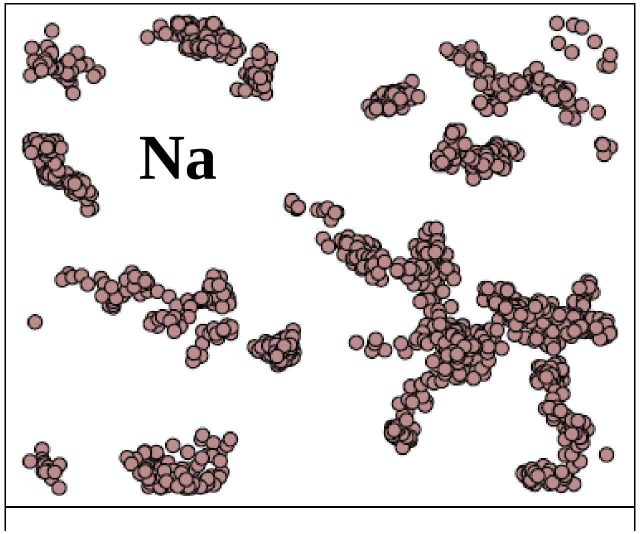
- ❑ Counting the number of states available to a particle =determining the available volume in phase space.

- ❑ Given the Heisenberg's uncertainty principle ($\Delta p \Delta r = h$), the smallest "cell" in phase space is $dV_{min} = h^3$ with h the Planck constant $h = 6.62.10^{-34}$ J.s

- ❑ The time evolution of a system can be described as a trajectory through phase space.

- ❑ Systems that are ergodic and able to reach equilibrium are able to explore all parts of phase space that fulfill E=cst. The trajectory will visit all points in phase space with constant energy hypersurface with equal probability.

- ❑ **Non-ergodic** systems have parts of the phase space that are inaccessible. **This is the case for glasses.** This means that glasses are not able to explore the complete E=cst surface during the timescale of interest.

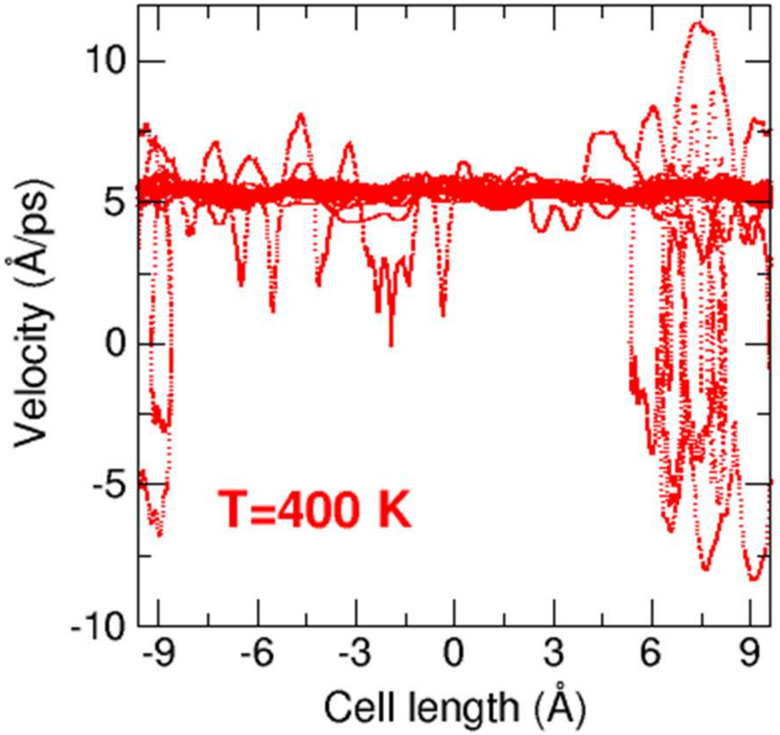


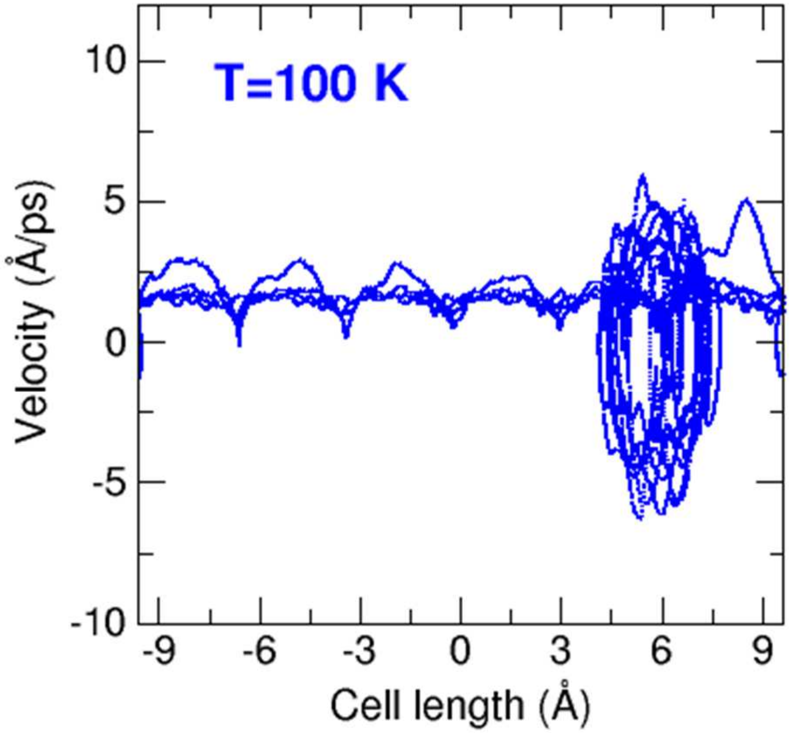
SiO_2-$2Na_2O$, 1500 K, 20ps

PRB 83, 184118 (2011)

Phase space of water at high and low temperature

1024 H_2O molecular system using a SPC model
Phase spaced explore by an oxygen atom





More on statistical mechanics and phase space in the additional notes, see web Page

Molecular Dynamics Simulations

Advantages of Molecular Dynamics:

- ☐ the only input in the model – description of interatomic or intermolecular interaction
- ☐ no assumptions are made about the processes/mechanism to be investigated
- ☐ provides a detailed molecular/atomic-level information
- ☐ **Results of the “*computational experiment*” may lead to the discover new physics/mechanisms!**

Limitations…we’ll see later on

Current applications of MD simulations

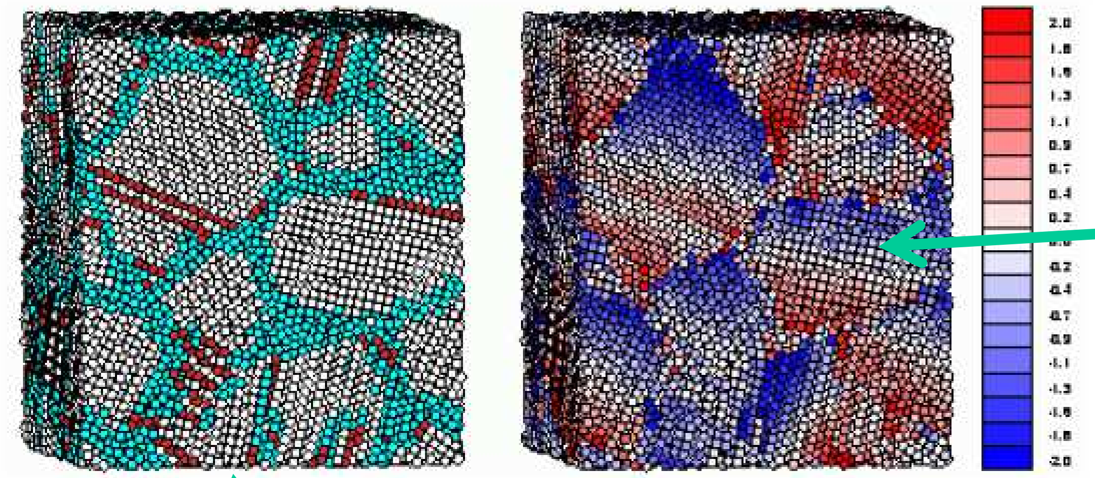
Wide range of problems in different research fields

- ☐ **Chemistry and biochemistry**: molecular structures, reactions, drug design, vibrational relaxation and energy transfer, structure of membranes, dynamics of large biomolecules,…
- ☐ **Statistical mechanics and physics:** theory of liquids, correlated many-body motion, **liquid-to glass transition**, phase transitions, structure and properties of small clusters
- ☐ **Materials science:** point, linear and planar defects in crystals, microscopic mechanisms of fracture, extreme conditions, melting, **glass properties**,…

matthieu.micoulaut@upmc.fr

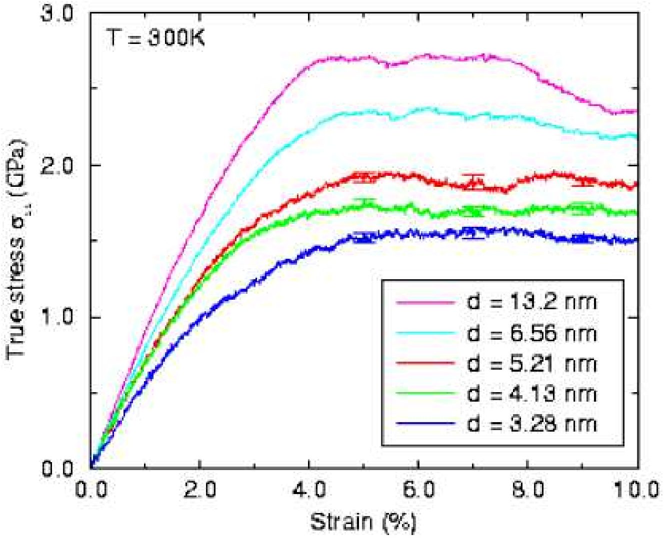
A few examples …

1 Deformation of nanocrystalline copper (100, 000 atoms after 10% deformation): interaction between dislocation and grain boundaries in metals



Atoms colored relative to their homogeneous deformation. **Blue atoms** have moved downwards, **red atoms** upwards. Major part of the deformation is in grain boundaries

Atoms in the grain boundaries are blue. Atoms in stacking faults are red. One sees stacking faults left behind the partial dislocations

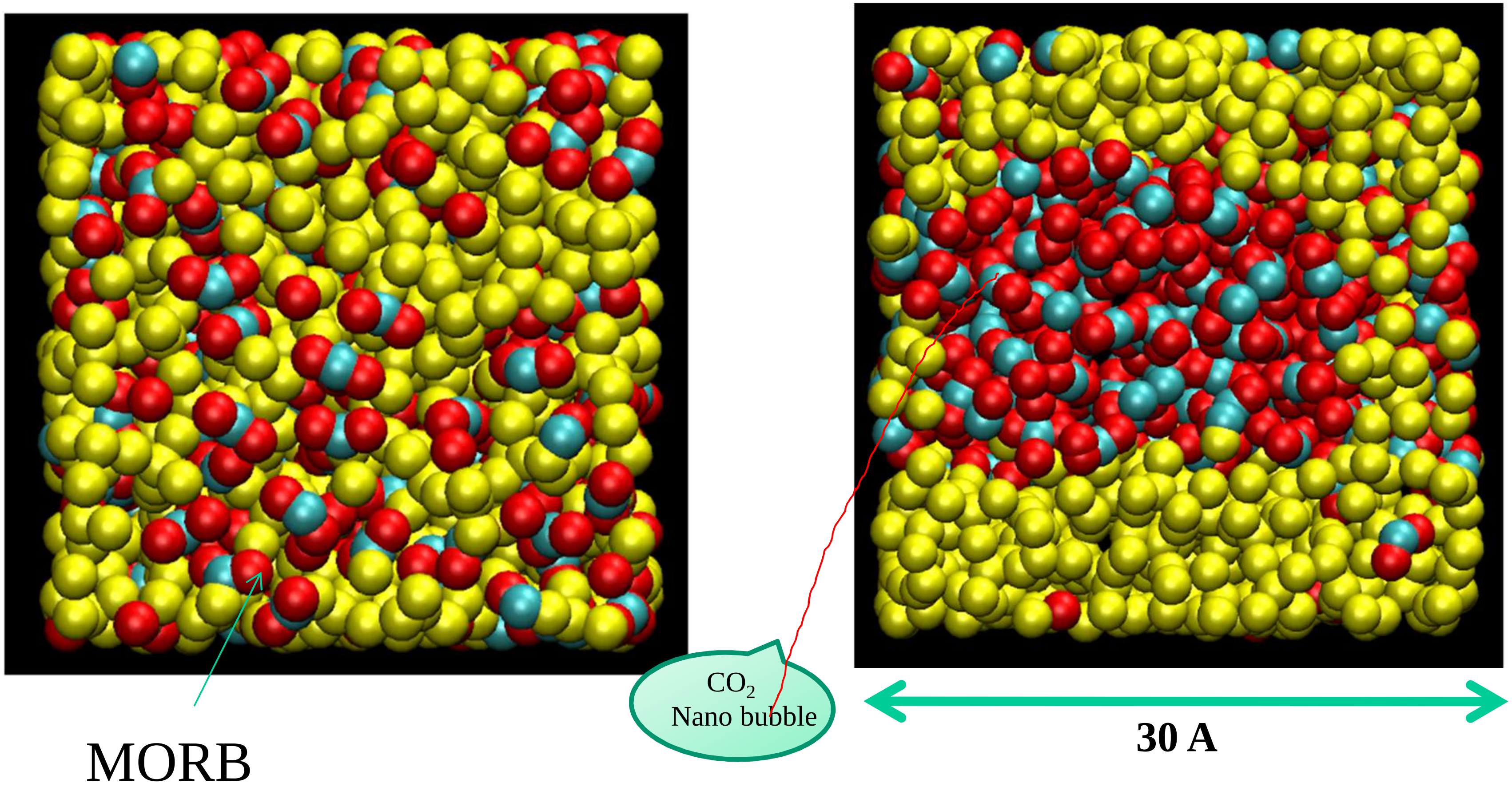


Average stress versus strain for each grain size.

Schiotz et al. PRB 2001

A few examples …

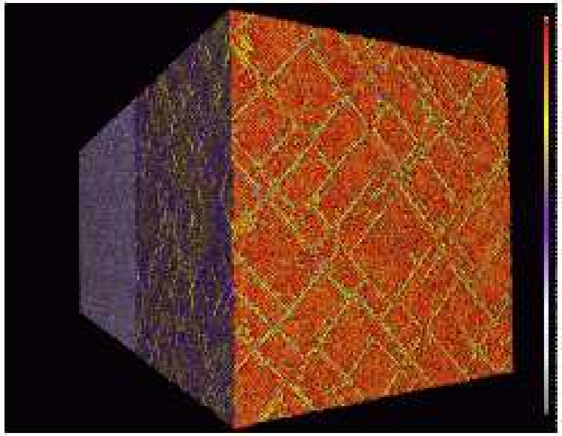
2 Bubble growth in a basaltic MORB melt (10, 000 atoms relaxed from 100 kbar to 20 kbar)



CO_2 saturated melt at 1873K and 100 kbar

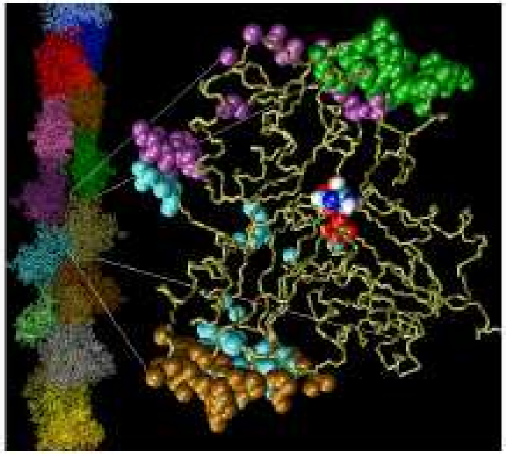
after pressure drop to 20 kbar

A few more examples …



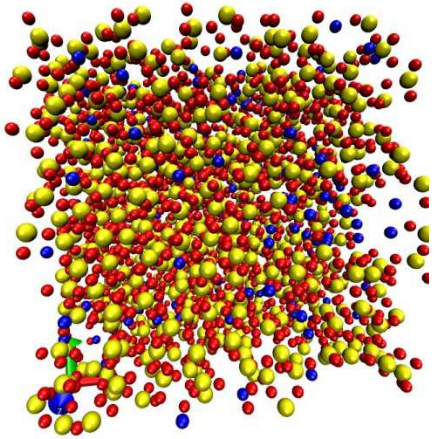
Shock-wave induced plasticity

Holian, Lomdahl, Science 1998



Actin filaments

W. Wriggers, U. Illinois



Pressure induced rigidity in glasses

Bauchy et al. PRL 2013

Molecular Dynamics Simulations : The idea

- Very similar to an experiment
  (**Numerical experiment**)

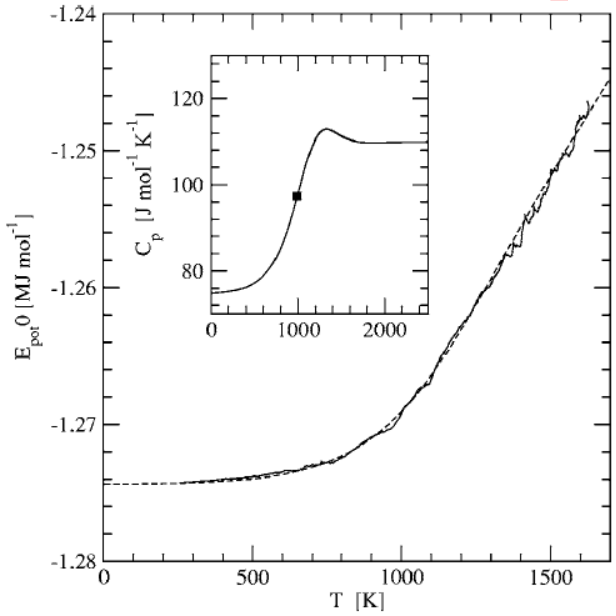
☐ Preparation of a sample of a given material

☐ Connect the sample to a measuring instrument.

☐ Measure a property of interest over a certain time interval

☐ In the case of glasses: melting of a crystal, and quench to a glass. Thermal properties behave like those determined in calorimetry.

MD-GeO_2



PRE 73, 031504 (2006)

1) Prepare a system of N particles of a given material and solve Newton's equation of motion until the properties do not change any more with time (**thermal equilibrium**).

2) Once equilibrated, one computes system properties.

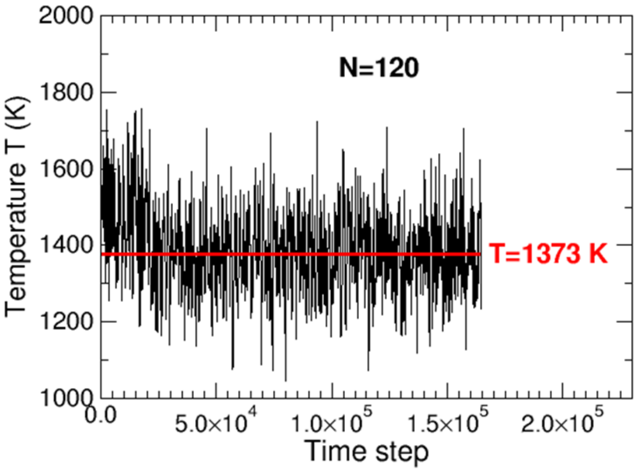
- Determined properties are averages over **the system** (i.e. the N particles) and **the measuring time**

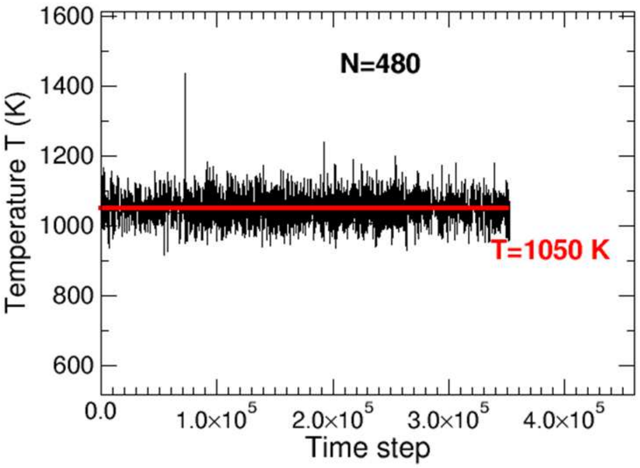
- Remember equipartition theorem (kinetic theory of gases). Per N_f degrees of freedom, one has :

$$\left\langle \frac{1}{2} m {v_i}^2 \right\rangle = \frac{1}{2} k_B T$$

Leads to a natural microscopic definition of the temperature :
where $N_f=3N-3$.
Fluctuations of the temperature are of $1/\sqrt{N_f}$

$$\boxed{\sum_{i=1}^{N} \frac{m_i}{N_f k_B} {v_i}^2 = T(t)}$$





l-$GeSe_2$

Molecular Dynamics Simulations : Structure of the program

The general structure remains always the same.

1) We read the parameters specifying the conditions of the run : N, T, V, time step,… **datinput**

2) We initialize the system (initial $\mathbf{r}_i$ and $\mathbf{v}_i$). **initial**

3) We compute the forces (**f**) from the interaction potential (force field, **en**). **forces**

4) We integrate the Newton's equations of motion (core of the simulation) and update the trajectories. **integrate**

5) We compute various quantities by averaging **stats**

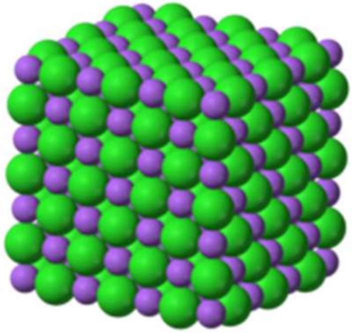
```
Program md

call datinput
call initial
t=0
do while (t.lt.tmax)
        call forces(f,en)
        call integrate(f,en)
        t=t+delt
        call stats
enddo
stop
end
```

Molecular Dynamics Simulations : Structure of the program

A. Initialization :

Initial positions: several options

- Generate the corresponding crystal (e.g. a-quartz for SiO_2) and melt it at HT
- Setup an « *easy* » pseudo-crystal (unphysical crystal), e.g. NaCl, cubic lattice
- Put positions at random (sometimes dangerous)
- Request a stable configuration from your colleague and relabel the atoms.

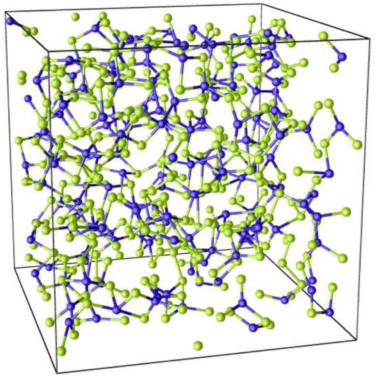


Initial velocities : standard strategy

- Setup random velocities
- Compute the velocity center of mass and set it to zero
- Rescale the velocities by a factor $\sqrt{T_{desired}/T(t)}$

to reach the desired temperature, as one has :

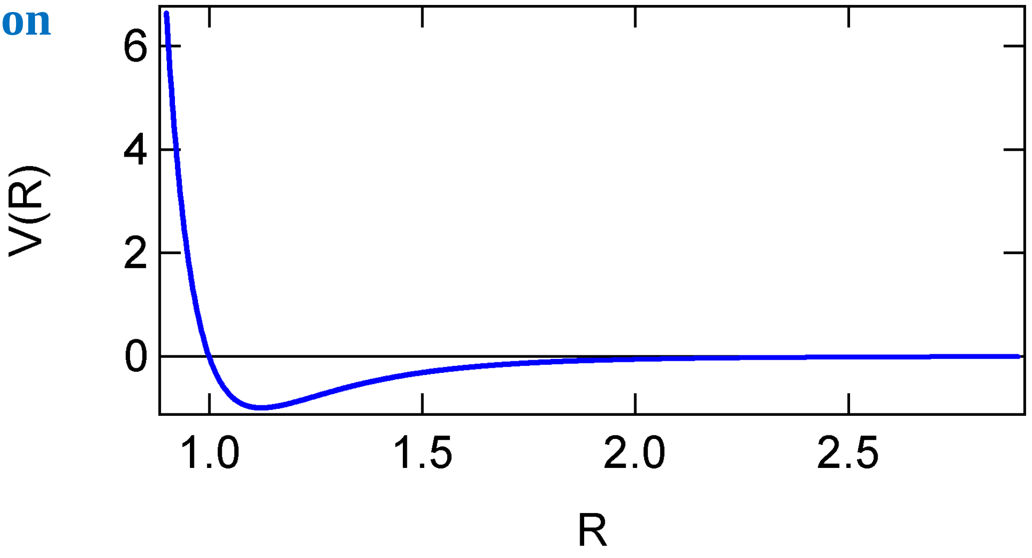
$$\sum_{i=1}^{N} \frac{m_i}{N_f k_B} {v_i}^2 = T\,(t)$$



Not critical given that the position update does not depend on the velocities (see below).

Molecular Dynamics Simulations : Structure of the program

B) Force calculation



1) Starting point is the interaction potential V(R). e.g. the Lennard-Jones potential

All the model approximations are contained in the parameters of V(R) !

2) Calculation of the force acting on every particle.

For the x-component of the force, we have :

$$F_x(R) = -\frac{\partial V(R)}{\partial x} = -\left(\frac{x}{R}\right)\frac{\partial V(R)}{\partial R}$$

Molecular Dynamics Simulations : Structure of the program

C) Integrating the equations of motion. The Verlet algorithm

Having the interactions (forces), we want to solve for any particle Netwton's law:

$$m\frac{d^2\mathbf{r}}{dt^2} = \mathbf{f}(\mathbf{r}(t)).$$

which can be Taylor expanded as a time series to :

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{\mathbf{f}(\mathbf{r}(t))}{2m}(\Delta t)^2 + \frac{d^3\mathbf{r}}{dt^3}(\Delta t)^3 + \mathcal{O}((\Delta t)^4)$$

$$\mathbf{r}(t-\Delta t) = \mathbf{r}(t) - \mathbf{v}(t)\Delta t + \frac{\mathbf{f}(\mathbf{r}(t))}{2m}(\Delta t)^2 - \frac{d^3\mathbf{r}}{dt^3}(\Delta t)^3 + \mathcal{O}((\Delta t)^4).$$

adding the two equations leads to :

$$\boxed{\mathbf{r}(t+\Delta t) + \mathbf{r}(t-\Delta t) = 2\mathbf{r}(t) + \frac{\mathbf{f}(\mathbf{r}(t))}{m}(\Delta t)^2 + \mathcal{O}((\Delta t)^4).}$$

which is of accuracy $(\Delta t)^4$.

It has to be realized that there is no need of the velocities for the position update.

But the velocities can also be computed :

$$\mathbf{v}(t) = \frac{\mathbf{r}(t+\Delta t) - \mathbf{r}(t-\Delta t)}{2\Delta t} + \mathcal{O}((\Delta t)^2)$$

Comments :

- Most of the time is spent in computing the forces f(t).
- A slow variation of f(t) is needed.
  Energy conservation at short times.
  Determines the time step Δt.
- The Verlet algorithm is time-reversing (Δt changed into – Δt).
  Allows to retrace MD trajectories
-

An alternative integration scheme: the Leapfrog algorithm

Velocities are calculated at half-timestep Δt/2, to yield :

$$\mathbf{v}(t+\Delta t/2) = \frac{\mathbf{r}(t+\Delta t) - \mathbf{r}(t)}{\Delta t},$$
$$\mathbf{v}(t-\Delta t/2) = \frac{\mathbf{r}(t) - \mathbf{r}(t-\Delta t)}{\Delta t},$$

$$\mathbf{v}(t+\Delta t/2) = \frac{\mathbf{r}(t+\Delta t)-\mathbf{r}(t)}{\Delta t},$$
$$\mathbf{v}(t-\Delta t/2) = \frac{\mathbf{r}(t)-\mathbf{r}(t-\Delta t)}{\Delta t},$$

adding the two equations leads to :

$$\mathbf{r}(t-\Delta t) = \mathbf{r}(t) - \mathbf{v}(t-\Delta t/2)\Delta t.$$
$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t+\Delta t/2)\Delta t$$

By using again: $\mathbf{v}(t+\Delta t/2) = \mathbf{v}(t-\Delta t/2) + \frac{\mathbf{f}(t)}{m}\Delta t + \mathcal{O}((\Delta t)^3).\ \mathcal{O}((\Delta t)^4).$

one obtains: $\boxed{\mathbf{v}(t+\Delta t/2) = \mathbf{v}(t-\Delta t/2) + \frac{\mathbf{f}(t)}{m}\Delta t + \mathcal{O}((\Delta t)^3).}$

Trajectories are identical to the Verlet algorithm.

The potential energy (function of ***r**(t)*) and the kinetic energy (function of ***v**(t)*) may differ given that these are computed at different half time-steps.

The general strategy is therefore:

1) Given r(t) and v(t) at one time point t, first compute the forces on each atom using the force field V(R).

2) Do an update of the positions $r \leftarrow r + v\delta t + \frac{f}{2m}\delta t^2.$

3) Do a partial update of the velocity array based on the current forces

$$v \leftarrow v + \frac{f}{2m}\delta t$$

4) Compute the new forces f(t+δt) using the new positions $r(t + \delta t).$

5) Finish the update of the velocity array

$$v \leftarrow v + \frac{f}{2m}\delta t$$

6) Go back to 1)

matthieu.micoulaut@upmc.fr

C) Integrating the equations of motion: On the choice of the timestep

- ❑ Netwon's equations conserve the total energy E, and therefore the numerical solutions of the system should behave the same way.

- ❑ Accuracy of the numerical solutions can be checked by the extent to which the energ remains constant.

- ❑ Two kinds of energy conservation:

-short term: usual fluctuations in total energy from step to step, around the mean value.
-long term: over large number of time steps, the energy can drift away from its (equilibrium) initial value.

- ❑ Choosing the timestep: for numerical stability and accuracy in conserving the energy, one typically needs to pick a time step that is at least one order of magnitude lower than the fastest time scale in the system.

Bond vibrations: 10^{-13}s=100fs MD time step: 0.5-1.0 fs

Choosing the timestep

- ☐ Gaining time: one does not want to spend useless computer time by having too small timesteps if not necessary.
- ☐ Time step limits the length of the MD trajectory and the simulation.
- ☐ Too large timesteps can cause a MD simulation to become unstable with E increasing rapidly with time.
- ☐ **Exploding :** large time steps propagate the positions of 2 atoms to be nearly overlapping

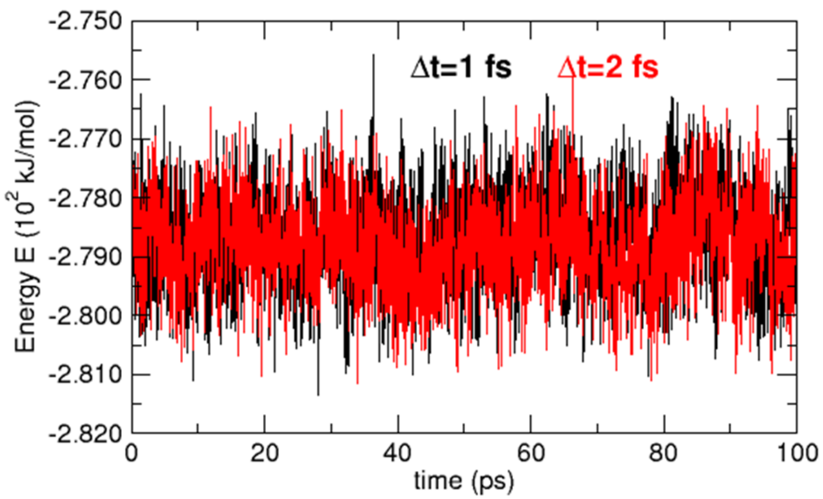
$$r \leftarrow r + v\delta t + \frac{f}{2m}\delta t^2.$$

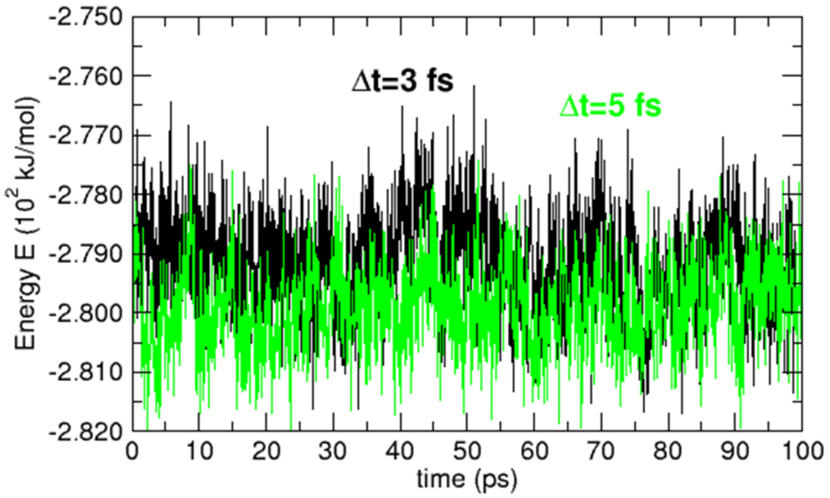
- ☐ **Devastating atomic collisions !!** The repulsive forces then create a strong force that propels these atoms apart.

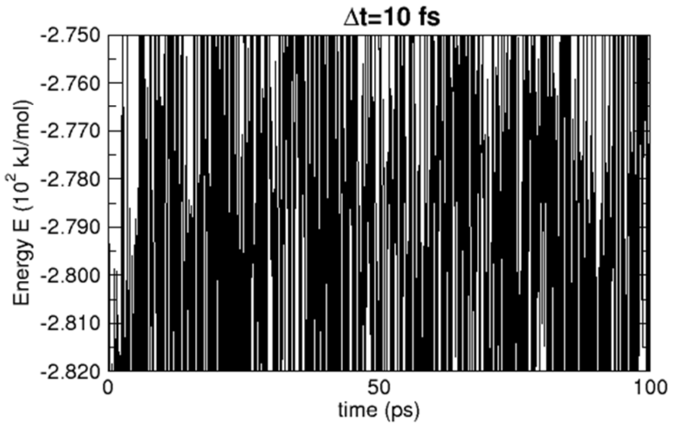
$$v \leftarrow v + \frac{f}{2m}\delta t$$

Choosing the timestep : Effect of the timestep on the energy :

Energy of a liquid GeO_2 at T=2000K and P=0 kbar over a trajectory of 100 ps=10^{-10}s



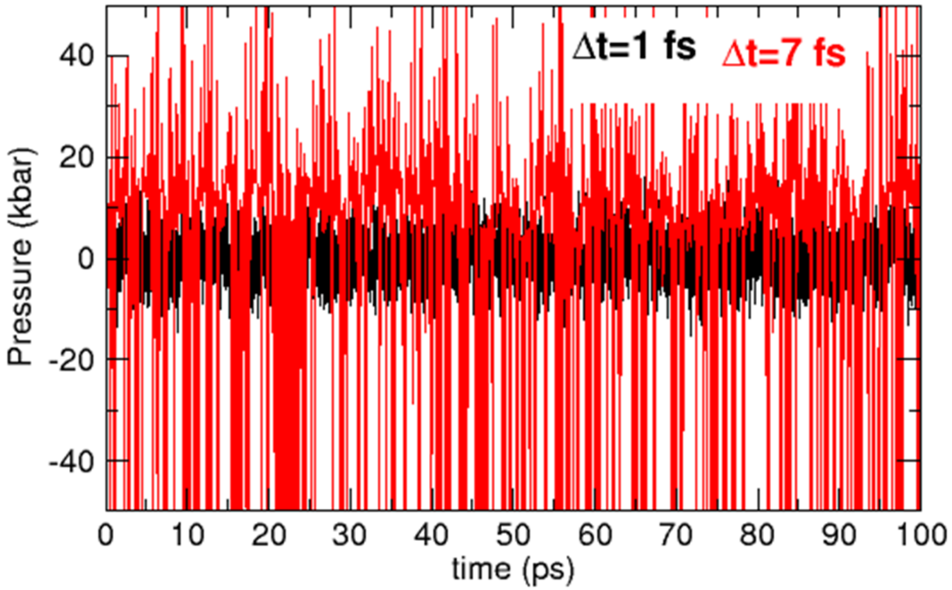




Conclusion : a good (and still safe) choice would be 2 fs.

Choosing the timestep : Effect of the timestep on the calculated pressure :

Energy of a liquid GeO_2 at T=2000K and P=0 kbar

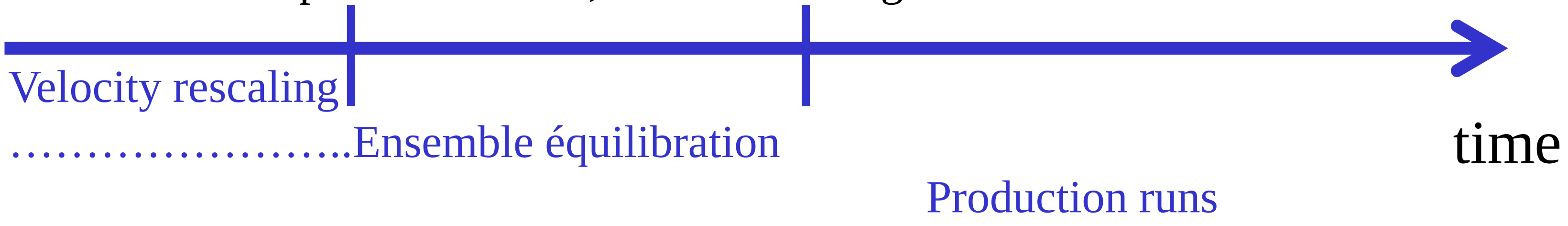


Increasing δt causes large fluctuations in the calculated pressure as it is directly related to the force f(t) (Virial theorem)

Molecular Dynamics Simulations : Structure of the program

D) NVT Ensembles

General idea: we want to put a thermodynamic system of N particles in contact with a thermostat or a barostat.
Performed before production runs, at an initial stage.



Velocity rescaling

……………………..Ensemble équilibration

time

Production runs

Velocity scaling: Remember that the kinetic temperature is given by

$$\boxed{\sum_{i=1}^{N} \frac{m_i}{N_f k_B} v_i^{\ 2} = T\,(t)}$$

One cas thus rewrite the velocities such as :

$$v^{new} = v^{old}\sqrt{\frac{T_{desired}}{T(t)}}$$

- No time-reversible, results do not correspond to any Ensemble. Not recommended for production runs.

Berendsen algorithm: To maintain the temperature, the system is coupled to an external heat bath with fixed temperature T_0.
Velocities are scaled at each step, such that the rate of change of temperature is proportional to the difference in temperature:

$$\frac{dT(t)}{dt} = \frac{1}{\tau}(T_0 - T(t))$$

τ being the rise time (strength of the coupling to a hypothetical heat bath).

- ❑ No time-reversible, results do not correspond to any Ensemble. Not recommended for production runs.

Andersen algorithm: Coupling of the system with a bath is performed through a stochastic process. Modifies velocities of particles by the presence of instantaneous forces.

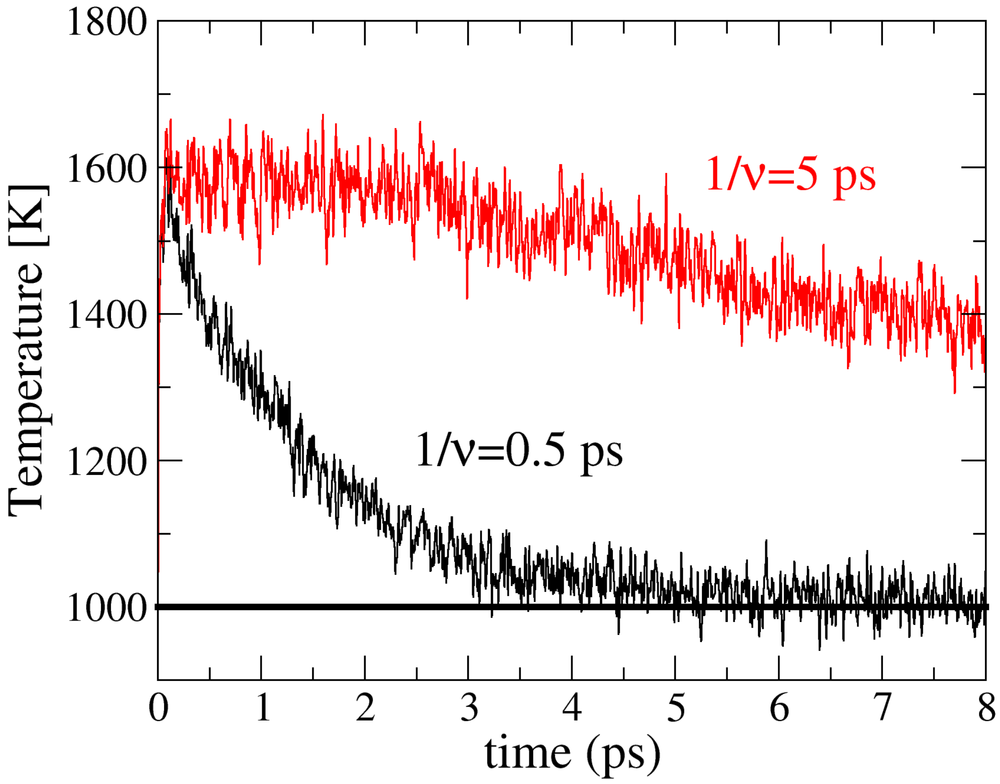
- ❑ Between these stochastic "collisions", the system evolves with the usual Newtonian dynamics.
- ❑ Coupling strength is controlled by a collision frequency denoted by ν.
- ❑ Stochastic collisions are assumed to be totally uncorrelated. This leads to a Poissonian distribution of collisions P(ν,t):

$$P(\nu,t) = e^{-\nu t}$$

Produces a canonical distribution

Effect of the rise time in a Berendsen thermostat

At fixed density, changing a **GeO_2** liquid from T=1600K to T=1000 K.
2 different rise times 1/ν



<u>Nosé-Hoover algorithm:</u> One adds a new internal variable **s** (a time scaling variable) to the equation of motion.

☐ s is associated with a "mass" Q>0 and a velocity ds/dt which acts as a friction coefficient able to change the particle velocities (and thus temperature).

The Lagrangian of the system is given by:

Kinetic E

Potential E

$$\mathcal{L} = \sum_{i=1}^{N} \frac{m_i s^2}{2} \left(\frac{d\mathbf{r_i}}{dt} \right)^2 - U(\mathbf{r}^N) + \frac{Q}{2} \left(\frac{ds}{dt} \right)^2 - \frac{L}{\beta} \ln(s)$$

and using the Lagrange equations:

$$\mathbf{p}_i = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_i} = m_i s^2 \dot{\mathbf{r}}_i$$

$$p_s = \frac{\partial \mathcal{L}}{\partial \dot{s}} = Q\dot{s}$$

we obtain the Hamiltonian

$$\mathcal{H} = \underbrace{\sum_{i=1}^{N} \frac{(\mathbf{p}_i)^2}{2m_i s^2} + U(\mathbf{r}^N)}_{\text{Physical system}} + \underbrace{\frac{p_s^2}{2Q} + \frac{L}{\beta} \ln(s)}_{\text{Thermostat}}$$

Physical system

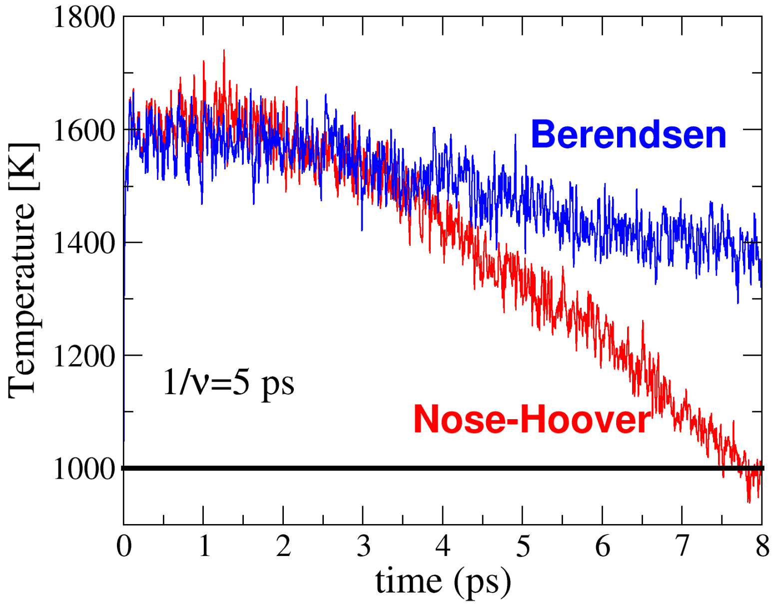
Thermostat

$$\mathcal{H} = \sum_{i=1}^{N} \frac{(\mathbf{p}_i)^2}{2m_i s^2} + U(\mathbf{r}^N) + \frac{p_s^2}{2Q} + \frac{L}{\beta}\ln(s)$$

- ❑ The logarithmic term (ln s) is needed to have the correct scaling of time.
  virtual time Δt'=Δt/s and v=sv' or p=sp'

- ❑ The magnitude of Q determines the coupling between the reservoir and the real system.

- ❑ Compatible with Canonical Ensemble



Conclusion

- ☐ MD simulation = solving Newton's equation of motion for N particle in interaction (V(r))
- ☐ Output: Position **r** and velocities **v** of each particle
- ☐ Statistical mechanics treatment of (**r**,m**v**)=(**r**,**p**). Notion of phase space for statistical averages

Next lecture Space correlations

Additional reading: Notes on Ensembles (Web page). You will need these notes for lecture 5