

Statistical Mechanics and Applications on Simulations



Practical session 2

Calculation of a surface tensión (L/V) and a interfacial tensión (L/L) from molecular dynamics simulations:

- n-hexane (ST)
- n-hexane/water (IFT)

Ramón Sayós and Gerard Alonso

1



Interfacial or surface tension (IFT or ST, γ) arises from unbalanced attractive forces between two different phases. This leads to an excess of free energy at interface (i.e., surface free energy). Formally, γ is the force per unit length that opposes the growth of the surface area. For a system at constant pressure (or volume), temperature and composition:

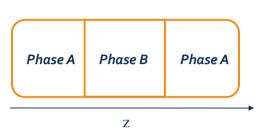
$$\gamma = \left(\frac{\partial G}{\partial \mathcal{A}}\right)_{N,P,T} = \left(\frac{\partial A}{\partial \mathcal{A}}\right)_{N,V,T}$$

where G and A are the Gibbs and Helmholtz free energies, respectively, and A is the surface area. The common units for IFT are dynes/cm or mN/m.

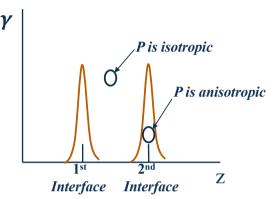


 γ can be calculated using the Kirkwood-Buff method (i.e., pressure tensor method), based on that this tensor is diagonal for a planar interface in an inhomogeneous fluid at hydrostatic equilibrium, being the independent components (i.e., normal P_n and tangential P_t ones) a function of only z:

$$\gamma = \int_{-\infty}^{\infty} dz [P_n(z) - P_t(z)] = \frac{L_z}{2} \left(P_{zz} - \frac{P_{XX} + P_{YY}}{2} \right) (2 interfaces)$$



 $\langle \gamma \rangle = \frac{1}{2} \left(\langle P_{zz} \rangle - \frac{\langle P_{xx} \rangle + \langle P_{yy} \rangle}{2} \right) \langle L_z \rangle$



3



$$P_{\alpha\beta} = \frac{1}{V} \left[\sum_{i=1}^{N} m_i v_{i\alpha} v_{i\beta} + \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} F_{ij\alpha} r_{ij\beta} \right]$$

$$x, y \text{ or } z \qquad \qquad forces \qquad distances$$

$$between \qquad between$$

$$velocities \qquad i \text{ and } j \quad i \text{ and } j$$

Л

Calculation of N-hexane surface tension



Computational method:

- Molecular Dynamics (numerical integration of Newton equations)
- TraPPE force field (which fits exp. phase equilibrium data):

$$\begin{split} &U_{total} = \sum\limits_{i=1}^{N}\sum\limits_{j>i}^{N}U_{ij}, \qquad U_{ij} = 4\epsilon_{ij} \Bigg[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \Bigg] + \frac{1}{4\pi\epsilon_{o}}\frac{q_{i}q_{j}}{r_{ij}} + \\ &+ K_{ij}\Big(r_{ij-}r_{ij}^{o}\Big) + K'_{ij}\Big(\theta_{ij-}\theta_{ij}^{o}\Big) + ... \end{split}$$

• TraPPE-UA (united-atom) representations for alkyl groups (i.e., hydrogen atoms are modelled implicitly along with the carbon atoms they are bonded to in a single pseudoatom type:

$$H_3C$$
 CH_3
 $C_6H_{14} \Rightarrow C_2 C_4$

5



- LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator) is an open-source classical molecular dynamics code
- Orthorhombic simulation cell: 30 Å x 30 Å x 90 Å (1/3 hexane + 2/3 vacuum):



$$N_{molec} = \frac{L_{x} \cdot L_{y} \cdot L_{z}(\text{Å}^{3})}{3} \cdot 10^{-24} \cdot \rho_{L}(\text{g/cm}^{3}) \cdot \frac{1}{M(\text{g/mol})} \cdot 6.023 \text{x} 10^{23}$$

$$(\rho_{L}(T,P))$$

- Firstly, there will be a minimization using a soft potential and later with TraPPE force field
- Canonical ensemble (NVT) $(T = 25^{\circ} C)$; Nose-Hoover thermostat

10011101C 10101011C 11101C 1010111101 1000001C Institut de Quinica Teorica 1000001C

- Inputs:

run.in: read all simulation parameters (i.e., L_x , L_y , L_z , T, steps ...).

hexane.in: contains the structure of a single hexane molecule (read by run.in).

potential.ff: contains the force field parameters (read by run.in).

- Outputs:

lammps.out: shows the simulation parameters and the different calculated properties at several timesteps.

lammps.ave: gives the total energy and IFT at several timesteps.

lammps.xyz: gives the geometry of the system at different timesteps (can be open with AVOGADRO in QUBUNTU).

(log.lammps, err.lammps, out.lammps)

7



- Execution:

A) Using QUBUNTU OS (LAMMPS_2017):

- open a terminal
- create directories /PS2/ST and /PS2/IFT
- cd to the folder /PS2/ST
- download the files from UAM (Moodle) or from cerqt2.qt.ub.es:
 e.g., scp your_user@cerqt2.qt.ub.edu:/work/MEAS/PS2/ST/*.
- lammps < run.in > lammps.out & (execute in background)



- B) Using an external cluster, IQTCUB (LAMMPS_2015):
 - Connect to the cluster (ssh): ssh -l -X your user cerqt2.qt.ub.edu (or ssh -l -Y)
 - Make several directories in your home:
 mkdir PS2, mkdir./PS2/ST, mkdir./PS2/IFT
 - Copy input files and scripts from /work/MEAS/PS2/ST and /work/MEAS/PS2/IFT into your directories ST and IFT: cd PS2/ST

cp /work/MEAS/PS2/ST/*. (the same for IFT)

Send your calculation with LAMMPS to a batch queue:

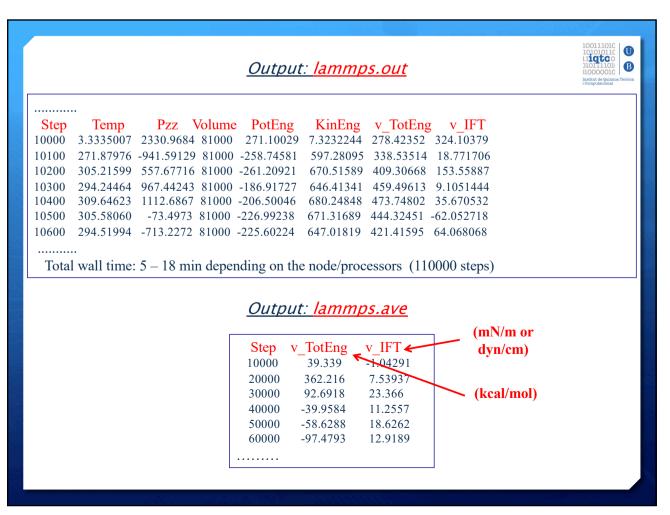
```
qsub -N name lammps_sub (put a name for the calculation)
qstat (show the status of your job)
```

9

Input: run.in #==Simulation Settings==# units real # Define units in Å, fs, kcal/mol, K, atm boundary p p p # The simulation cell is periodic in all directions of the space 3 # The simulation cell is 3-Dimensional dimension atom style full # Define the system as a molecular set (atoms with bonds or no-bonds, etc) =Simulation Variables= variable Cutoff # Cutoff distance for intermolecular interactions (Å) equal 12 variable T equal 298.15 # Temperature (K) timestep 2 # Timestep (fs) thermo # Print thermodynamic output each 100 steps #== Box Parameters and Molecular PPTs ==# variable X equal 30.0 # x variable equal 30.0 # y variable equal 90.0 # z dimension length (Å) \mathbf{Z} variable Rho1 # Density of hexane at 298.15 K (g/cm3) equal 0.65478 variable PM1 equal 86.1756 # Molar mass of hexane (g/mol) #-----# molecule mol1 hexane.in # Load a molecule file # Calculate the number of molecules to add in the sym box as a function of Rho, PM and # XYZ cell parameters. This variable fills 1/3 of the simulation cell (1 var per molec) variable N1 equal $floor(\{X\}^*\{Y\}^*(\{Z\}/3)/1e24^*6.023E23^*\{Rho1\}/\{PM1\})$ = Box Creation = lattice custom 1. a1 \$X 0. 0. a2 0. \$Y 0. a3 0. 0. \$Z basis 0. 0. 0.

```
region sym block 0 1 0 1 0 1 units lattice
                                                                                                B
create box 4 sym bond/types 2 angle/types 2 dihedral/types 1 extra/bond/per/atom 10
            extra/angle/per/atom 10 extra/dihedral/per/atom 40 extra/special/per/atom 25
                   = Region Definition & Atom Creation=
                                                                # cubic region
 region Liquid block 0 1 0 1 0 0.33 units lattice
 create atoms 0 random ${N1} 298437 Liquid mol mol1 301709 # add molecs to liquid region: seeds
                      OUTPUT VARIABLES=
 variable IFT equal (-((zhi-zlo)/2.0)*(((pxx+pyy)/2.0)-pzz))*0.0101325 # Calculate IFT (mN/m)
 variable
            TotEng equal etotal
                                                                        # Calculate Etotal
 fix averages all ave/time 1 10000 10000 v TotEng v IFT file lammps.ave # Write file lammps.ave
 dump trj all xyz 10000 lammps.xyz
                                          # Write traj. each 10000 timesteps
 thermo style custom step temp pzz vol pe ke etotal v IFT # What to print (you may add new keys)
     ==== Minimization Process With Soft Potential at 10 K and rigid molecules=====#
include potential.ff
                               # Load the whole FF
 pair style soft 3
                                # Change the LJ pair by Soft
 pair coeff * * 1
                                # Soft parameters
 fix label1 all rigid/nvt molecule temp 10 10 200
                                                     # Time evolution scheme (10 K)
 run 10000
                                                     # run during 10000 steps
                                                     # Erase time evolution scheme
 unfix label1
#===== Simulation with appropriate Potential and non-rigid molecules=====#
                                          # Load the whole FF
 include potential.ff
 fix label 1 all nvt temp $T $T 200
                                          # Time evolution scheme (298.15 K)
 run 100000
                                          # run during 100000 steps
 unfix label1
                                          # Erase time evolution scheme
```

11



Tasks to do about ST:



- 1. $T = 25 \,^{\circ} \, C$, 100.000 steps (200 ps)
 - Prepare lammps.out file to be plotted:

cp lammps.out lammps.data

- Remove all file lines keeping from step 10.100 to 110.000:

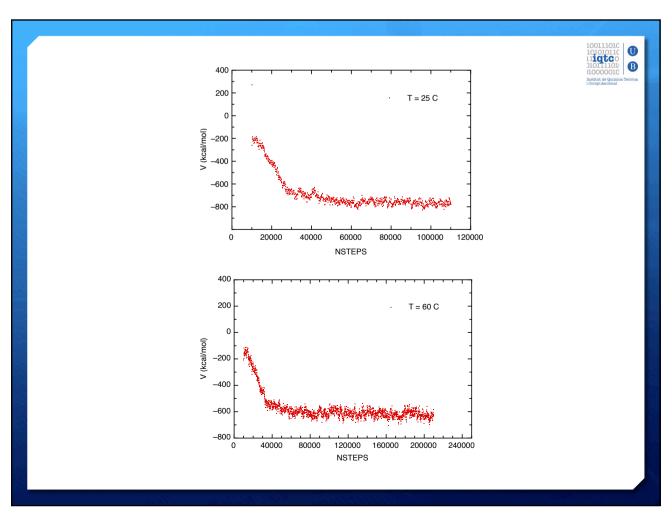
vi lammps.data

- put the cursor over the line (second time) with the column beginning with 10000 and write dgg
- put the cursor after the line beginning with 110000 and write dG
- write :wg (save the file)
- Plot several variables (T, Pe, Ke, Te, γ) every 100 steps:
 gnuplot (open previously a terminal and write gnuplot)
 plot "lammps.data" u 1:2 w lp (plot T vs. time)
 plot "lammps.data" u 1:5 w lp, "plot" u 1:6 w lp, "plot" u 1:7 w lp
 (plot Pe, Ke, Te vs. time)

When the system is equilibrated?

- Plot y vs. time: plot "lammps.data" u 1:8 w lp Why are too oscillations in y?

13





- Average the variables of file lammps.data:
 - Open file with LibreOffice Calc (with space + merge delimiters)
 - Average each column (variable) using AVERAGE function
 - Check the Equipartition Principle using the kinetic energy:

$$K = N \cdot \frac{3}{2} k_b T$$
 ($N = N_{molec} \cdot 6$)

- Plot the average variables of lammps.ave file:
 - Delete the first two lines and save the file
 - Write gnuplot
 plot "lammps.ave" u 1:2 w lp (plot Te vs. time)
 plot "lammps.ave" u 1:3 w lp (plot y vs. time)
- Open lammps.ave file with LibreOffice Calc:
 - Average y column using AVERAGE function
 - Compare calculated y with the experimental value

15



In file run.in:

fix averages all ave/time 1 10000 10000 v_TotEng v_IFT file lammps.ave

- · ave/time = style name of this fix command
- Nevery = use input values every this many timesteps
- · Nrepeat = number of times to use input values for calculating averages
- · Nfreq = calculate averages every this many timesteps

As example: Nevery=2, Nrepeat=6 and Nfreq=100

Values at steps 90, 92, 94, 96, 98 and 100 are used at step 100

Values at steps 190, 192, 194, 196, 198 i 200 at step 200
......

Change to: ave/time 10 1000 10000 in ST calculation and compare the results (i.e., averages); use also new random seed numbers

Other tasks about ST calculation:



- 2. Do a calculation with $T = 25 \,^{\circ}$ C, 200.000 steps, 400 ps (better average); you can also use the long_run_ST.ave file (2000000 steps, 4 ns).
- 3. Do a calculation with $T = 60 \,^{\circ}$ C, 200.000 steps (temperature effect on γ).
- 4. Visualize the lammps.xyz files with AVOGADRO at several steps. At the end some molecules appear in the vapour/gas phase.
- 5. Justify why we did not put initially n-hexane molecules in the vapour/gas box (i.e., 2/3 of total volume)

Experimental data (NIST web/Springer Materials) for n-hexane

T (° C)	ρ _L (g/cm³)	ρ_V (g/cm ³)	P _v (torr)	γ (dyn/cm)
25	0.65478	0.7132X10 ⁻³	151.45	17.90, 18.60
40	0.64083	1.2652X10 ⁻³	279.71	16.37, 16.46
60	0.62163	2.4523X10 ⁻³	571.80	14.32

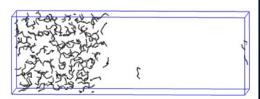
$$(T_{boil} = 68.75 \, ^{\circ} \, C)$$

17



- 6. Check the mass density (n-hexane) along the Z direction of the simulation cell:
 - Use Z_histo.f90 fortran code to generate the Z-distributions.
 - Compile and mount the executable for Z_histo.f90 code:
 - gfortran -o Z_histo.exe Z_histo.f90
 - Extract the last frame from the trajectory (lammps.xylplotz):
 - tail -n \$((\$(head -n1 lammps.xyz) +2)) lammps.xyz > lastep.xyz
 - Execute:
 - ./Z_histo.exe < lastep.xyz (creates oil.Z_histo file) *
 - Plot the file (density vs. Z distance):
 - gnuplot
 - plot "oil.Z_histo" w l

Check the density of n-hexane and the width and position of the interface



* creates also water.Z_histo file for IFT water/hexane calculations



Calculation of N-hexane/water interfacial tension

Computational method:

- Molecular Dynamics
- TraPPE force field for n-hexane and SPC force field for water
- Orthorhombic simulation cell: 30 Å x 30 Å x 90 Å
 (1/3 hexane + 1/3 water + 1/3 vacuum). Some of the space will be left without molecules to prevent an unphysical mixing of water and n-hexane during the relaxation with the artificial "soft" potential



• Isothermal-isobaric ensemble (NPT) $(T = 25^{\circ} C, P = 1 \text{ atm})$; Nose-Hoover thermostat and barostat

19

- Inputs:

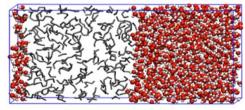


run.in: read all simulation parameters (i.e., L_x , L_y , L_z , T, P, Steps 1...).

hexane.in: contains the structure of a single hexane molecule (read by run.in).

water.in: contains the structure of a single water molecule (read by run.in).

potential.ff: contains the force field parameters (read by run.in).



- Tasks to do about IFT:
- 1. Calculation at $T = 25 \,^{\circ}$ C, P = 1 atm, 200.000 steps (200 ps).
- 2. Calculation at $T = 60 \,^{\circ}$ C, P = 1 atm, 200.000 steps.
- 3. Do the same type of plots and analysis as for ST (now oil.Z_histo and water.Z_histo files for two density plots:

plot "oil.Z_histo" w l, "water.Z_histo" w l



Experimental data (NIST web/JANAF/Springer) for water

T (° C)	ρ _L (g/cm³)	ρ _V (g/cm³)	P _v (torr)	γ (dyn/cm)
25	0.9971	2.3075X10 ⁻⁵	23.78	71.98
40	0.9922	5.1242X10 ⁻⁵	55-39	69.59
60	0.9832	1.3X10 ⁻⁴	149.41	66.24

Experimental data (*) for n-hexane/water

T (° C)	γ (dyn/cm)
25	50.38
40	48.92
60	47.29

* (J. Chem. Eng. Data 46 (2001) 1086

21



Additional slides

Statistical errors

Limited sampling of the Configurational Space

With the mean value we must provide also the standard deviation

From N configurations:

$$\langle A \rangle = \frac{1}{N} \sum A_i$$

$$\langle A \rangle = \frac{1}{N} \sum A_i$$
 $\sigma(A) = \sqrt{\frac{\sum (A_i - \langle A \rangle)^2}{N}}$

The standard deviation of the averaged value will be

$$\sigma(\langle A \rangle) = \frac{\sigma(A)}{\sqrt{M}}$$

M are independent configurations. M<N

In general consecutive configurations are highly correlated

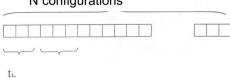
Relaxation time: number of steps required by the system to lose memory of previous configurations

It can be estimated previously to the simulation

23

Statistical errors

N configurations



The N configurations are split into n_b blocks with t_b steps each block The average value of the property is calculated in each block:

$$\langle A \rangle_b = \frac{1}{t_b} \sum_{i=1}^{t_b} A_i$$

$$\sigma^2(\langle A \rangle_b) = \frac{1}{n_b} \sum_{i=1}^{n_b} (\langle A \rangle_b - \langle A \rangle)^2$$

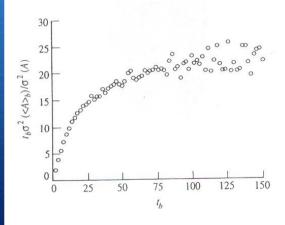
Is the variance of the block averages. Inversely proportional to t_b

Statistical errors



The statistical inefficiency or relaxation time is calculated as:

$$s = \lim_{t_b \to \infty} \frac{t_b \sigma^2(\langle A_b \rangle)}{\sigma^2(A)}$$



With N configuration, just M= N/s will be uncorrelated

The error in the averaged value when calculated from a finite simulation:

$$\sigma(\langle A \rangle) = \sigma(A) \sqrt{\frac{s}{N}}$$