

Evaluating Diffusion Coefficients in LAMMPS

Types of Diffusion Coefficients Diffusion

Diffusion occurs in various systems. Diffusion coefficients therefore take on different meanings depending upon the type of system they are describing. In this document, we examine two types of diffusion: (1) self-diffusion and (2) Fickian diffusion. The diffusion coefficients describing these processes are not the same but they are related.

Self-Diffusion

We often think of diffusion as a response to a concentration gradient. Remember that concentration of component A, c_A , is the product of the inverse of the molar volume, \bar{V} , and the mole fraction, x_A .

$$c_A = \frac{1}{\bar{V}} x_A \quad (1)$$

Thus a concentration gradient can be due to either a gradient in the molar volume (equivalently the density) or a gradient in composition or both. In a system at equilibrium, there is no concentration gradient. In a pure component system, the concept of a concentration gradient only exists if there is an external force, for example a pressure gradient, giving rise to a density gradient. In a multicomponent system at equilibrium, there is neither a composition gradient nor a density gradient.

Yet in a system at equilibrium, either single component or multicomponent, atoms and/or molecules (collected in the term particles) still experience Brownian motion. This motion is called self-diffusion and is described by a self-diffusion coefficient, D_{self} . Each particle experiences a random motion that allows it to move in space without any corresponding change in the averaged density or composition profile. In a single component system, there is one self-diffusion coefficient, D_{self} . In a multicomponent system, each species has its own self-diffusion coefficient, $D_{self,A}$, $D_{self,B}$, etc. The self-diffusion coefficient measured in molecular dynamics (MD) simulations is most like the isotopic tracer diffusion coefficient measured using, for example, Pulse-Field Gradient (PFG) Nuclear Magnetic Resonance (NMR). In this case, there is no gradient. While the self-diffusivity of only the isotope active to NMR is measured, it is generally assumed (especially for elements heavier than hydrogen) that the self-diffusivity of all isotopes is the same. Self-diffusion coefficients are relatively easy to get from an MD simulation. In part this is because, self-diffusion coefficients rely on a single-particle correlation function, as shown below. This means we can average our result over all particles of a given species to increase the statistical reliability of the result.

Fickian (or transport) Diffusion

In a multicomponent case, we often think of Fick's law as the definition of diffusion, just as we think of Fourier's law as the definition of heat conduction. Thus, Fickian diffusion describes a mixing process in which a non-uniform distribution of species becomes more uniform. From the point of view of thermodynamics, this is an entropy-generating process. The underlying driving force for this mixing process is again the Brownian motion of the particles, but the conventions by which diffusion coefficients are introduced to describe this process are completely different from self-diffusion. Often in an undergraduate course, Fick's law is presented as

$$J_A = -D_A \nabla c_A \quad (2)$$

Where J_A is a diffusive flux and D_A is the diffusion coefficient of A. Without further explanation, this is a meaningless and generic version of Fick's law. It convey an idea but cannot be unambiguously used

in a quantitative sense. Because MD simulations are intended to deliver quantitative results, we need to be much more careful than this.

Clearly, a Fickian diffusivity is not defined for a single component system. More subtly, multiple diffusion coefficients exist for even a binary system, D_{AB} , the diffusion coefficient of A relative to B, and D_{BA} , the diffusion coefficient of B relative to A. Unless one is careful, these diffusion coefficients are not the same. We will discuss this further below.

Functionality of Diffusion Coefficients

Like any thermodynamic or transport property, the self-diffusion coefficient is a function of thermodynamic state, for example, temperature, T , density, ρ , and composition, x . Since the diffusivity is frequently most sensitive to temperature, often its dependence on the other two variables, density (or its conjugate variable, pressure) and composition, is ignored. Ignoring the dependence of the diffusivity on pressure and composition is an approximation.

Transport properties can in principle also be functions of the non-equilibrium fields, such as the concentration gradient. (Shear-thinning polymers provide an example of a transport property, the shear viscosity, which is a function of the non-equilibrium field strength, the shear rate.) Typically, one makes the assumption that a diffusion coefficient is not a function of field strength. There is no theory behind this, only empirical evidence. In essence, we obtain reasonable agreement with experiment when we truncate the Taylor series describing the diffusive flux as a function of the concentration gradient at the linear term.

Expressions for the Self-diffusion coefficients from MD

The derivation for the expression of a self-diffusion comes from a Green-Kubo integral. For the case of the self-diffusivity, the argument of the Green-Kubo integral is the velocity [auto correlation function](#) (VACF). The diffusivity can equivalently be obtained from the mean square displacement (MSD). This is the more common approach, although both methods are formally equivalent. In Chapter 7, Haile has a simple and clear derivation of the MSD form from the VACF form.

In brief, the self-diffusion coefficient in the α direction can be obtained from the integration of the VACF.

$$D_{self} = \frac{1}{d} \int d\tau \langle v_{i,\alpha}(t + \tau) v_{i,\alpha}(t) \rangle \quad (3)$$

In this expression, d is the dimensionality of the system the angled brackets indicate an ensemble average, which is an average over both all particles $i = 1$ to N and an average over all time origins, t . This latter fact means that every time step in the MD simulation can be used as a time origin in the calculation of the VACF.

The self-diffusion coefficient can be also obtained from the mean square displacement (MSD)

$$D_{self} = \frac{1}{2d} \lim_{t \rightarrow \infty} \frac{\langle r_{i,\alpha}(t + \tau) - r_{i,\alpha}(t) \rangle^2}{\tau} \quad (4)$$

Where d is the dimensionality of the system. In this expression, often called Einstein's relation for the diffusivity, the angled brackets again indicate an ensemble average. The positions used in this calculation cannot have had periodic boundary conditions (PBCs) applied to them. If they have, then the trajectories have to be unfolded. In expressions for both the VACF and MSD, infinity appears. Of

course, one doesn't go to infinity observation times. At some point the effects become negligible. We discuss the impact of the choice of the infinite time limit in several examples below.

Lennard-Jones Potential and Lennard-Jones Fluid

The Lennard-Jones potential (also termed the LJ potential or 12-6 potential), named after the British Physics Sir John Edward Lennard-Jones, is an pair potential describing the interaction between a pair of particles. The Lennard-Jones potential models soft repulsive and attractive interactions. It is considered an archetype model for simple yet realistic intermolecular interactions. Indeed, it is a good approximation to ensembles of neutral particles like noble gases.

The commonly used expression for the Lennard-Jones potential is

$$V_{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (5)$$

where r is the distance between two interacting particles, ϵ is the depth of the potential well (usually referred to as 'dispersion energy'), and σ is the distance at which the particle-particle potential energy V is zero (often referred to as 'size of the particle'). The Lennard-Jones potential has its minimum at a distance of $r = r_m = 2^{1/6}\sigma$, where the potential energy has the value $V = -\epsilon$ (see Fig. 1).

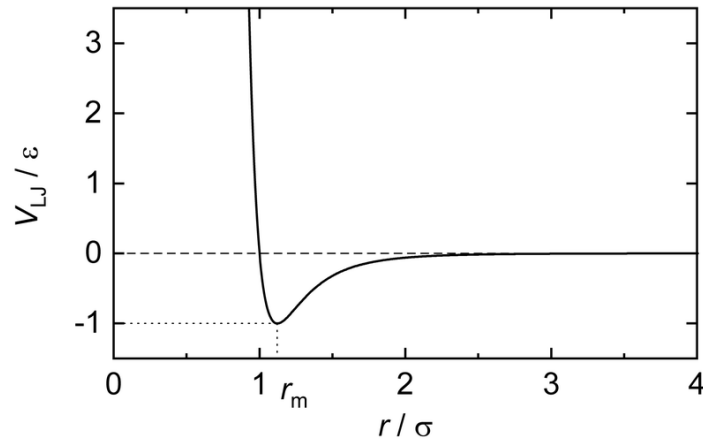


Figure 1. Graph of the Lennard-Jones potential function: Intermolecular potential energy V_{LJ} as a function of the distance of a pair of particles.

The Lennard-Jones potential is a simplified model that yet describes the essential features of interactions between simple atoms and molecules: Two interacting particles repel each other at very close distance, through the term of power of 12 term, while attract each other at moderate distance through the power of 6 term, not interacting at infinite distance (see Fig. 1). The Lennard-Jones potential is a pair potential, i.e. no three- or multi-body interactions are covered by the potential.

A Lennard-Jones substance is an ensemble of independent particles only interacting through a Lennard-Jones potential. Its thermodynamical properties can be obtained through statistical mechanics or molecular simulations. Figure 2 shows the phase diagram of the Lennard-Jones fluid. The mean intermolecular interaction of a Lennard-Jones particle strongly depends on the thermodynamic state, i.e. temperature and pressure (or density). For solid states, the attractive Lennard-Jones interaction plays a dominant role – especially at low temperatures. For liquid states, no ordered structure is present compared to solid states. The mean potential energy per particle is negative. For gaseous states, attractive interactions of the Lennard-Jones potential play a minor role – since they are far distanced. The main part of the internal energy is stored as kinetic energy for gaseous states. At supercritical states, the attractive Lennard-Jones interaction plays a minor role. With increasing temperature, the mean kinetic energy of the particles increases and exceeds the energy

well of the Lennard-Jones potential. Hence, the particles mainly interact by the potentials' soft repulsive interactions and the mean potential energy per particle is accordingly positive.

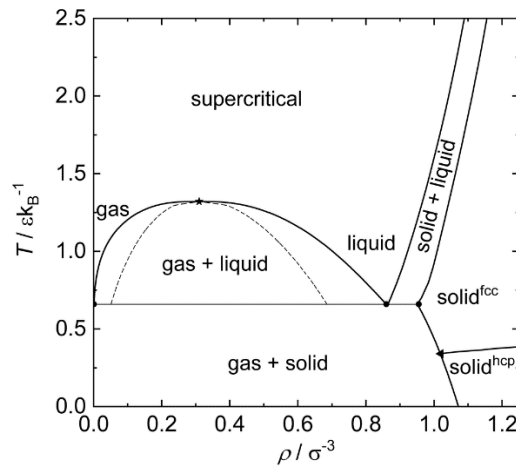


Figure 2. Phase diagram of the Lennard-Jones substance.

Calculating the Self-Diffusion Coefficient of a Lennard-Jones Fluid in LAMMPS

In the current assignment, the self-diffusion coefficient of a 2D Lennard-Jones fluid will be determined in LAMMPS. The system will be simple enough for the simulations to be carried out in seconds in a standard Linux laptop using only 1 CPU. Therefore, if LAMMPS is not already installed, it can be downloaded and installed as indicated in the [installation instructions](#) ([git](#) download and [cmake](#) build are recommended).

Attached is provided an input file for the calculation of the self-diffusion coefficient of a 2D Lennard-Jones fluid through the evolution of the mean squared displacement of a reference particle (in.ms.d.2d). The input starts with the definitions of some variables useful later for the calculation, namely box size (x and y) in Lennard-Jones units (σ for length and ϵ for energy), density and temperature of the fluid and cutoff of the interactions. Later, the input defines a collection of atoms (atom_style atomic) in a 2D rearrangement (dimension 2) in LJ units, creating the lattice with the given size and density (see sections lattice, region and create_box).

One of the most important sections of the input comes with the *pair_style* and *pair_coeff* lines. The first defines that two particles within a given cutoff will be interacting with a Lennard-Jones potential, while beyond the cutoff there will be no interaction. *Pair_coeff* defines that any particle i will interact with any particle j following a LJ potential (as defined previously) defined by the values of σ and ϵ 1 and 1, respectively.

The following five lines define the mass, initial velocities, and NVE ensemble with a Langevin thermostat moving in a 2D grid.

The next block is an dynamics to equilibrate the fluid, for 5000 timesteps. In the next section, the calculation of the mean squared displacement (MSD) is defined. Then the slope of the MSD is estimated from the finite difference between the current point and the previous, from which the diffusion coefficient is estimated from Eq. 4 (twopoint). Be aware that the dimensionality of the system is 2, $d = 2$. Additionally, the full MSD up to the current point is fit to a straight line, and the diffusion coefficient is again estimated through Eq. 4. In *thermo_style* the format of the output is set, where per each 1000th point (defined later in thermo 1000) the step number is displayed, the value of temperature, MSD, the D_{self} as calculated by finite difference and D_{self} as calculated by the whole fit. Finally the productive run for the calculation of the diffusion coefficient is invoked, saving every 1000th step (thermo) and propagating the dynamics 100000 steps (run).

Another input file is attached, in.vacf.2d. The syntax is rather similar to the previous file, the only difference being that the calculation of the diffusion coefficient will be done by time integration of the velocity autocorrelation function (Eq. 3). Compute calls the calculation of the velocity autocorrelation function (VACF). Then, in fix and variable the autocorrelation factor is integrated over time. Be aware that the factor $\frac{1}{2}$ arises from the system being 2 dimensional.

Running LAMMPS is pretty simple:

```
$ lmp < input_file > output_file
```

An log.lammps file will be also produced, with almost the same info as *output_file*.

The following tasks are required:

1. Run first the input in.msd.2d.
 - a. Plot the MSD and from its fit calculate D_{self} from Eq. 4
 - i. Is the plot linear, and if so, from which time?
 - b. Compare this value with the one obtained in the output.
2. Repeat the calculation but setting the equilibration run to only 1 timestep (no equilibration)
 - a. Plot the MSD versus time. How does it compare to the previous calculation?
3. Run the second input in.vacf.2d
 - a. Plot the VACF.
 - i. How converged is the VACF?
 - ii. How converged is the value of D_{self} displayed in the output
 - iii. How does this value compare with the one calculated through MSD.
4. Repeat the first calculation increasing the box size of x and y to 50 each.
 - a. Repeat the calculation and compare the value of the calculated diffusion coefficient to the original calculation. What is the origin of the discrepancy?
 - b. Increase the number of steps of the equilibration run to 10000 and repeat the simulation. How does now compare the diffusion calculated? Has your answer to question 4a) changed?
5. Repeat 1 with a density of 1 and a equilibration 100000 steps.
 - a. How have D_{self} changed? How can this change be explained?
6. Repeat 1 with a temperature of 0.7 and an equilibration of 10000 steps.
 - a. How have D_{self} changed? How can this change be explained?
7. Repeat 1 changing the values of σ and ϵ (line *pair_coeff*).
 - a. How can the differences observed in D_{self} be explained?