

The results of a simulation can take different forms depending on the objectives one seeks. A first and basic evaluation that must be made is to establish what the thermodynamic phase of the system is. To this end there exist various tools. In addition to the visual examination of the temporal evolution using appropriate visualization tools (in our course for example *Jmol*), there are more quantitative methods.

The most important of these methods is the calculation of the *pair correlation function* $g(r)$ = probability density that at a distance r from an atom there are another = average number $dn(r)$ of atoms in a spherical shell of thickness dr around r divided by dr . It is immediate to understand that if a system is a homogeneous solid at 0 temperature, the respective $g(r)$ will be a set of δ functions centered at the lattice distances, allowing to reconstruct the crystalline structure, if unknown. The integral of each δ will give the number of first neighbors (*coordination number*), second neighbors etc.. At the other extreme, a perfect gas, if we neglect the excluded volume effects, will have a proportional $g(r)$ to r^2 , since each atom moves independently of the others. Since this dependence is purely geometric and appears in all situations, and since it is much more convenient and efficient to deal with constant functions, the usual definition is

$$g(r)dr = dn(r)/(4\pi r^2)$$

which in the case of the perfect gas is a constant. What happens in the intermediate cases between a cold solid and a perfect gas? If we consider a crystal at finite T the δ s will widen due to thermal agitation, and will gradually tend to overlap. The space between the peaks where $g = 0$ will reduce as T increases until it disappears, with the exception of the $r \sim 0$ zone where it will remain 0 due to excluded volume effects (repulsive core). In a liquid those spaces will simply be throughs in the function $g(r)$, but if the fluid is sufficiently dense a structure of peaks and throughs remains, providing valuable information on coordinations and distances between atoms. As T increases, in particularly in supercritical conditions, the structure eventually disappears, leaving an almost steplike function.

The $g(r)$ is therefore valuable for the analysis, and moreover it is easy to connect to the experimental results, since its Fourier transform is the static structure factor $S(k)$, i.e. the result of the scattering experiments, both X-rays and neutrons.

The definition given above requires, to be useful, that the system is

entirely composed of equivalent atoms. If it is not, however, it is not difficult to extend it for example to binary (and similarly ternary etc.) mixtures of atoms of species A and B , introducing functions g_{AA} = average number $dn_{AA}(r)$ of atoms of type A around an atom A in a spherical shell of thickness dr around r divided by dr , and similarly g_{AB} and g_{BB} . For spatially inhomogeneous systems, for example surfaces of solids, it is also possible to define a series of different correlation functions. For example, for the case of the surface the averages can be limited to the distances relative to the atoms found on the surface obtaining a function $g_1(r)$, then one can define a second one limiting the averages to the distances relative to the atoms belonging to the immediately lower layer and obtaining a second function $g_2(r)$, then define a third etc., in order to observe the differences in the chemical-physical environment as the depth increases.

A second analysis method consists in calculating the *diffusion coefficient* D . Imagining for simplicity a homogeneous system, this coefficient is defined by the *diffusion equation*

$$\frac{\partial \rho}{\partial t} = D \nabla^2 \rho$$

where ρ is the concentration of particles, a function of time and space. Note the similarity to the equation of heat (and also with that of Schrödinger...). The coefficient D , which at equilibrium obeys the *Arrhenius law*

$$D = D_0 e^{-\beta Q}$$

where the *pre-exponential coefficient* D_0 and the (*free*) *activation energy* Q are, in the first approximation, constants, connects to microscopic quantities via the *Einstein relation*

$$D = \lim_{t \rightarrow \infty} \langle r^2 \rangle / 6t$$

where $\langle r^2 \rangle$ is the mean square displacement (ensemble mean at fixed time t), of the particles from the starting point (coordinates at $t = 0$). Note that for free particles, with ballistic motion, $r \sim t$ therefore $\langle r^2 \rangle \propto t^2$ and a finite limit does not exist. Diffusive motion therefore occurs only in the presence of interactions between particles that transform the motion into a random walk, in which $r \sim \sqrt{t}$. Vice versa, for a solid system, if we neglect the motions that take the atoms out of their lattice sites, $\langle r^2 \rangle$ is a constant, closely linked to the *Debye-Waller factor*. Note also that the limit $t \rightarrow \infty$ is essential,

because by terminating the calculation at the appropriate t you can obtain more or less any desired value... Used correctly, however, it allows, among other things, to distinguish between a solid, for which $\langle r(t) \rangle$ oscillates around a constant value, from a fluid, for which $\langle r(t) \rangle$ is an increasing straight line unless there are fluctuations.

The diffusion coefficient and the pair correlation function are special cases of a set of quantities called *correlation functions* that express the dynamic connections within a system. Any mechanical quantity can be correlated by defining functions of this type. Another important example is the *velocity autocorrelation function* defined as

$$C(t) = \lim_{\sigma \rightarrow \infty} \frac{1}{\sigma} \int_0^\sigma d\tau \langle \mathbf{v}_I(t + \tau) \cdot \mathbf{v}_I(\tau) \rangle$$

where the average is also in this case the ensemble average, at a fixed t , on all I atoms. Let's consider two extreme cases: 1) $\mathbf{v}_I(t)$ is constant with respect to t for every I (maximum correlation), 2) $\mathbf{v}_I(t)$ is a random variable (zero correlation). In case (1), $C(t) = \langle \mathbf{v}_I^2 \rangle$ is a non-zero constant (it is proportional to the total average kinetic energy), in case (2) the integral over τ will give 0 since each value of the positive integral will correspond to a negative value equal in magnitude. It shows therefore we actually obtain a measure of how much the system "remembers" its past dynamics: the greater the modulus of C , the greater the "memory".

Finally, if we set $\mathbf{v}_I(t) = A \cos \omega t$, then

$$\begin{aligned} C(t) &= \lim_{\sigma \rightarrow \infty} \left[\frac{A^2}{\sigma} \int_0^\sigma d\tau \cos \omega \tau^2 \cos \omega t - \frac{A^2}{\sigma} \int_0^\sigma d\tau \cos \omega \tau \sin \omega \tau \sin \omega t \right] \\ &= \frac{A^2}{2} \cos \omega t. \end{aligned}$$

It can therefore be seen how the C also provides information on the vibrational spectrum of the system. Taking the Fourier transform $C(\omega)$ we therefore obtain the power spectrum of the system dynamics and from this, exploiting the equilibrium hypothesis, also the vibrational density of states, which is its product with a Boltzmann factor. For a solid, if the eigenvectors of the dynamical matrix are known the phonon frequencies can also be obtained.

Finally, we observe that the value at zero frequency

$$\begin{aligned}
C(\omega = 0) &= \lim_{\sigma' \rightarrow \infty} \int_0^{\sigma'} dt C(t) = \lim_{\sigma' \rightarrow \infty} \int_0^{\sigma'} dt \lim_{\sigma \rightarrow \infty} \frac{1}{\sigma} \int_0^{\sigma} d\tau \langle \mathbf{v}_I(t + \tau) \cdot \mathbf{v}_I(\tau) \rangle \\
&= \lim_{\sigma, \sigma' \rightarrow \infty} \frac{1}{\sigma} \left\langle \int_0^{\sigma} d\tau \mathbf{v}_I(\tau) \cdot \int_0^{\sigma'} dt \mathbf{v}_I(t + \tau) \right\rangle \\
&= \lim_{\sigma, \sigma' \rightarrow \infty} \frac{1}{\sigma} \langle (\mathbf{r}_I(\sigma) - \mathbf{r}_I(0)) \cdot (\mathbf{r}_I(\sigma') - \mathbf{r}_I(0)) \rangle \propto \lim_{\sigma \rightarrow \infty} \langle r^2 \rangle / \sigma \propto D.
\end{aligned}$$

Usually, however, the direct calculation of the mean square displacement is preferable.