Introduction to correlation, Kirkwood-Buff theory, Onsagers regression theory, and Green's functions

Winter School for Theoretical Chemistry and Spectroscopy Han-sur-Lesse, 10-14 December 2018

> by Bernd Ensing

Content

- Spatial correlation
- Time correlation
- Application: Anisotropy decay in water
- Reaction rate theory, reactive-flux method (Bennet-Chandler)

The structure of materials is governed interaction between the particles (atoms, molecules,...) and the external conditions (T, p,...).



The radial distribution function (RDF), aka the paircorrelation function, measures the (excess) probability to find a particle *j* at a distance, *r*, from a reference particle *i*, relative to the average density of particles *j*.

average density:

 $\rho_j = \frac{n_j}{V}$

The probability to find a certain configuration of particle positions is given bij Boltzmann's distribution:

$$P(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \frac{1}{Z} e^{-\beta V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)} \qquad \beta = 1/k_B T$$
$$Z = \int \cdots \int e^{-\beta V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)} d\mathbf{r}_1 \dots d\mathbf{r}_N$$

The probability to find particles 1 and 2 at a certain position:

$$P(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{Z} \int \cdots \int e^{-\beta V(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N)} d\mathbf{r}_3 \ldots d\mathbf{r}_N$$

The probability that any of the particles is at those positions is given by the 2-particle density: NI

$$\rho(\mathbf{r}_1, \mathbf{r}_2) = \frac{N!}{(N-2)!} P(\mathbf{r}_1, \mathbf{r}_2)$$



Define the correlation function $g(\mathbf{r}_1, \mathbf{r}_2)$

$$g(\mathbf{r}_1, \mathbf{r}_2) = \frac{\rho(\mathbf{r}_1, \mathbf{r}_2)}{\rho^2} = \frac{V^2 N!}{N^2 (N-2)!} \frac{1}{Z} \int \cdots \int e^{-\beta V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)} d\mathbf{r}_3 \dots d\mathbf{r}_N$$

is equal to one if there is no correlation (= no interaction)

The radial distribution function (RDF), g(r), considers the correlation of particles "2" at a distance *r* from particles "1"

The number of particles at a distance between *r* and *r+dr* from any central particle:

• at very small distances: g(0) = 0

• at very large distances: $g(\infty) = 1$

The RDF can be obtained from:

- MD or MC simulation
- neutron or X-ray diffraction (structure factor)

 $4\pi r^2 \rho g(r) dr$

microscopy (for colloids and other large particles)











From website of Dr R K Thomas FRS and Prof J Penfold: http://rkt.chem.ox.ac.uk/lectures/liqsolns/liquids.html

Liquid water, DFT-MD simulation



From Ambuj Tiwari (UvA)



Fig. 3. MD simulations of the investigated solution. (**A**) Snapshot of the classical MD simulation. (**B**) Radial distribution functions (RDFs)

of different H-bond donor and acceptor atoms in the solution around water atoms, obtained from the classical (solid lines) and DFT (dashed lines) MD simulations. In the top panel, radial distributions of the water O and H atoms are compared to these functions in bulk water (dotted lines). (**C** and **D**) Three-dimensional structural density plots of the first coordination shell of a water molecule, in neat water (C) and in N2H5TFA solution (D), showing the distribution of different H-bond donor and acceptor atoms. Color code for isosurfaces: red, water O atom; gray, water H atom; orange, TFA O atom; blue, hydrazinium H atom; green, hydrazinium N atom. (**E** to **G**) Three-dimensional structural density plots of the H-bond acceptor atoms (water O, TFA O, and hydrazinium N) in the first and second coordination shells of a water molecule (in the center) in neat water at 1 bar (E), in neat water at 6 kbar (F), and in N2H5TFA solution at 1 bar (G).

Thermodynamic properties

The RDF can be used to compute various thermodynamic properties, such as:

- potential energy
- pressure
- compressibility

total energy:

$$U = U^{\text{kin}} + U^{\text{intra}} + \frac{1}{2}N\rho \int_{0}^{\infty} dr \, 4\pi r^{2}g(r)u(r)$$
translation energy
$$U^{\text{kin}} = \frac{3}{2}Nk_{B}T$$
potential energy
of a particle

Pressure

The pressure in the canonical ensemble is computed for a system with pairwise interactions from the virial equation:

$$P = \rho k_B T + \frac{1}{3V} \left\langle \sum_{i < j} \mathbf{f}(\mathbf{r}_{ij}) \cdot \mathbf{r}_{ij} \right\rangle$$

using instead the radial distribution function:

$$P = \rho k_B T - \frac{2}{3} \pi \rho^2 \int_0^\infty dr \, \frac{du(r)}{dr} r^3 g(r) dr$$
avoid double force

Isothermal compressibility

The isothermal compressibility is: $\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_{T,N}$

which is linked to the fluctuations in the number of particles in an open system

$$\langle N \rangle \rho k_B T \kappa_T = \left\langle (N - \langle N \rangle)^2 \right\rangle = \langle N^2 \rangle - \langle N \rangle^2$$

integration of the RDF over all particle 1 and 2 positions gives something similar:

$$\int_{V} d^{3}r_{1} \int_{V} d^{3}r_{2}\rho^{2}g(r_{12}) = \left\langle N(N-1)\right\rangle = \left\langle N^{2}\right\rangle - \left\langle N\right\rangle$$

add $\langle N \rangle$ and subtract $\langle N \rangle^2$:

$$\langle N \rangle \rho k_B T \kappa_T = \rho \int_V d^3 r_1 \rho \int_V d^3 r_2 g(r_{12}) + \langle N \rangle - \rho \int_V d^3 r_1 \rho \int_V d^3 r_2 = \rho \int_V d^3 r_1 \rho \int_V d^3 r_2 (g(r_{12}) - 1) + \langle N \rangle$$
compressibility:
$$\kappa_T = (k_B T)^{-1} \int_{\mathbb{R}^3} d^3 r (g(r) - 1) + (\rho k_B T)^{-1}$$

Dynamical properties

Thus far: spatial correlations between particles.

A material or liquid in equilibrium has also an underlying dynamics of the particles. Particle positions and velocities continuously change. Most quantities fluctuate in time.



The dynamics of microscopic processes can be captured by time-correlation functions

$$\langle A(0)B(t)\rangle = \lim_{T \to \infty} \frac{1}{T} \int_0^T d\tau A(\tau)B(\tau+t)$$

$$\langle A(0)A(t)\rangle = \lim_{T \to \infty} \frac{1}{T} \int_0^T d\tau A(\tau)A(\tau+t)$$

time auto-correlation function (correlate with itself)

At very short time intervals, the correlation is maximal $\cdot \langle A(0)A(0) \rangle = \langle A^2 \rangle$

At very long times, the correlation has decayed to zero

- $\langle A(0)A(\infty)\rangle = \langle A\rangle\langle A\rangle$
- $\langle A(0)B(\infty)\rangle = \langle A\rangle\langle B\rangle$

Often, the correlation function is normalized by dividing it by $\langle A(0)A(0) \rangle$

Time correlation functions

In equilibrium, any time can be taken as the t=0 moment, so, $\langle A(t)A(t + \Delta t) \rangle$ is independent of t. Therefore:

$$\frac{d}{dt}\langle A(t)A(t+\Delta t)\rangle = \langle \dot{A}(t)A(t+\Delta t)\rangle + \langle A(t)\dot{A}(t+\Delta t)\rangle = 0$$

and

$$\langle \dot{A}(t)A(t+\Delta t) \rangle = -\langle A(t)\dot{A}(t+\Delta t) \rangle$$

but note that A(t) and $A(t+\Delta t)$ are permutable: $\langle A(t)A(t+\Delta t)\rangle = \langle A(t+\Delta t)A(t)\rangle$

$$\left. \frac{d}{dt} \langle A(0)A(t) \rangle \right|_{t=0} = 0$$

every correlation function starts with a horizontal slope

time derivative

We can also derive:

erive:
$$\frac{d^2}{dt^2} \langle A(t)A(t+\Delta t) \rangle = -\langle \dot{A}(t)\dot{A}(t+\Delta t) \rangle$$

$$\frac{d^2}{dt^2} \langle A(0)A(t)\rangle \bigg|_{t=0} = -\langle \dot{A^2}(t)\rangle < 0$$

every correlation function starts with a negative curvature

Example: self-diffusion



Diffusive dynamics is characterised by a mean square displacement that is proportional to time.

~~~

in one dimension, for large enough t:

$$\left\langle \left( x(t) - x(0) \right)^2 \right\rangle = 2Dt$$

(Einstein equation)

$$\begin{aligned} \langle x^2(t) \rangle &= \int_0^t dt' \int_0^t dt'' \langle v_x(t') v_x(t'') \rangle = 2 \int_0^t dt' \int_0^{t'} dt'' \langle v_x(t') v_x(t'') \rangle \\ D &= \frac{1}{2} \frac{d \langle x^2(t) \rangle}{dt} = \int_0^t dt' \langle v_x(0) v_x(t''-t') \rangle \end{aligned} \stackrel{\text{example table tabl$$

in 3D:

$$D = \frac{1}{3} \int_0^\infty d\tau \langle v(0)v(\tau) \rangle$$
 Green-Kubo relation (relates a transport property, here D, to a time correlation function)

# On the slowdown mechanism of water dynamics around small amphiphiles

Two amphiphilic molecules in water

Urea

Tetramethyl urea







Wagner H. Brandeburgo Evert Jan Meijer Sietse T. van der Post



#### Fit function:

$$C_2(t) = A_0 \left( \cos\left(\frac{2\pi t}{\tau_{\text{sub}}}\right) \cdot e^{\frac{-t}{\tau_{\text{d}}}} + e^{\frac{-t}{\tau_{\text{fast}}}} \right) + (1 - A_0) e^{\frac{-t}{\tau_{\text{mid}}}}$$

Water partitioning







#### Water partitioning



- hydrophilic solvent water
- hydrophobic solvent water
- bulk-like solvent water
- ---- pure water

| System method                                                                      | Fit interval (ps) | $\tau_{\rm fast} \left( ps \right)$ | $\tau_{mid}$ (ps)              | $\tau_{\rm slow}$ (ps)                |
|------------------------------------------------------------------------------------|-------------------|-------------------------------------|--------------------------------|---------------------------------------|
| Pure water CMD<br>All water                                                        | 0-50              | 0.3                                 | 2.4 (1.9)                      | _                                     |
| Pure water AIMD<br>All water                                                       | 0-15              | 0.5                                 | 3.1 (2.1)                      | _                                     |
| Pure water fs-IR<br>All water                                                      | 2-10              | _                                   | — (2.3)                        | _                                     |
| Urea CMD<br>All water<br>Hydrophilic water<br>Hydrophobic water<br>Bulk-like water | 0-12              | 0.6<br>0.5<br>0.6<br>0.3            | 3.6<br>6.1<br>5.3<br>2.7       |                                       |
| TMU CMD<br>All water<br>Hydrophilic water<br>Hydrophobic water<br>Bulk-like water  | 0-20              | 0.3<br>0.3<br>0.3<br>0.3            | 3.0 (2.2)<br>5.1<br>3.3<br>2.9 | 9.0 (∞)<br>>20<br>15.5<br>—           |
| TMU AIMD<br>All water<br>Hydrophilic water<br>Hydrophobic water                    | 0–12              | 0.2<br>0.2<br>0.1                   | 2.0 (2.8)<br>3.9<br>2.3        | >12 ( $\infty$ )<br>>12<br>>12<br>>12 |
| TMU fs-IR<br>All water                                                             | 2-10              | _                                   | — (2.6)                        | — (∞)                                 |

#### H-bond life times and breaking mechanism

| System method     | $\tau_{\text{H-bond}} \text{ (ps)}$ | Switches (%) |
|-------------------|-------------------------------------|--------------|
| Pure water CMD    |                                     |              |
| All water         | 2.1                                 | 74.6         |
| Pure water AIMD   |                                     |              |
| All water         | 2.2                                 | 75.9         |
| Urea CMD          |                                     |              |
| All water         | 2.6                                 | 75.6         |
| Hydrophilic water | 3.1                                 | 73.8         |
| Hydrophobic water | 2.7                                 | 74.4         |
| Bulk-like water   | 2.4                                 | 77.0         |
| TMU CMD           |                                     |              |
| All water         | 2.9                                 | 68.8         |
| Hydrophilic water | 5.1                                 | 56.6         |
| Hydrophobic water | 3.1                                 | 65.6         |
| Bulk-like water   | 2.4                                 | 74.6         |
| TMU AIMD          |                                     |              |
| All water         | 4.1                                 | 62.3         |
| Hydrophilic water | 6.0                                 | 54.8         |
| Hydrophobic water | 3.9                                 | 61.2         |

#### Model

| <ul> <li>Jump model by Laage and</li> </ul>       |
|---------------------------------------------------|
| Hynes, <i>J. Phys. Chem. B</i>                    |
| 112,14230–14242 (2008)                            |
| <ul> <li>Hydrogen bond life times</li> </ul>      |
| <ul> <li>Associative / dissociative H-</li> </ul> |
| bond breaking                                     |
| Close to hydrophobic surface:                     |
| dissociative mechanism                            |
| becomes important (excluded                       |
| volume effect)                                    |
| <ul> <li>After dissociative breaking,</li> </ul>  |
| bond often restores, resetting                    |
| the timer (new life time)                         |

# PEG and POM

The solvation mystery unraveled from the solvent dynamics

- polyethylene glycol, [-CH<sub>2</sub>-CH<sub>2</sub>-O-]<sub>n</sub>
- polyoxymethylene, [-CH<sub>2</sub>-O-]<sub>n</sub>







### In real life (experiment)

- $\cdot$  no direct access to free energy
- reaction rate measurement







### Rare event simulation

Macroscopic phenomenological theory

Chemical reaction:  $A \leftrightarrow B$ 

$$\frac{dc_A(t)}{dt} = -k_{A\to B}c_A(t) + k_{B\to A}c_A(t)$$
$$\frac{dc_B(t)}{dt} = +k_{A\to B}c_A(t) - k_{B\to A}c_A(t)$$

Total number of molecules:

Equilibrium:

$$\frac{d[c_A(t) + c_B(t)]}{dt} = 0 \qquad \qquad \frac{dc_A(t)}{dt} = \frac{dc_B(t)}{dt} = 0 \qquad \frac{\langle c_A \rangle}{\langle c_B \rangle} = \frac{k_{B \to A}}{k_{A \to B}}$$

### Rare event simulation

Macroscopic phenomenological theory

Make a small perturbation:

$$c_A(t) = \langle c_A \rangle + \Delta c_A(t) \qquad c_B(t) = \langle c_B \rangle + \Delta c_A(t)$$
$$\frac{d\Delta c_A(t)}{dt} = -k_{A \to B} \Delta c_A(t) - k_{B \to A} \Delta c_A(t)$$

$$\Delta c_A(t) = \Delta c_A(0) \exp[-(k_{A \to B} + k_{B \to A})t]$$
  
=  $\Delta c_A(0) \exp[-t/\tau]$ 

$$\tau = \left(k_{A \to B} + k_{B \to A}\right)^{-1}$$
$$= k_{A \to B}^{-1} \left(1 + \langle c_A \rangle / \langle c_B \rangle\right)^{-1} = \frac{\langle c_B \rangle}{k_{A \to B}}$$

### Rare event simulation



Microscopic linear response theory

$$(q - q^*) = \begin{cases} 0 & \text{if } q - q^* < 0 \quad (\text{Reactant A}) \\ 1 & \text{if } q - q^* > 0 \quad (\text{Product B}) \end{cases}$$

Perturbation add bias to increase concentration *c*<sub>A</sub>

$$H = H_0 - \epsilon g_A (q - q^*)$$

$$g_A(q - q^*) = 1 - \theta(q - q^*) = \theta(q^* - q)$$

 $\Delta c_A = \left\langle c_A \right\rangle_{\epsilon} - \left\langle c_A \right\rangle_0$  $\Delta c_A = \left\langle g_A \right\rangle_{\epsilon} - \left\langle g_A \right\rangle_0$ 

 $\langle g_A \rangle \rightarrow$  probability to be in state A

#### Linear response theory: static

$$\langle \Delta A \rangle = \langle A \rangle - \langle A \rangle_0 \qquad H = H_0 - \epsilon B$$

$$\left\langle A\right\rangle_{0} = \frac{\int d\Gamma A \exp[-\beta(H_{0})]}{\int d\Gamma \exp[-\beta(H_{0})]} \qquad \left\langle A\right\rangle = \frac{\int d\Gamma A \exp[-\beta(H_{0} - \epsilon B)]}{\int d\Gamma \exp[-\beta(H_{0} - \epsilon B)]}$$

$$\left\langle \frac{\partial \Delta A}{\partial \epsilon} \right\rangle = \frac{\int d\Gamma \beta AB \exp[-\beta(H_0 - \epsilon B)] \int d\Gamma \exp[-\beta(H_0 - \epsilon B)]}{\left\{ \int d\Gamma \exp[-\beta(H_0 - \epsilon B)] \right\}^2} - \frac{\int d\Gamma A \exp[-\beta(H_0 - \epsilon B)] \int d\Gamma \beta B \exp[-\beta(H_0 - \epsilon B)]}{\left\{ \int d\Gamma \exp[-\beta(H_0 - \epsilon B)] \right\}^2} = \beta \left\{ \left\langle AB \right\rangle_0 - \left\langle A \right\rangle_0 \left\langle B \right\rangle_0 \right\}$$

Very small perturbation: linear response theory

$$\Delta c_A = \left\langle g_A \right\rangle_{\epsilon} - \left\langle g_A \right\rangle_0 \qquad \qquad H = H_0 - \epsilon g_A (q - q^*)$$

How does the response ( $\Delta c$ ) depend on the perturbation ( $\Delta \epsilon$ )?

Switch of the perturbation: dynamic linear response

$$\Delta c_A(t) = \Delta c_A(0) \frac{\Delta g_A(0) \Delta g_A(t)}{\langle c_A \rangle \langle c_B \rangle}$$
$$= \Delta c_A(0) \exp[-t/\tau]$$

holds for sufficiently long times

$$\exp[-t/\tau] = \frac{\left\langle \Delta g_A(0) \Delta g_A(t) \right\rangle}{\left\langle c_A \right\rangle \left\langle c_B \right\rangle}$$

Derivative 
$$-\frac{1}{\tau} \exp[-t/\tau] = \frac{\langle g_A(0)\dot{g}_A(t)\rangle}{\langle c_A\rangle\langle c_B\rangle} = \frac{\langle \dot{g}_A(0)g_A(t)\rangle}{\langle c_A\rangle\langle c_B\rangle}$$

 $\Delta$  has disappeared because of derivative

For sufficiently short *t* 

$$k_{A\to B}(t) = \frac{\left\langle \dot{g}_A(0)g_A(t)\right\rangle}{\left\langle c_A\right\rangle}$$
$$\dot{g}_A(q-q^*) = \dot{q}\frac{\partial g_A(q-q^*)}{\partial q} = -\dot{q}\frac{\partial g_B(q-q^*)}{\partial q}$$
$$k_{a\to B}(t) = \frac{\left\langle \dot{q}(0)\frac{\partial g_B(q(0)-q^*)}{\partial q}g_B(t)\right\rangle}{\left\langle c_A\right\rangle}$$

Stationary  

$$\frac{d}{dt} \langle A(t)B(t+t') \rangle = 0$$

$$\langle A(t)\dot{B}(t+t') \rangle + \langle \dot{A}(t)B(t+t') \rangle = 0$$

$$\langle A(t)\dot{B}(t+t') \rangle = -\langle \dot{A}(t)B(t+t') \rangle$$

## Eyring's transition state theory

$$k_{a \to B}(t) = \frac{\left\langle \dot{q}(0) \frac{\partial g_B(q(0) - q^*)}{\partial q} g_B(t) \right\rangle}{\left\langle c_A \right\rangle}$$
$$= \frac{\left\langle \dot{q}(0) \delta(q(0) - q^*) \theta(q(t) - q^*) \right\rangle}{\left\langle \theta(q^* - q) \right\rangle}$$

Correlation between velocity of states that are at the top of the barrier at t=0 and in the product state B some time t later.

Let us consider the limit  $t \rightarrow 0^+$ :

$$\lim_{t \to 0^+} = \theta(q(t) - q^*) = \theta(\dot{q}(t))$$
$$k_{a \to B}^{\text{TST}}(t) = \frac{\langle \dot{q}(0)\delta(q(0) - q^*)\theta(\dot{q}) \rangle}{\langle \theta(q^* - q) \rangle}$$

## **Bennett-Chandler** approach

(or Reactive flux method)

$$k_{a \to B}(t) = \frac{\left\langle \dot{q}(0)\delta(q(0) - q^*)\theta(q(t) - q^*)\right\rangle}{\left\langle \theta(q^* - q)\right\rangle}$$

$$k_{a \to B}(t) = \frac{\left\langle \dot{q}(0)\delta(q(0) - q^*)\theta(q(t) - q^*) \right\rangle}{\left\langle \delta(q(0) - q^*) \right\rangle} \times \frac{\delta(q(0) - q^*)}{\theta(q^* - q)}$$
Conditional average:  $\dot{q}(0)\theta(q(t) - q^*)$ 
given that we start on top of barrier
Probability to find  $q$ 
on barrier top

#### Computational scheme:

- Determine the probability with free energy calculation
- Compute conditional average from "shooting" trajectories from barrier top



- The structure of matter and fluids is governed by interactions between particles, which causes spatial correlations.
- The best know is the pair-correlation function, or radial distribution function, which can be obtained from MD or MC simulation, neutron or X-ray diffraction, or microscopy.
- Dynamics can be probed by time-correlation functions.
- Green-Kubo relations connect macroscopic transport properties to microscopic time correlation functions.
- The orientational dynamics of water molecules in solutions can be captured by the decay of the dipole auto-correlation function.
- The Bennett-Chandler approach, or reactive flux method, allows for accurate estimation of reaction rates, by combining a free energy calculation (giving the transition state theory estimate of the rate) with a correction using the transmission coefficient.