165.144 Simulation of condensed matter

2025 Ass.Prof. Esther Heid esther.heid@tuwien.ac.at

Summary chapter 3

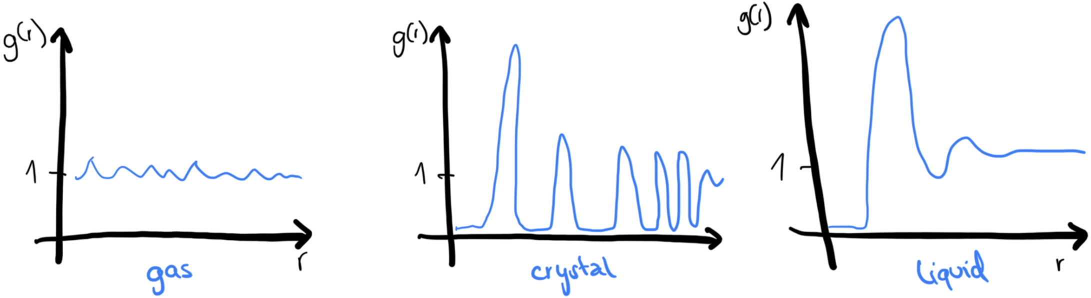
- Utilizing cell subdivison and neighbor lists, we can make our MD code fast enough to study larger systems

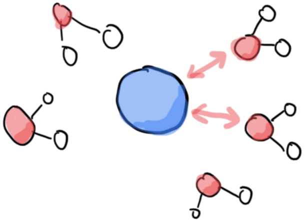
Summary chapter 3

- Utilizing cell subdivison and neighbor lists, we can make our MD code fast enough to study larger systems
- Before starting a measurement, the systems needs to reach equilibrium

Summary chapter 3

- Utilizing cell subdivison and neighbor lists, we can make our MD code fast enough to study larger systems
- Before starting a measurement, the systems needs to reach equilibrium
- Static properties like the spatial structure and organization of particles in a system are directly accessible from equilibrium simulations





Chapter 4: Dynamic properties

Properties of a system

Static properties describe measurements of thermodynamic properties, spatial structure and organization

- Potential/kinetic energy, heat capacity, pressure, temperature
- Radial distribution functions
- Number of neighbors
- Order

Properties of a system

Static properties describe measurements of thermodynamic properties, spatial structure and organization

- Potential/kinetic energy, heat capacity, pressure, temperature
- Radial distribution functions
- Number of neighbors
- Order

Dynamic properties describe the movements of atoms

- Diffusion coefficients
- Conductivity
- Time-dependent spectroscopy
- Viscosity

Dynamic properties

Dynamic properties are related to nonequilibrium properties. For example, transport coefficients like the diffusion, shear and bulk viscosity, and thermal conductivity arise from mechanical perturbations (concentration, velocity, thermal gradients).

Dynamic properties

Dynamic properties are related to nonequilibrium properties. For example, transport coefficients like the diffusion, shear and bulk viscosity, and thermal conductivity arise from mechanical perturbations (concentration, velocity, thermal gradients).

Linear response theory relates the reaction of an equilibrium system to a small external perturbation via equilibrium correlation functions. We can thus compute dynamic properties from equilibrium trajectories.

Dynamic properties

Dynamic properties are related to nonequilibrium properties. For example, transport coefficients like the diffusion, shear and bulk viscosity, and thermal conductivity arise from mechanical perturbations (concentration, velocity, thermal gradients).

Linear response theory relates the reaction of an equilibrium system to a small external perturbation via equilibrium correlation functions. We can thus compute dynamic properties from equilibrium trajectories.

Namely, the Green-Kubo formalism connects a macroscopic property (=response property of a system) to its equilibrium fluctuations. For example, diffusion is the response to concentration inhomogeneity.

Diffusion

Diffusion coefficients D accessible from the mean-square displacement (**Einstein expression**):

$$D = \lim_{t\to\infty} \frac{\langle \Delta r^2(t) \rangle}{6t} = \lim_{t\to\infty} \frac{1}{6Nt} \sum_j \Delta r_j^2$$

Diffusion

Diffusion coefficients D accessible from the mean-square displacement (**Einstein expression**):

$$D = \lim_{t\to\infty} \frac{\langle \Delta r^2(t) \rangle}{6t} = \lim_{t\to\infty} \frac{1}{6Nt} \sum_j \Delta r_j^2$$

Derivation:

- Fick's law with ρ... local density or concentration: $\frac{\partial \rho(\mathbf{r},t)}{\partial t} = D\nabla^2 \rho(\mathbf{r}, t)$

Diffusion

Diffusion coefficients D accessible from the mean-square displacement (**Einstein expression**):

$$D = \lim_{t\to\infty} \frac{\langle \Delta r^2(t) \rangle}{6t} = \lim_{t\to\infty} \frac{1}{6Nt} \sum_j \Delta r_j^2$$

Derivation:

- Fick's law with ρ... local density or concentration: $\frac{\partial \rho(\mathbf{r},t)}{\partial t} = D\nabla^2 \rho(\mathbf{r},t)$
- Multiply by Δr^2 and integrate along $\mathbf{r}$: $\frac{\partial}{\partial t} \int \mathrm{d}\mathbf{r} \Delta r^2 \rho(\mathbf{r},t) = D \int \mathrm{d}\mathbf{r} \Delta r^2 \nabla^2 \rho(\mathbf{r},t)$

Diffusion

Diffusion coefficients D accessible from the mean-square displacement (**Einstein expression**):

$$D = \lim_{t\to\infty} \frac{\langle \Delta r^2(t) \rangle}{6t} = \lim_{t\to\infty} \frac{1}{6Nt} \sum_j \Delta r_j^2$$

Derivation:

- Fick's law with ρ... local density or concentration: $\frac{\partial \rho(\mathbf{r},t)}{\partial t} = D\nabla^2 \rho(\mathbf{r},t)$
- Multiply by Δr^2 and integrate along $\mathbf{r}$: $\frac{\partial}{\partial t} \int \mathrm{d}\mathbf{r} \Delta r^2 \rho(\mathbf{r},t) = D \int \mathrm{d}\mathbf{r} \Delta r^2 \nabla^2 \rho(\mathbf{r},t)$
- Integrate by parts and assume ρ is normalized to 1: $\frac{\partial}{\partial t} \int \mathrm{d}\mathbf{r} \Delta r^2 \rho(\mathbf{r},t) = 2D \cdot d = 6D$

Diffusion

Diffusion coefficients D accessible from the mean-square displacement (**Einstein expression**):

$$D = \lim_{t\to\infty} \frac{\langle \Delta r^2(t) \rangle}{6t} = \lim_{t\to\infty} \frac{1}{6Nt} \sum_j \Delta r_j^2$$

Derivation:

- Fick's law with ρ... local density or concentration: $\frac{\partial \rho(\mathbf{r},t)}{\partial t} = D \nabla^2 \rho(\mathbf{r},t)$
- Multiply by Δr^2 and integrate along $\mathbf{r}$: $\frac{\partial}{\partial t} \int \mathrm{d}\mathbf{r} \Delta r^2 \rho(\mathbf{r},t) = D \int \mathrm{d}\mathbf{r} \Delta r^2 \nabla^2 \rho(\mathbf{r},t)$
- Integrate by parts and assume ρ is normalized to 1: $\frac{\partial}{\partial t} \int \mathrm{d}\mathbf{r} \Delta r^2 \rho(\mathbf{r},t) = 2D \cdot d = 6D$
- Integral is an expectation value since ρ behaves like a probability: $\frac{\partial}{\partial t} \langle \Delta r^2(t) \rangle = 6D$

Diffusion

Diffusion coefficients D accessible from the mean-square displacement (**Einstein expression**):

$$D = \lim_{t\to\infty} \frac{\langle \Delta r^2(t) \rangle}{6t} = \lim_{t\to\infty} \frac{1}{6Nt} \sum_j \Delta r_j^2$$

Derivation:

- Fick's law with ρ... local density or concentration: $\frac{\partial \rho(\mathbf{r},t)}{\partial t} = D \nabla^2 \rho(\mathbf{r},t)$
- Multiply by Δr^2 and integrate along $\mathbf{r}$: $\frac{\partial}{\partial t} \int \mathrm{d}\mathbf{r} \Delta r^2 \rho(\mathbf{r},t) = D \int \mathrm{d}\mathbf{r} \Delta r^2 \nabla^2 \rho(\mathbf{r},t)$
- Integrate by parts and assume ρ is normalized to 1: $\frac{\partial}{\partial t} \int \mathrm{d}\mathbf{r} \Delta r^2 \rho(\mathbf{r},t) = 2D \cdot d = 6D$
- Integral is an expectation value since ρ behaves like a probability: $\frac{\partial}{\partial t} \langle \Delta r^2(t) \rangle = 6D$
- $\langle \Delta r^2(t) \rangle$ is linear at long time scales: $\lim_{t\to\infty} \frac{\langle \Delta r^2(t) \rangle}{t} = 6D$

Diffusion

Diffusion coefficients D accessible from the mean-square displacement (**Einstein expression**):

$$D = \lim_{t\to\infty} \frac{\langle \Delta r^2(t)\rangle}{6t} = \lim_{t\to\infty} \frac{1}{6Nt} \sum_j \Delta r_j^2$$

Derivation:

- Fick's law with ρ... local density or concentration: $\frac{\partial \rho(\mathbf{r},t)}{\partial t} = D\nabla^2 \rho(\mathbf{r},t)$
- Multiply by Δr^2 and integrate along $\mathbf{r}$: $\frac{\partial}{\partial t} \int \mathrm{d}\mathbf{r} \Delta r^2 \rho(\mathbf{r},t) = D \int \mathrm{d}\mathbf{r} \Delta r^2 \nabla^2 \rho(\mathbf{r},t)$
- Integrate by parts and assume ρ is normalized to 1: $\frac{\partial}{\partial t} \int \mathrm{d}\mathbf{r} \Delta r^2 \rho(\mathbf{r},t) = 2D \cdot d = 6D$
- Integral is an expectation value since ρ behaves like a probability: $\frac{\partial}{\partial t} \langle \Delta r^2(t) \rangle = 6D$
- $\langle \Delta r^2(t) \rangle$ is linear at long time scales: $\lim_{t\to\infty} \frac{\langle \Delta r^2(t)\rangle}{t} = 6D$
- $\langle \Delta r^2(t) \rangle$ is the average over the particles: $\lim_{t\to\infty} \frac{\sum_j \Delta r_j(t)^2}{Nt} = 6D$

Diffusion

Diffusion coefficients D accessible from the mean-square displacement (**Einstein expression**):

$$D = \lim_{t\to\infty} \frac{\langle \Delta r^2(t) \rangle}{6t} = \lim_{t\to\infty} \frac{1}{6Nt} \sum_j \Delta r_j^2$$

Derivation:

- Fick's law with ρ... local density or concentration: $\frac{\partial \rho(\mathbf{r},t)}{\partial t} = D\nabla^2 \rho(\mathbf{r},t)$
- Multiply by Δr^2 and integrate along $\mathbf{r}$: $\frac{\partial}{\partial t} \int \mathrm{d}\mathbf{r} \Delta r^2 \rho(\mathbf{r},t) = D \int \mathrm{d}\mathbf{r} \Delta r^2 \nabla^2 \rho(\mathbf{r},t)$
- Integrate by parts and assume ρ is normalized to 1: $\frac{\partial}{\partial t} \int \mathrm{d}\mathbf{r} \Delta r^2 \rho(\mathbf{r},t) = 2D \cdot d = 6D$
- Integral is an expectation value since ρ behaves like a probability: $\frac{\partial}{\partial t} \langle \Delta r^2(t) \rangle = 6D$
- $\langle \Delta r^2(t) \rangle$ is linear at long time scales: $\lim_{t\to\infty} \frac{\langle \Delta r^2(t) \rangle}{t} = 6D$
- $\langle \Delta r^2(t) \rangle$ is the average over the particles: $\lim_{t\to\infty} \frac{\sum_j \Delta r_j(t)^2}{Nt} = 6D$

We thus relate the miscroscopic dynamics ($\langle \Delta r^2(t) \rangle$) with a macroscopic property (D)

Diffusion

For periodic boundary conditions, we need the true atomic displacements without periodic wraparound (we need to "unfold the trajectory")! Alternative without unfolding: **Green-Kubo expression** based on the velocity autocorrelation function:

$$D = \frac{1}{3} \int_0^\infty \langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle \, dt$$

Diffusion

For periodic boundary conditions, we need the true atomic displacements without periodic wraparound (we need to "unfold the trajectory")! Alternative without unfolding: **Green-Kubo expression** based on the velocity autocorrelation function:

$$D = \frac{1}{3} \int_0^\infty \langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle \, dt$$

Derivation:

- Start with: $D = \lim_{t\to\infty} \frac{1}{6t} \langle \Delta r^2(t) \rangle$

Diffusion

For periodic boundary conditions, we need the true atomic displacements without periodic wraparound (we need to "unfold the trajectory")! Alternative without unfolding: **Green-Kubo expression** based on the velocity autocorrelation function:

$$D = \frac{1}{3} \int_0^\infty \langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle \, dt$$

Derivation:

- Start with: $D = \lim_{t\to\infty} \frac{1}{6t} \langle \Delta r^2(t) \rangle$
- Reformulate $\Delta \mathbf{r}(t)$ as $\int_0^t \mathrm{d}t' \mathbf{v}(t')$: $D = \lim_{t\to\infty} \frac{1}{6t} \int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t'' \langle \mathbf{v}(t') \mathbf{v}(t'') \rangle$

Diffusion

For periodic boundary conditions, we need the true atomic displacements without periodic wraparound (we need to "unfold the trajectory")! Alternative without unfolding: **Green-Kubo expression** based on the velocity autocorrelation function:

$$D = \frac{1}{3}\int_0^\infty \langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle \, dt$$

Derivation:

- Start with: $D = \lim_{t\to\infty} \frac{1}{6t} \langle \Delta r^2(t) \rangle$
- Reformulate $\Delta \mathbf{r}(t)$ as $\int_0^t \mathrm{d}t' \mathbf{v}(t')$: $D = \lim_{t\to\infty} \frac{1}{6t} \int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t'' \langle \mathbf{v}(t')\mathbf{v}(t'') \rangle$
- One can show that $\int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t'' \langle \mathbf{v}(t')\mathbf{v}(t'') \rangle = 2\int_0^\tau \mathrm{d}\tau (t-\tau) \langle \mathbf{v}(\tau)\mathbf{v}(0) \rangle$, which at $\lim_{t\to\infty}$ turns into $2t \int_0^\infty \mathrm{d}\tau \langle \mathbf{v}(\tau)\mathbf{v}(0) \rangle$: $D = \frac{1}{3}\int_0^\infty \mathrm{d}\tau \langle \mathbf{v}(\tau)\mathbf{v}(0) \rangle$

Diffusion

For periodic boundary conditions, we need the true atomic displacements without periodic wraparound (we need to "unfold the trajectory")! Alternative without unfolding: **Green-Kubo expression** based on the velocity autocorrelation function:

$$D = \frac{1}{3}\int_0^\infty \langle \mathbf{v}(t)\cdot\mathbf{v}(0)\rangle \, dt$$

Derivation:

- Start with: $D = \lim_{t\to\infty}\frac{1}{6t}\langle\Delta r^2(t)\rangle$
- Reformulate $\Delta\mathbf{r}(t)$ as $\int_0^t \mathrm{d}t'\mathbf{v}(t')$: $D = \lim_{t\to\infty}\frac{1}{6t}\int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t''\langle\mathbf{v}(t')\mathbf{v}(t'')\rangle$
- One can show that $\int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t''\langle\mathbf{v}(t')\mathbf{v}(t'')\rangle = 2\int_0^\tau \mathrm{d}\tau(t-\tau)\langle\mathbf{v}(\tau)\mathbf{v}(0)\rangle$, which at $\lim_{t\to\infty}$ turns into $2t\int_0^\infty \mathrm{d}\tau\langle\mathbf{v}(\tau)\mathbf{v}(0)\rangle$: $D = \frac{1}{3}\int_0^\infty \mathrm{d}\tau\langle\mathbf{v}(\tau)\mathbf{v}(0)\rangle$

Important (for Einstein or Green-Kubo relation): For very small box-sizes, we need to include a correction factor!

Shear viscosity

The shear viscosity η can be obtained via

$$\eta = \lim_{t\to\infty} \frac{1}{6TVt} \left\langle \sum_{x<y} \left[\sum_j m_j r_{jx}(t) v_{jy}(t) - \sum_j m_j r_{jx}(0) v_{jy}(0) \right]^2 \right\rangle$$

where $\sum_{x<y}$ denotes a sum over the three pairs of components xy, yz and zx. It characterizes the rate at which some component (e.g. y) of momentum diffuses in a perpendicular (x) direction. This expression is unfortunately unusable with periodic boundary conditions!

Shear viscosity

The shear viscosity η can be obtained via

$$\eta = \lim_{t\to\infty} \frac{1}{6TVt} \left\langle \sum_{x<y} \left[\sum_j m_j r_{jx}(t) v_{jy}(t) - \sum_j m_j r_{jx}(0) v_{jy}(0) \right]^2 \right\rangle$$

where $\sum_{x<y}$ denotes a sum over the three pairs of components xy, yz and zx. It characterizes the rate at which some component (e.g. y) of momentum diffuses in a perpendicular (x) direction. This expression is unfortunately unusable with periodic boundary conditions!

Alternative: Green-Kubo expression based on the pressure tensor autocorrelation function:

$$\eta = \frac{V}{3T} \int_0^\infty \left\langle \sum_{x<y} P_{xy}(t) P_{xy}(0) \right\rangle dt$$

$$\text{with } P_{xy} = \frac{1}{V} \left[\sum_j m_j v_{jx} v_{jy} + \frac{1}{2} \sum_{i\neq j} r_{ijx} f_{ijy} \right]$$

Thermal conductivity

The thermal conductivity λ can be obtained via

$$\lambda = \lim_{t\to\infty} \frac{1}{6T^2Vt} \left\langle \sum_x \left[\sum_j r_{jx}(t) e_j(t) - \sum_j r_{jx}(0) e_j(0) \right]^2 \right\rangle$$

where $e_j = \frac{1}{2} m v_j^2 + \frac{1}{2} \sum_{i \neq j} u(r_{ij}) - \langle e \rangle$ is the instantaneous excess energy of atom j, $\langle e \rangle$ is the mean energy and $\sum_x$ is a sum over vector components. This expression is unfortunately unusable with periodic boundary conditions!

Thermal conductivity

The thermal conductivity λ can be obtained via

$$\lambda = \lim_{t\to\infty} \frac{1}{6T^2Vt} \left\langle \sum_x \left[\sum_j r_{jx}(t) e_j(t) - \sum_j r_{jx}(0) e_j(0) \right]^2 \right\rangle$$

where $e_j = \frac{1}{2} m v_j^2 + \frac{1}{2} \sum_{i\neq j} u(r_{ij}) - \langle e \rangle$ is the instantaneous excess energy of atom j, $\langle e \rangle$ is the mean energy and $\sum_x$ is a sum over vector components. This expression is unfortunately unusable with periodic boundary conditions!

Alternative: Green-Kubo expression based on the heat flux autocorrelation function:

$$\lambda = \frac{V}{3T^2} \int_0^\infty \langle \mathbf{S}(t) \mathbf{S}(0) \rangle \, dt$$

$$\text{with } \mathbf{S} = \frac{1}{V} \left[\sum_j e_j \mathbf{v}_j + \frac{1}{2} \sum_{i\neq j} \mathbf{r}_{ij} (\mathbf{f}_{ij} \cdot \mathbf{v}_j) \right]$$

Dynamic properties

(Demonstration of dynamic properties of different systems)

Measuring a property: Sources of error

Sources of error:

- Finite-size effects, interaction cutoffs, approximations to E_{pot}, numerical integration (systematic, reproducible errors)

Measuring a property: Sources of error

Sources of error:

- Finite-size effects, interaction cutoffs, approximations to E_{pot}, numerical integration (systematic, reproducible errors)
- Inclusion of un-equilibrated data (systematic bias)

Measuring a property: Sources of error

Sources of error:

- Finite-size effects, interaction cutoffs, approximations to E_{pot}, numerical integration (systematic, reproducible errors)
- Inclusion of un-equilibrated data (systematic bias)
- Inadequate sampling errors due to rare events not sampled often enough (often not well reproducible)

Measuring a property: Sources of error

Sources of error:

- Finite-size effects, interaction cutoffs, approximations to E_{pot}, numerical integration (systematic, reproducible errors)
- Inclusion of un-equilibrated data (systematic bias)
- Inadequate sampling errors due to rare events not sampled often enough (often not well reproducible)
- Random fluctuations in the measurements lead to a statistical error (upon which confidence estimates like error-bars are built)

Measuring a property: Collective variables and central limit theorem

- Given a trajectory (coordinates and velocities of N atoms over T timesteps), we calculate the value of a collective variable from the positions of all atoms at a given T.

$$\begin{array}{lcccccccc}
\text{Trajectory:} & \vec{x}(t_1) & \vec{x}(t_2) & \vec{x}(t_3) & \vec{x}(t_4) & \vec{x}(t_5) & \vec{x}(t_6) & \vec{x}(t_7) & \vec{x}(t_8) \\
 & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
 & A_1 & A_2 & A_3 & A_4 & A_5 & A_6 & A_7 & A_8
\end{array}$$

Random variables $\longrightarrow$ Gaussian probability distribution for large numbers

Measuring a property: Collective variables and central limit theorem

- Given a trajectory (coordinates and velocities of N atoms over T timesteps), we calculate the value of a collective variable from the positions of all atoms at a given T.
- Central limit theorem: Treat values of the collective variable at different times as independent random variables

$$\begin{array}{lcccccccc} \text{Trajectory:} & \vec{x}(t_1) & \vec{x}(t_2) & \vec{x}(t_3) & \vec{x}(t_4) & \vec{x}(t_5) & \vec{x}(t_6) & \vec{x}(t_7) & \vec{x}(t_8) \\ & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ & A_1 & A_2 & A_3 & A_4 & A_5 & A_6 & A_7 & A_8 \end{array}$$

Random variables $\longrightarrow$ Gaussian probability distribution for large numbers

Measuring a property: Collective variables and central limit theorem

- Given a trajectory (coordinates and velocities of N atoms over T timesteps), we calculate the value of a collective variable from the positions of all atoms at a given T.
- Central limit theorem: Treat values of the collective variable at different times as independent random variables
- If we run our trajectories longer, we will have more certainty in our estimates of the ensemble averages

Trajectory: $\vec{x}(t_1)$ $\vec{x}(t_2)$ $\vec{x}(t_3)$ $\vec{x}(t_4)$ $\vec{x}(t_5)$ $\vec{x}(t_6)$ $\vec{x}(t_7)$ $\vec{x}(t_8)$

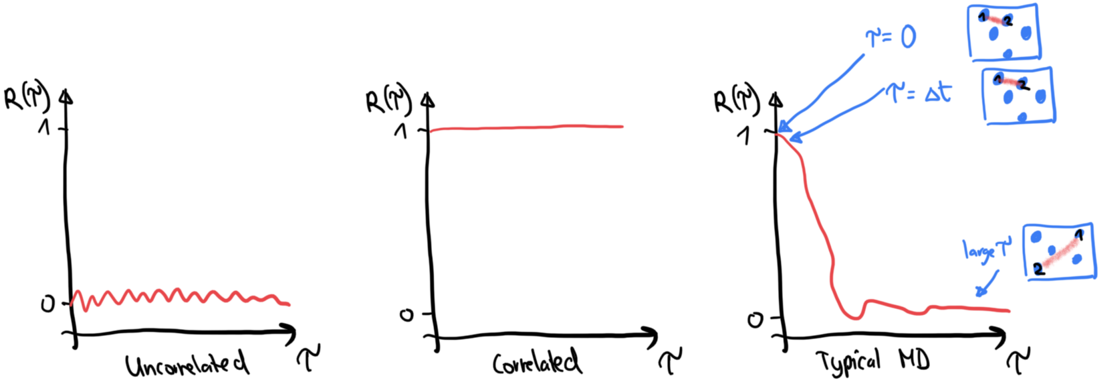
$\downarrow$ $\downarrow$ $\downarrow$ $\downarrow$ $\downarrow$ $\downarrow$ $\downarrow$ $\downarrow$

A_1 A_2 A_3 A_4 A_5 A_6 A_7 A_8

Random variables $\rightarrow$ Gaussian probability distribution for large numbers

Measuring a property: The autocorrelation function

$$R(\tau) = \frac{\langle (A_t - \langle A \rangle)(A_{t+\tau} - \langle A \rangle) \rangle}{\langle (\delta A)^2 \rangle}$$



We thus cannot use the central limit theorem to analyse trajectory data directly!

Measuring a property: Estimating statistical error

To average over measurements from a fluctuating property, we have to keep in mind that the measurements are correlated for consecutive timesteps:

- Average property $\langle A \rangle = \frac{1}{M} \sum_\mu A_\mu$ with variance (if A_μ are independent) $\sigma^2(A) = \frac{1}{M} \sum_\mu (A_\mu - \langle A \rangle)^2 = \langle A^2 \rangle - \langle A \rangle^2$

Measuring a property: Estimating statistical error

To average over measurements from a fluctuating property, we have to keep in mind that the measurements are correlated for consecutive timesteps:

- Average property $\langle A \rangle = \frac{1}{M} \sum_\mu A_\mu$ with variance (if A_μ are independent) $\sigma^2(A) = \frac{1}{M} \sum_\mu (A_\mu - \langle A \rangle)^2 = \langle A^2 \rangle - \langle A \rangle^2$
- Obtain the variance of the mean as $\sigma^2(\langle A \rangle) = \frac{1}{M} \sigma^2(A)$

Measuring a property: Estimating statistical error

To average over measurements from a fluctuating property, we have to keep in mind that the measurements are correlated for consecutive timesteps:

- Average property $\langle A \rangle = \frac{1}{M} \sum_\mu A_\mu$ with variance (if A_μ are independent) $\sigma^2(A) = \frac{1}{M} \sum_\mu (A_\mu - \langle A \rangle)^2 = \langle A^2 \rangle - \langle A \rangle^2$
- Obtain the variance of the mean as $\sigma^2(\langle A \rangle) = \frac{1}{M}\sigma^2(A)$
- However, A_μ are usually not independent if they stem from an MD simulation, so that $\sigma^2(\langle A \rangle)$ is underestimated, since the number of independent measurements is much less than M.

Measuring a property: Estimating statistical error

To average over measurements from a fluctuating property, we have to keep in mind that the measurements are correlated for consecutive timesteps:

- Average property $\langle A \rangle = \frac{1}{M} \sum_\mu A_\mu$ with variance (if A_μ are independent) $\sigma^2(A) = \frac{1}{M} \sum_\mu (A_\mu - \langle A \rangle)^2 = \langle A^2 \rangle - \langle A \rangle^2$
- Obtain the variance of the mean as $\sigma^2(\langle A \rangle) = \frac{1}{M} \sigma^2(A)$
- However, A_μ are usually not independent if they stem from an MD simulation, so that $\sigma^2(\langle A \rangle)$ is underestimated, since the number of independent measurements is much less than M.
- To deal with this correctly we would need to follow how soon A_μ becomes de-correlated from A_0. In practice we can instead often get away with block averaging

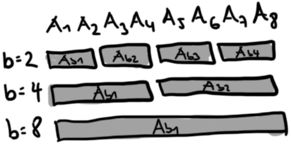
Measuring a property: Block averaging

- Make a series of successive block sizes $b = 1, 2, 4, \ldots$ For each b, compute $\sigma^2(\langle A \rangle_b) = \frac{1}{M_b - 1} \sum_b (A_b^2 - \langle A \rangle_b^2)$ with A_b a typical block average and $\langle A \rangle_b$ the overall average

Measuring a property: Block averaging

- Make a series of successive block sizes $b = 1, 2, 4, \ldots$ For each b, compute $\sigma^2(\langle A \rangle_b) = \frac{1}{M_b - 1} \sum_b (A_b^2 - \langle A \rangle_b^2)$ with A_b a typical block average and $\langle A \rangle_b$ the overall average
- $\sigma^2(\langle A \rangle_b)$ will reach a plateau if the total measurement period exceeds the longest correlation time, this is our best estimate for $\langle (\delta A)^2 \rangle$

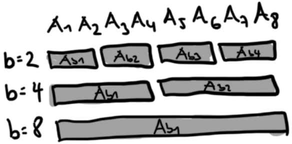
good estimate for $\langle (\delta A)^2 \rangle$ where A_b become uncorrelated



Measuring a property: Block averaging

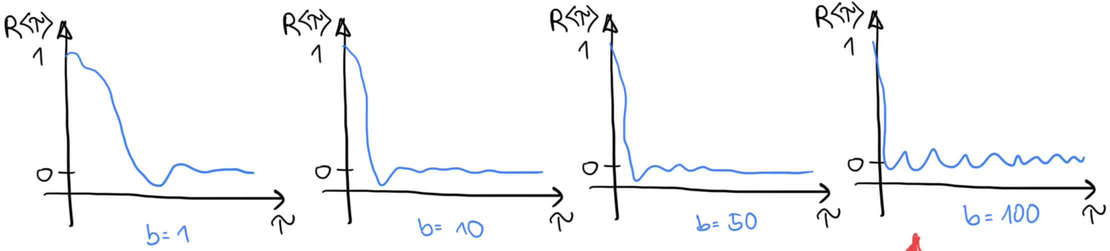
- Make a series of successive block sizes $b = 1, 2, 4, \ldots$ For each b, compute $\sigma^2(\langle A\rangle_b) = \frac{1}{M_b - 1}\sum_b (A_b^2 - \langle A\rangle_b^2)$ with A_b a typical block average and $\langle A\rangle_b$ the overall average
- $\sigma^2(\langle A\rangle_b)$ will reach a plateau if the total measurement period exceeds the longest correlation time, this is our best estimate for $\langle(\delta A)^2\rangle$
- The onset of the plateau gives an indication to the extent to which the samples are correlated. If there is no plateau, the variance estimate is unreliable

good estimate for $\langle(\delta A)^2\rangle$
where A_b become uncorrelated



Measuring a property: Block averaging

The autocorrelation function of block-averaged data shows that the data is now uncorrelated, we can therefore use the central limit theorem to analyse our data and obtain meaningful errorbars



Data becomes uncorrelated and can be analyzed using the central limit theorem

$$\rightarrow \langle(\delta A)^2\rangle = \frac{1}{N_b-1}\sum_b \left(A_b - \langle A\rangle\right)^2$$