Stat-mech on a computer: Simulating and sampling

MSE 421 - Ceriotti

Averages in the canonical ensemble

- The canonical ensemble describes a system in contact with a thermal reservoir at temperature $T = 1/k_B\beta$
- Averages of observables can be computed as integrals over the canonical Boltzmann distribution

$$\langle A\rangle = \int \mathrm{d}\mathbf{q}\mathrm{d}\mathbf{p}\, P\left(\mathbf{p},\mathbf{q}\right) A\left(\mathbf{p},\mathbf{q}\right) \qquad \langle A\rangle = \sum_\nu A_\nu P_\nu$$

- The goal of a simulation is to perform these averages/integrals by sampling **the states that have a non-negligible probability**

Atoms and molecules

- Classical dynamics of a system of atoms with positions $\mathbf{q}$ and momenta $\mathbf{p}$, described by the Hamiltonian

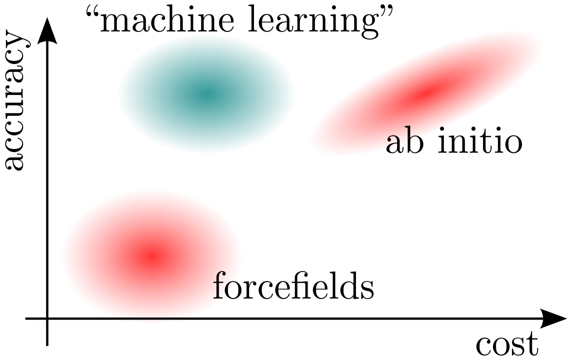
$$H\left(\mathbf{p},\mathbf{q}\right)=K\left(\mathbf{p}\right)+V\left(\mathbf{q}\right)=\sum_i \frac{\mathbf{p}_i^2}{2m_i}+V\left(\mathbf{q}\right)$$

  - We assume V is given (force fields, electronic structure calculation)
  - In all cases, strictly Born-Oppenheimer
  - Nuclei are treated as classical, distinguishable particles
- The configuration and momentum parts of $P\left(\mathbf{p},\mathbf{q}\right)$ factor out, and the momentum part $P\left(\mathbf{p}\right)$ is trivial

$$P\left(\mathbf{p},\mathbf{q}\right)=e^{-\beta H(\mathbf{p},\mathbf{q})}=P\left(\mathbf{p}\right)\cdot P\left(\mathbf{q}\right)=\frac{e^{-\beta\sum_i \frac{\mathbf{p}_i^2}{2m_i}}}{\int \mathrm{d}^{3N}\mathbf{p}\, e^{-\beta\sum_i \frac{\mathbf{p}_i^2}{2m_i}}}\cdot\frac{e^{-\beta V(\mathbf{q})}}{\int \mathrm{d}^{3N}\mathbf{q}\, e^{-\beta V(\mathbf{q})}}$$

Potentials, potentials, potentials!

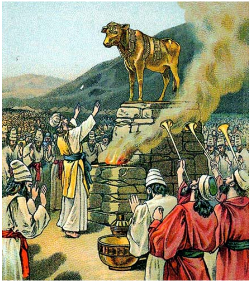
- Actually, constructing a good model for $V(\mathbf{q})$ is the most crucial step! Accuracy/cost trade-off



Potentials, potentials, potentials!

- Actually, constructing a good model for $V(\mathbf{q})$ is the most crucial step! Accuracy/cost trade-off
- Empirical forcefields try to represent interactions with "simple" functional forms. Time-consuming and error-prone fitting procedure, but then easy to use and inexpensive. A must for large systems

$$V(\mathbf{q}) \sim \underbrace{\sum_{ij} \frac{z_i z_j}{|\mathbf{q}_i - \mathbf{q}_j|}}_{\text{electrostatics}} + \underbrace{\sum_{\text{bonds}} k_i (\mathbf{q}_i - \mathbf{q}_i')^2}_{\text{bonded terms}} - \underbrace{\sum_{ij} \frac{A_{ij}}{|\mathbf{q}_i - \mathbf{q}_j|^6}}_{\text{dispersion}} + \ldots$$

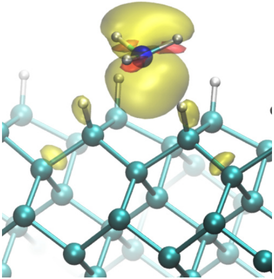


Potentials, potentials, potentials!

- Actually, constructing a good model for $V(\mathbf{q})$ is the most crucial step! Accuracy/cost trade-off

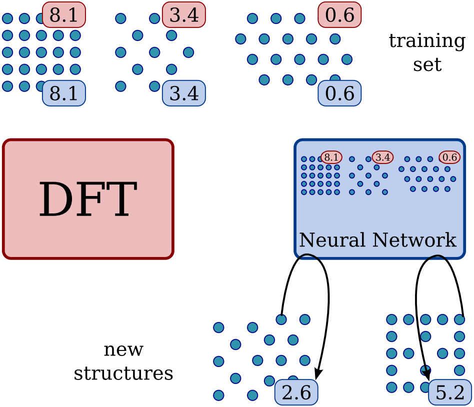
- Ab initio calculations offer transferability and predictive power by solving on-the-fly the electronic structure problem: expensive! Typically limited to few 100's atoms

$$\hat{H}(\mathbf{q})\,\Psi_{\text{electrons}} = E\Psi_{\text{electrons}} \quad \rightarrow \quad V(\mathbf{q}) \equiv E$$



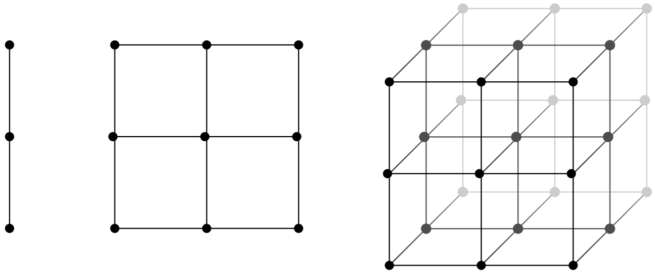
Potentials, potentials, potentials!

- Actually, constructing a good model for $V(\mathbf{q})$ is the most crucial step! Accuracy/cost trade-off
- Recent developments try to "machine-learn" ab initio energies, trying to get nearly as accurate with much less effort



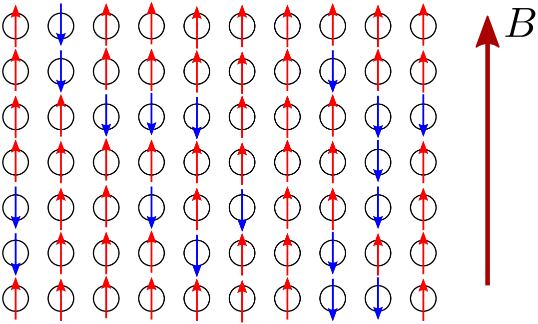
The curse of dimensionality

- Statistical averaging is basically just integration. Should be easy do do numerically....
- Performing the integral by quadrature is impractical, because the number of grid points grows **exponentially** with dimensionality



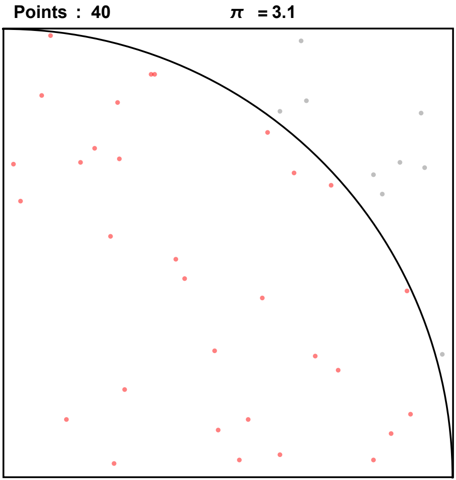
Sampling by enumeration

- Now think about computing averages for a spin lattice with N sites. Enumeration of all possible spin states also grows exponentially: 2^N possible configurations must be traversed to cover all possibilities!
- What saves us is (1) that often just a small fraction of microstates have $P_\nu \gg 0$ and (2) that we don't need to sample *all* states, only a representative fraction



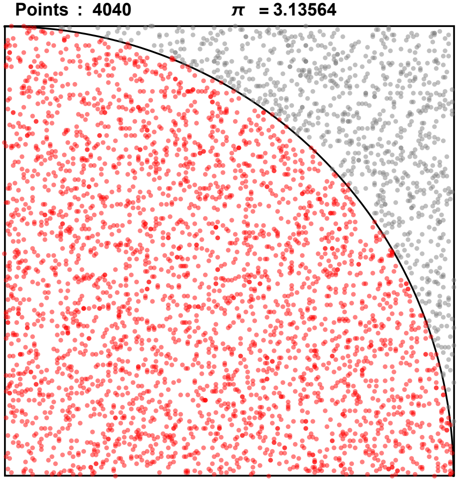
Monte Carlo integration

- To compute integrals without a grid, we could just generate uniformly distributed random values, and compute the value of the function there
- For instance, one can use this to compute the area of a circle relative to an enclosing square by tossing random values on a square, and counting how many have $x^2 + y^2 < 1$ [the fraction converges to $\pi/4$]



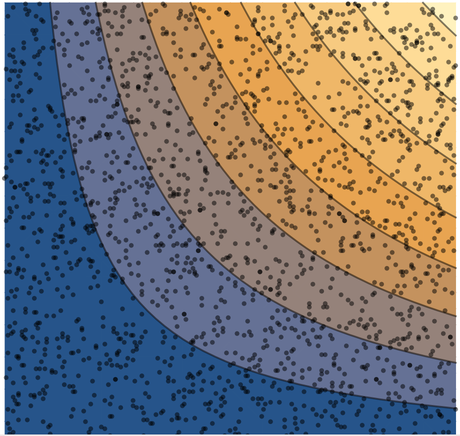
Monte Carlo integration

- To compute integrals without a grid, we could just generate uniformly distributed random values, and compute the value of the function there
- For instance, one can use this to compute the area of a circle relative to an enclosing square by tossing random values on a square, and counting how many have $x^2 + y^2 < 1$ [the fraction converges to $\pi/4$]



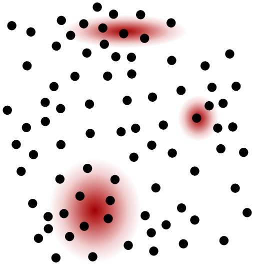
Integration by importance sampling

- We can compute $\int f(x)\,\mathrm{d}x$ as area $\times \sum_i f(x_i)/N$
- One could think of using the same trick: randomly-distributed points $\mathbf{q}_i$, and computing $\langle A\rangle \approx \sum_i P(\mathbf{q}_i)\,A(\mathbf{q}_i)\,/\sum_i P(\mathbf{q}_i)$. But $P(\mathbf{q})$ is often very irregular and most points do not contribute to the integral
- Ideally, one would want to generate $\mathbf{q}_i$'s distributed as $P(\mathbf{q})$, and compute $\langle A\rangle \approx \frac{1}{M}\sum_i A(\mathbf{q}_i)$. This is **importance sampling**.



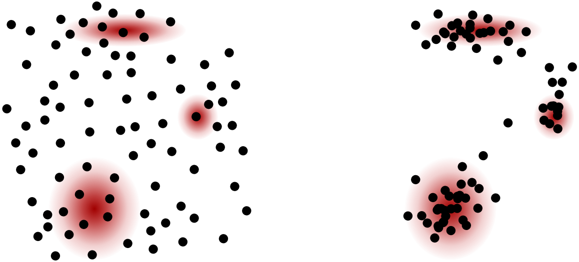
Integration by importance sampling

- We can compute $\int f(x)\,\mathrm{d}x$ as area $\times \sum_i f(x_i)/N$
- One could think of using the same trick: randomly-distributed points $\mathbf{q}_i$, and computing $\langle A\rangle \approx \sum_i P(\mathbf{q}_i)\,A(\mathbf{q}_i)/\sum_i P(\mathbf{q}_i)$. But $P(\mathbf{q})$ is often very irregular and most points do not contribute to the integral
- Ideally, one would want to generate $\mathbf{q}_i$'s distributed as $P(\mathbf{q})$, and compute $\langle A\rangle \approx \frac{1}{M}\sum_i A(\mathbf{q}_i)$. This is **importance sampling**.



Integration by importance sampling

- We can compute $\int f(x)\,\mathrm{d}x$ as area $\times \sum_i f(x_i)\,/N$
- One could think of using the same trick: randomly-distributed points $\mathbf{q}_i$, and computing $\langle A\rangle \approx \sum_i P(\mathbf{q}_i)\,A(\mathbf{q}_i)\,/\sum_i P(\mathbf{q}_i)$. But $P(\mathbf{q})$ is often very irregular and most points do not contribute to the integral
- Ideally, one would want to generate $\mathbf{q}_i$'s distributed as $P(\mathbf{q})$, and compute $\langle A\rangle \approx \frac{1}{M}\sum_i A(\mathbf{q}_i)$. This is **importance sampling**.



Ergodicity of sampling

- Underlying assumption behind all of this is **ergodicity**:

$$\lim_{M\to\infty}\frac{1}{M}\sum_i A(\mathbf{q}_i) = \int \mathrm{d}\mathbf{q}\, P(\mathbf{q})\, A(\mathbf{q})$$

  - This is an additional condition, that is very hard to prove in practice.
  - We always have a finite M, so the question is: *how much* ergodic is the trajectory? How can we “measure” ergodicity?
  - As we saw, this is closely related to the statistical error on $\langle A\rangle$
- The key tool to assess the ergodicity of a trajectory for computing $\langle A\rangle$ is the autocorrelation function

$$c_{AA}(t) = \langle A(0)\, A(t)\rangle$$

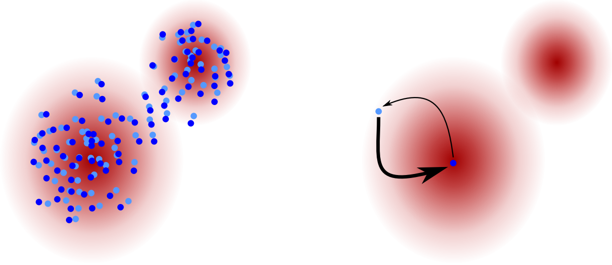
Using Monte Carlo for importance sampling

- How to generate a sequence of points distributed as $P(\mathbf{q})$?
  - Monte Carlo moves $\mathbf{q} \to \mathbf{q}'$ that generate a new point given a previous point consistent with P.
  - The move (typically) contains some randomness: probability $p(\mathbf{q} \to \mathbf{q}')$
- **Necessary** (& "sufficient") condition: P invariant under the move

$$\int d\mathbf{q}\, P(\mathbf{q})\, p(\mathbf{q} \to \mathbf{q}') = P(\mathbf{q}')$$

- **Sufficient** condition: **detailed balance**

$$P(\mathbf{q})\, p(\mathbf{q} \to \mathbf{q}') = P(\mathbf{q}')\, p(\mathbf{q}' \to \mathbf{q})$$



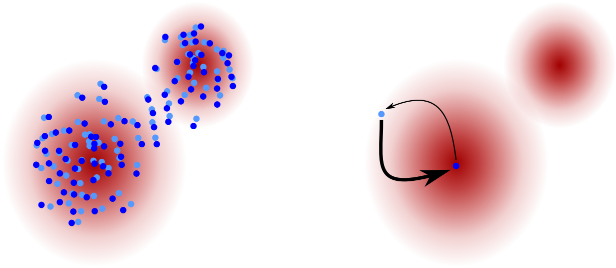
Using Monte Carlo for importance sampling

- How to generate a sequence of points distributed as $P(\mathbf{q})$?
  - Monte Carlo moves $\mathbf{q} \rightarrow \mathbf{q}'$ that generate a new point given a previous point consistent with P.
  - The move (typically) contains some randomness: probability $p(\mathbf{q} \rightarrow \mathbf{q}')$
- **Necessary** (& "sufficient") condition: P invariant under the move

$$\int d\mathbf{q}\, P(\mathbf{q})\, p(\mathbf{q} \rightarrow \mathbf{q}') = P(\mathbf{q}')$$

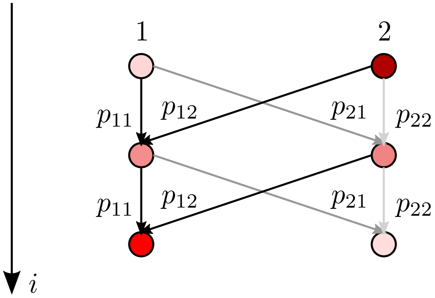
- **Sufficient** condition: **detailed balance**

$$p(\mathbf{q} \rightarrow \mathbf{q}') / p(\mathbf{q}' \rightarrow \mathbf{q}) = P(\mathbf{q}') / P(\mathbf{q}) = e^{-\beta\left[V(\mathbf{q}')-V(\mathbf{q})\right]}$$



"Equilibration"

- Detailed balance guarantees that the target distribution $P(\mathbf{q})$ is *conserved*. What if we start from the wrong distribution? Example: two-state problem, with equilibrium distribution $P_1 = 0.1$, $P_2 = 0.9$.
  1. Take $p(1 \rightarrow 2) = 0.45$, $p(2 \rightarrow 1) = 0.05$, $p(1 \rightarrow 1) = 0.55$, $p(2 \rightarrow 2) = 0.95$
  2. Say we are in state 1 with prob. x_1 and in 2 with prob. x_2. Let $p_{ij} = p(j \rightarrow i)$. One MC step is equivalent to a matrix-vector product $\mathbf{x}' \leftarrow \mathbf{p}\,\mathbf{x}$
  3. Repeat application of $\mathbf{p}$ corresponds to $\mathbf{p}^N$, that we can characterize in terms of the N-th powers of its eigenvalues. Let's see what are the eigenvalues and eigenvectors of $\mathbf{p}$



"Equilibration"

- Detailed balance guarantees that the target distribution $P(\mathbf{q})$ is *conserved*. What if we start from the wrong distribution? Example: two-state problem, with equilibrium distribution $P_1 = 0.1$, $P_2 = 0.9$.
  1. Take $p(1 \rightarrow 2) = 0.45$, $p(2 \rightarrow 1) = 0.05$, $p(1 \rightarrow 1) = 0.55$, $p(2 \rightarrow 2) = 0.95$
  2. Say we are in state 1 with prob. x_1 and in 2 with prob. x_2. Let $p_{ij} = p(j \rightarrow i)$. One MC step is equivalent to a matrix-vector product $\mathbf{x}' \leftarrow \mathbf{p}\,\mathbf{x}$
  3. Repeat application of $\mathbf{p}$ corresponds to $\mathbf{p}^N$, that we can characterize in terms of the N-th powers of its eigenvalues. Let's see what are the eigenvalues and eigenvectors of $\mathbf{p}$

$$\mathbf{p} = \begin{pmatrix} 0.55 & 0.05 \\ 0.45 & 0.95 \end{pmatrix} \rightarrow \lambda_a = 0.5;\ \mathbf{u}_a = \begin{pmatrix} 0.5 \\ -0.5 \end{pmatrix} \quad \lambda_b = 1;\ \mathbf{u}_b = \begin{pmatrix} 0.1 \\ 0.9 \end{pmatrix}$$

"Equilibration"

- Detailed balance guarantees that the target distribution $P(\mathbf{q})$ is *conserved*. What if we start from the wrong distribution? Example: two-state problem, with equilibrium distribution $P_1 = 0.1$, $P_2 = 0.9$.
  1. Take $p(1\rightarrow 2) = 0.45$, $p(2\rightarrow 1) = 0.05$, $p(1\rightarrow 1) = 0.55$, $p(2\rightarrow 2) = 0.95$
  2. Say we are in state 1 with prob. x_1 and in 2 with prob. x_2. Let $p_{ij} = p(j\rightarrow i)$. One MC step is equivalent to a matrix-vector product $\mathbf{x}' \leftarrow \mathbf{p}\,\mathbf{x}$
  3. Repeat application of $\mathbf{p}$ corresponds to $\mathbf{p}^N$, that we can characterize in terms of the N-th powers of its eigenvalues. Let's see what are the eigenvalues and eigenvectors of $\mathbf{p}$

$$\mathbf{p} = \begin{pmatrix} 0.55 & 0.05 \\ 0.45 & 0.95 \end{pmatrix} \rightarrow \lambda_a = 0.5;\ \mathbf{u}_a = \begin{pmatrix} 0.5 \\ -0.5 \end{pmatrix} \quad \lambda_b = 1;\ \mathbf{u}_b = \begin{pmatrix} 0.1 \\ 0.9 \end{pmatrix}$$

- Repeated application of $\mathbf{p}$ will progressively kill off the component $\mathbf{u}_a$, and let only $\mathbf{u}_b$ -- the *stationary* distribution

“Equilibration”

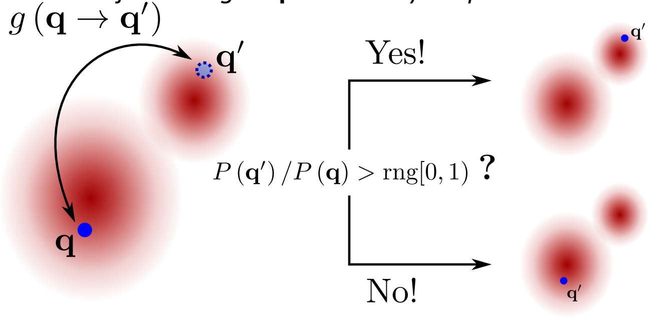
- Detailed balance guarantees that the target distribution $P(\mathbf{q})$ is *conserved*. What if we start from the wrong distribution? Example: two-state problem, with equilibrium distribution $P_1 = 0.1$, $P_2 = 0.9$.
  1. Take $p(1 \to 2) = 0.45$, $p(2 \to 1) = 0.05$, $p(1 \to 1) = 0.55$, $p(2 \to 2) = 0.95$
  2. Say we are in state 1 with prob. x_1 and in 2 with prob. x_2. Let $p_{ij} = p(j \to i)$. One MC step is equivalent to a matrix-vector product $\mathbf{x}' \leftarrow \mathbf{p}\,\mathbf{x}$
  3. Repeat application of $\mathbf{p}$ corresponds to $\mathbf{p}^N$, that we can characterize in terms of the N-th powers of its eigenvalues. Let's see what are the eigenvalues and eigenvectors of $\mathbf{p}$

$$\mathbf{p} = \begin{pmatrix} 0.8 & 0.3 & 0.5 \\ 0.1 & 0.3 & 0.1 \\ 0.1 & 0.4 & 0.4 \end{pmatrix}$$

- One can build a transition matrix that violates detailed balance, still it has a well-defined stationary distribution!

Accept/reject: Metropolis moves

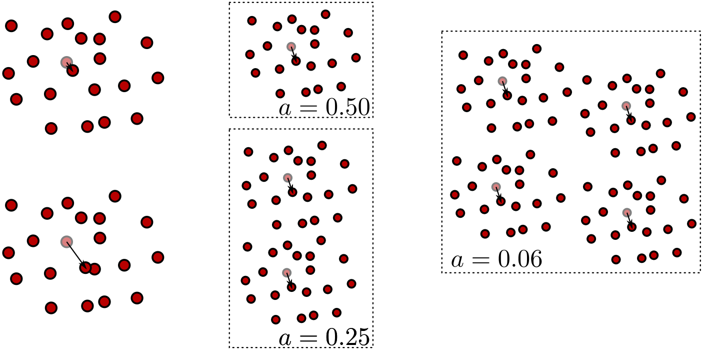
- A simple way to generate moves consistent with detailed balance. Split the move in two steps: generation/acceptance
  $p\left(\mathbf{q}\rightarrow\mathbf{q}'\right)=g\left(\mathbf{q}\rightarrow\mathbf{q}'\right)a\left(\mathbf{q}\rightarrow\mathbf{q}'\right)$
  - Generate a new configuration with a criterion that is symmetric:
    $g\left(\mathbf{q}\rightarrow\mathbf{q}'\right)=g\left(\mathbf{q}'\rightarrow\mathbf{q}\right)$
  - Accept the new configuration with probability
    $a\left(\mathbf{q}\rightarrow\mathbf{q}'\right)=\min\left[1,P\left(\mathbf{q}'\right)/P\left(\mathbf{q}\right)\right]=\min\left[1,e^{-\beta\left[V\left(\mathbf{q}'\right)-V\left(\mathbf{q}\right)\right]}\right].$
  - In case of reject *add again* **q** *to the list of samples!*



Metropolis et al., J. Chem. Phys 1953

Problems with Monte Carlo

- Must find a balance between having high acceptance, and making things happen!
- Fundamental problem: size-extensive changes in energy upon "global" moves reduce dramatically the acceptance
- One should develop "smart" moves, hard to generalize to arbitrary problems



Sampling by molecular dynamics

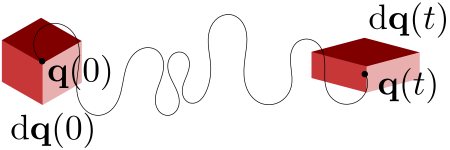
- An alternative: use Hamilton's equations to generate a trajectory

$$\dot{\mathbf{p}} = -\partial V/\partial \mathbf{q}, \qquad \dot{\mathbf{q}} = \mathbf{p}/m$$

- Does not obey detailed balance (but a generalization of it).
- Does fulfill the *necessary*, "global" condition: could prove cleanly with a Liouville operator formalism. The ingredients are:
    - Conservation of probability density

$$\frac{dP}{dt} \propto e^{-\beta H}\frac{dH}{dt}, \qquad \frac{dH}{dt} = \frac{\partial H}{\partial \mathbf{p}}\cdot\dot{\mathbf{p}} + \frac{\partial H}{\partial \mathbf{q}}\cdot\dot{\mathbf{q}} = -\frac{\partial H}{\partial \mathbf{p}}\cdot\frac{\partial H}{\partial \mathbf{q}} + \frac{\partial H}{\partial \mathbf{q}}\cdot\frac{\partial H}{\partial \mathbf{p}} = 0$$

    - Conservation of phase-space volume



Integrating the equations of motion

- MD requires discrete time step *dt* integration of Hamilton's equations
- Many integrators exist. My favourite: symmetric-split velocity Verlet
  - Time-reversible, **symplectic**, simple implementation, easy to extend
  - Monitor total energy fluctuations and drift to get a grip of the integration error

$$\mathbf{p} \leftarrow \mathbf{p} - \frac{\partial V}{\partial \mathbf{q}} \frac{dt}{2}$$

$$\mathbf{q} \leftarrow \mathbf{q} + \frac{\mathbf{p}}{m} dt$$

$$\mathbf{p} \leftarrow \mathbf{p} - \frac{\partial V}{\partial \mathbf{q}} \frac{dt}{2}$$

- Integration error leads to (small) departure from the (micro)canonical ensemble, but no problem with system size: MD assumes the error is split evenly across the system

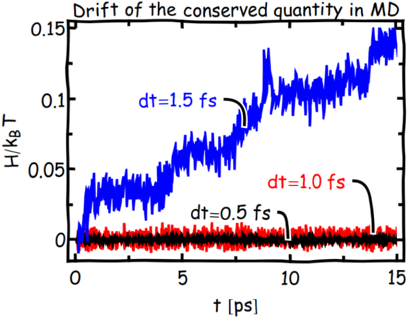
Integrating the equations of motion

- MD requires discrete time step *dt* integration of Hamilton's equations
- Many integrators exist. My favourite: symmetric-split velocity Verlet
  - Time-reversible, **symplectic**, simple implementation, easy to extend
  - Monitor total energy fluctuations and drift to get a grip of the integration error

$$\mathbf{p} \leftarrow \mathbf{p} - \frac{\partial V}{\partial \mathbf{q}} \frac{dt}{2}$$

$$\mathbf{q} \leftarrow \mathbf{q} + \frac{\mathbf{p}}{m} dt$$

$$\mathbf{p} \leftarrow \mathbf{p} - \frac{\partial V}{\partial \mathbf{q}} \frac{dt}{2}$$



- Integration error leads to (small) departure from the (micro)canonical ensemble, but no problem with system size: MD assumes the error is split evenly across the system

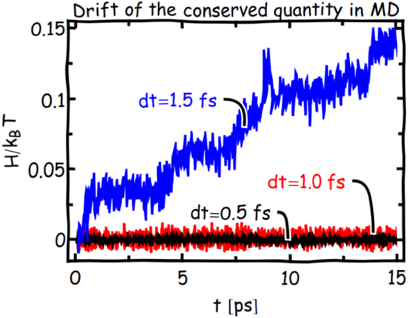
Integrating the equations of motion

- MD requires discrete time step *dt* integration of Hamilton's equations
- Many integrators exist. My favourite: symmetric-split velocity Verlet
  - Time-reversible, **symplectic**, simple implementation, easy to extend
  - Monitor total energy fluctuations and drift to get a grip of the integration error

$$\mathbf{p} \leftarrow \mathbf{p} - \frac{\partial V}{\partial \mathbf{q}} \frac{dt}{2}$$

$$\mathbf{q} \leftarrow \mathbf{q} + \frac{\mathbf{p}}{m} dt$$

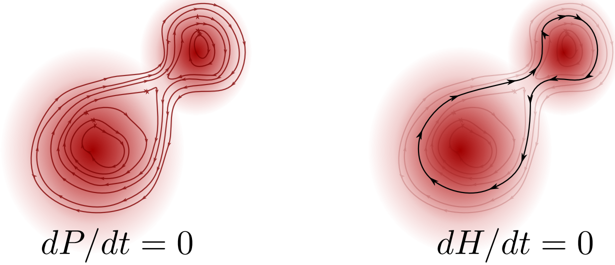
$$\mathbf{p} \leftarrow \mathbf{p} - \frac{\partial V}{\partial \mathbf{q}} \frac{dt}{2}$$



- Integration error leads to (small) departure from the (micro)canonical ensemble, but no problem with system size: MD assumes the error is split evenly across the system

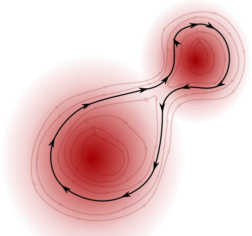
Conserved energy in Hamiltonian MD

- The probability distribution is conserved *globally*, i.e. if we started many trajectories from the canonical ensemble, we would sample the whole of it
- However, individual trajectories conserve the total energy $H(\mathbf{p}, \mathbf{q})$, and so only sample partially the canonical distribution
- We have an **ergodicity** problem with MD



Constant-temperature molecular dynamics

- Plain, Hamiltonian MD conserves energy and is not ergodic
- Must modify the dynamics to model the interaction with a heat bath (**thermostats**)
- Andersen thermostat: randomize atomic velocities every now and then
  - Exploits factorization of the canonical partition function
  - Simple, physically sound and effective

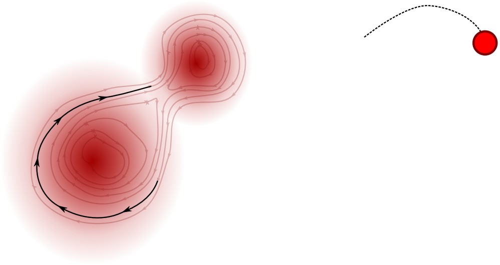


$$dH/dt = 0$$

Andersen, J. Chem. Phys. 1980

Constant-temperature molecular dynamics

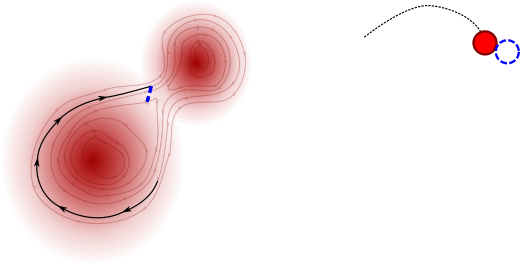
- Plain, Hamiltonian MD conserves energy and is not ergodic
- Must modify the dynamics to model the interaction with a heat bath (**thermostats**)
- Andersen thermostat: randomize atomic velocities every now and then
  - Exploits factorization of the canonical partition function
  - Simple, physically sound and effective



Andersen, J. Chem. Phys. 1980

Constant-temperature molecular dynamics

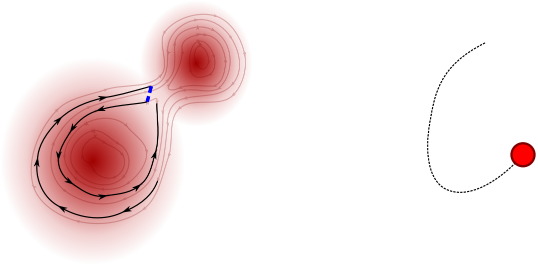
- Plain, Hamiltonian MD conserves energy and is not ergodic
- Must modify the dynamics to model the interaction with a heat bath (**thermostats**)
- Andersen thermostat: randomize atomic velocities every now and then
    - Exploits factorization of the canonical partition function
    - Simple, physically sound and effective



Andersen, J. Chem. Phys. 1980

Constant-temperature molecular dynamics

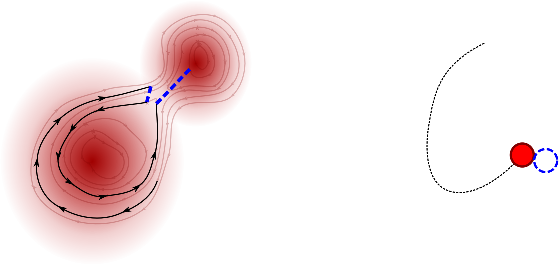
- Plain, Hamiltonian MD conserves energy and is not ergodic
- Must modify the dynamics to model the interaction with a heat bath (**thermostats**)
- Andersen thermostat: randomize atomic velocities every now and then
  - Exploits factorization of the canonical partition function
  - Simple, physically sound and effective



Andersen, J. Chem. Phys. 1980

Constant-temperature molecular dynamics

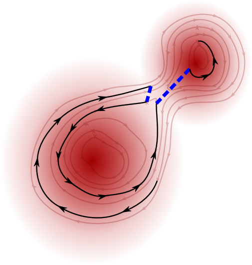
- Plain, Hamiltonian MD conserves energy and is not ergodic
- Must modify the dynamics to model the interaction with a heat bath (**thermostats**)
- Andersen thermostat: randomize atomic velocities every now and then
  - Exploits factorization of the canonical partition function
  - Simple, physically sound and effective



Andersen, J. Chem. Phys. 1980

Constant-temperature molecular dynamics

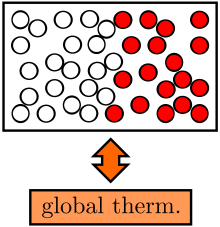
- Plain, Hamiltonian MD conserves energy and is not ergodic
- Must modify the dynamics to model the interaction with a heat bath (**thermostats**)
- Andersen thermostat: randomize atomic velocities every now and then
    - Exploits factorization of the canonical partition function
    - Simple, physically sound and effective



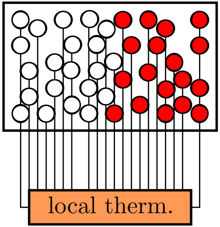
Andersen, J. Chem. Phys. 1980

Local and global thermostats

- "Global" thermostats enforce the correct distribution of the *total* kinetic energy (Berendsen, Nosé-Hoover, Bussi)
  - Gentle on the dynamics, but rely on the intrinsic Hamiltonian dynamics to relax internal degrees of freedom
- "Local" thermostats enforce the correct distribution of *each component* of the momentum (Andersen, Langevin, massive NHC)
  - More aggressive, slower collective dynamics, but also effective for poorly ergodic systems (e.g. crystals)



$$P(K) \propto K^{3N/2-1} e^{-\beta K}$$



$$P(\mathbf{p}) \propto e^{-\beta \mathbf{p}^2/2m}$$

Deterministic and stochastic thermostats

- Nosé-Hoover thermostat: extended-Lagrangian approach, **deterministic** equations of motion

$$\dot{q} = \frac{p}{m}, \quad \dot{p} = -\frac{\partial V}{\partial q} - p\frac{p_s}{Q}, \quad \dot{p}_s = \frac{p^2}{m} - \frac{1}{\beta}, \quad \dot{s} = \frac{p_s}{Q}$$

  - Not ergodic, must introduce chains to make the dynamics chaotic
  - The local version is not rotationally invariant
  - Integration is not straightforward, must use multiple time step

- Langevin-style thermostats: intrinsically **stochastic** dynamic

$$\dot{q} = \frac{p}{m}, \quad \dot{p} = -\frac{\partial V}{\partial q} - \gamma p + \sqrt{2m\gamma/\beta}\xi, \quad \langle \xi(t)\xi(0)\rangle = \delta(t)$$

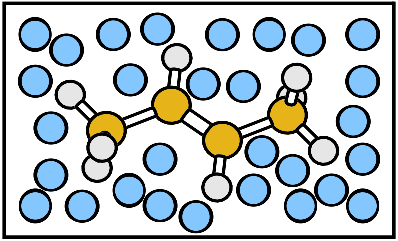
  - Ergodic, very natural model for coupling to a heat bath
  - Linear equations, very stable and easy to integrate
  - Require some care with the random number generator (parallelism!)

Nosé, J. Chem. Phys. 1984; Hoover, Phys. Rev. A 1985; Schneider & Stoll, Phys. Rev. B 1978

Other ensembles

- It is possible to modify the equations that underlie MD (or MC!) to sample several different ensembles

1. Plain Hamilton’s equations sample the microcanonical ensemble
2. A thermostat allows for energy fluctuations and enables canonical sampling
3. It is also possible to introduce a *barostat*, equations that allow for volume fluctuations to sample an isobaric/isothermal ensemble

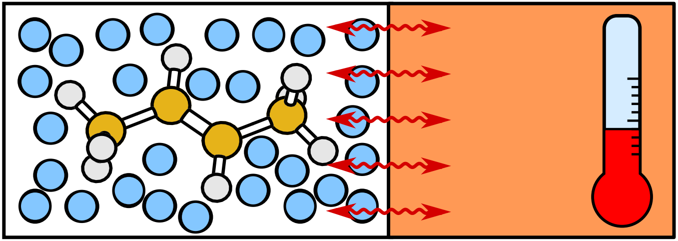


$$P \propto \delta(E - \bar{E})\delta(V - \bar{V})$$

Other ensembles

- It is possible to modify the equations that underlie MD (or MC!) to sample several different ensembles

1. Plain Hamilton's equations sample the microcanonical ensemble
2. A thermostat allows for energy fluctuations and enables canonical sampling
3. It is also possible to introduce a *barostat*, equations that allow for volume fluctuations to sample an isobaric/isothermal ensemble

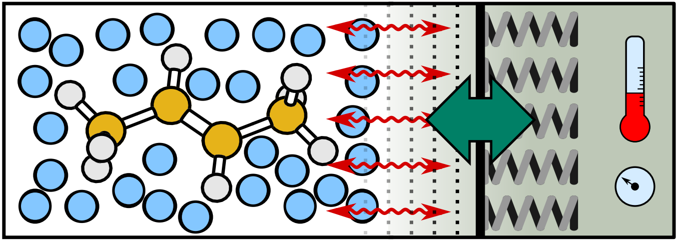


$$P \propto e^{-\beta E}\delta(V-\bar{V})$$

Other ensembles

- It is possible to modify the equations that underlie MD (or MC!) to sample several different ensembles

1. Plain Hamilton's equations sample the microcanonical ensemble
2. A thermostat allows for energy fluctuations and enables canonical sampling
3. It is also possible to introduce a *barostat*, equations that allow for volume fluctuations to sample an isobaric/isothermal ensemble



$$P \propto e^{-\beta E} e^{-\beta p V}$$