



University of Antwerp
| Faculty of Science

Parallel Programming

Project assignment

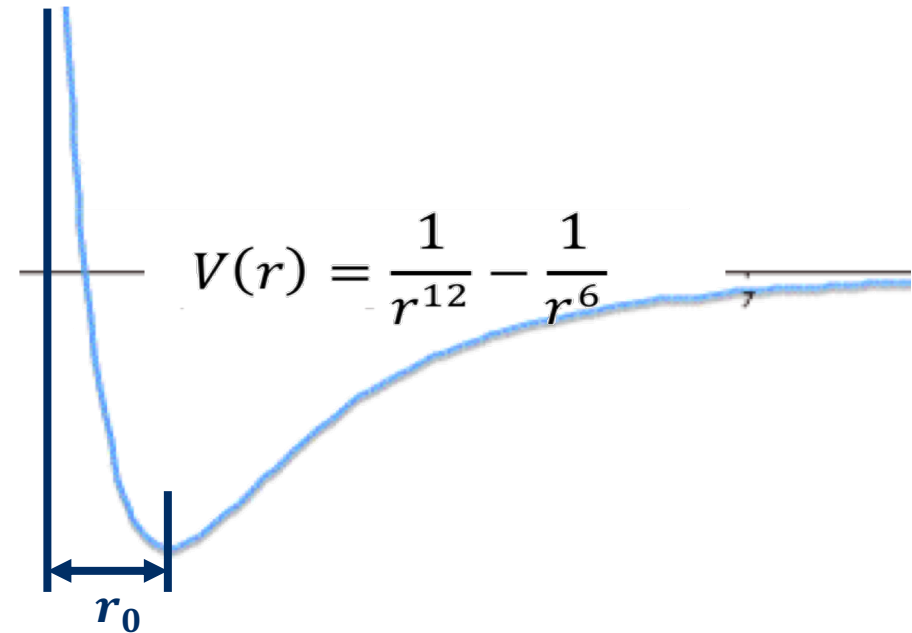
Domain decomposition for Molecular Dynamics

phase 1

sequential code

Molecular Dynamics – LJ potential

- interaction = Lennard-Jones potential (without coefficients)
- cut-off distance: $r_c = 3r_0$



- avoid sqrt when computing the distance: $V(r^2) = \frac{1}{(r^2)^6} - \frac{1}{(r^2)^3}$

Molecular Dynamics – Verlet list

- Verlet list of atom i , L_i is list of all atoms j for which

$$r_{ij} \leq r_0 \text{ or } r_{ij}^2 \leq r_0^2$$

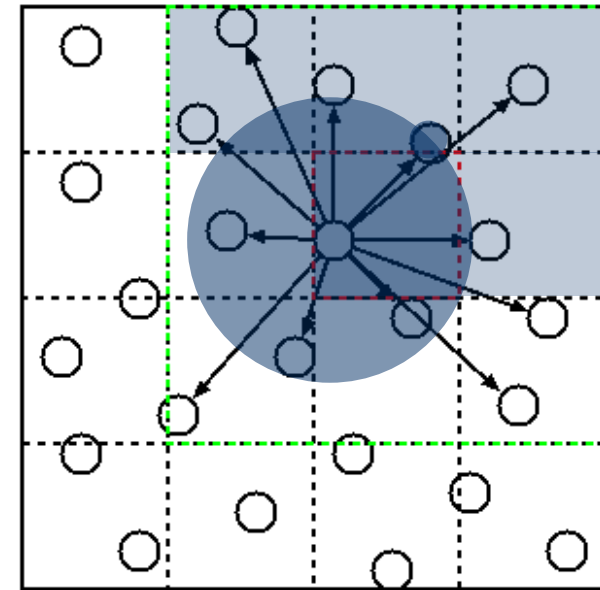
- and

$$i < j$$

- this excludes self-interaction ($i = j$), and avoids counting interactions twice ($r_{ij} = r_{ji}$)
- We stick to the two-dimensional case
 - easier to visualise

Molecular Dynamics

- Create a square grid, with grid size r_c
- Cell C_{kl} is list of atoms i for which
$$kr_c \leq x_i < (k+1)r_c$$
$$lr_c \leq y_i < (l+1)r_c$$
- Denote the cell of atom i as $C(i)$
- To build the Verlet list of atom i in $C(i) = C_{kl}$ you loop over all atoms $j \neq i$ in $C_{k,l}$, $C_{k,l+1}$, $C_{k-1,l+1}$, $C_{k+1,l+1}$.
- This way, every pair is visited only once!



Molecular dynamics – force and acceleration

- The force exerted by atom j on atom i is

$$\bar{F}_{ij} = \nabla V(r_{ij}) = \frac{dV(r^2)}{dr^2} \nabla(r^2) = \frac{dV(r^2)}{dr^2} \cdot 2\mathbf{r} \cdot \bar{\mathbf{r}}$$

- and

$$\bar{F}_{ji} = -\bar{F}_{ij}$$

- total force on atom i :

$$\bar{F}_i = \sum_{j \in L_i} \bar{F}_{ij}$$

- acceleration: $\bar{a}_i = \frac{\bar{F}_i}{m_i}$

Molecular dynamics – time evolution

Molecular dynamics – time evolution

- Atoms may move from one cell to another. Update the grid cells every time the Verlet list must be updated.

Molecular Dynamics - initialization

- you must control the location of atoms at $t = 0$. If, accidentally, two atoms are very close, the force will be very high and the atoms will move too fast, your simulation will explode. If they are too far the the force will be small and the atoms move slowly
- fixed timestep is problematic, select timestep as function of highest force or acceleration
- generate atoms in close packing (hexagonal in 2D, FCC in 3D), add a bit of noise to the positions
- This component will be provided by me

Strategy

- **make sure you have carried out personal-setup.pdf**
 - github account!
 - mpi installation (was forgotten in the VM)
- **see principles in first talk**
- **use micc**
 - put your code in a private github repository and give me access
 - you have a backup at any time (commit and push regularly)
 - I can check your code and investigate your problems to help you
 - the code is a deliverable for the evaluation

Strategy

- **start in Python**
- **write tests for every component**
 - LJ potential and force
 - grid construction
 - verlet list construction
 - total forces
 - time evolution
- **replace Python version with a C++ or Fortran version for**
 - LJ potential and forces
 - time evolution
 - (other components may remain in Python even if they are slow)
 - provide timings for both versions (e.g. using the `et_stopwatch` package, see micc tutorials)

phase 2

Parallelisation

Molecular Dynamics - Parallelization

- Divide the work across several processes, using mpi4py. Every core runs a separate process that does part of the work. E.g.

```
> mpirun -np 2 python md.py
```

runs 2 processes. If the tasks are independent, they must not exchange information. This is not always possible.

- Check the (api) documentation of mpi4py at <https://mpi4py.readthedocs.io/en/stable/>
- For a tutorial on MPI check <https://folk.idi.ntnu.no/elster/tdt4200-f09/gropp-mpi-tutorial.pdf>

Molecular Dynamics - Parallelization

- **domain composition**

- Divide the space in domains, one per process. Each domain has its own atoms.
- Let domain boundaries coincide with cell boundaries.
- Domain needs information from neighbouring domains. Cells on the boundary must be known to each domain touching that boundary in order to compute the Verlet list. This is the halo concept, also known as ghost cells.
- Atoms may move from one cell to another and from one domain to another.
- Think carefully about which information must be communicated and when
- The amount of communication is proportional to the area of the boundaries
- The amount of work (load) is proportional to the volume of the domains

Molecular Dynamics - Parallelization

- **Communication is slow and is not useful work on its own**
- **In general we want to minimize the amount of data communicated.**
- **Try to overlap computation and communication. While waiting for the communication to finish, do some other useful computation.**
 - check the tutorial by Gropp and make sure you understand these concepts:
 - one-sided vs two-sided communication
 - blocking vs non-blocking communication

Molecular Dynamics - Parallelization

- if the atom density is uniform across domains, each process has approximately the same amount of work and the processes will finish their timestep more or less simultaneously.
- If not, some processes will finish their timestep well before others and will have to wait for the last process before the next timestep can be started. Load balancing needed
 - Dynamical domain decomposition. E.g. using only boundaries perpendicular to the x-axis, letting the boundaries move along the x-axis so they have approximately the same atom density.
 - Much more domains (= small tasks) than processes. The tasks are moved around to waiting processes.
 - We won't go there...

Molecular Dynamics - Parallelization

- **start simple**
- **Divide the space in two domains along a straight boundary parallel with cell boundaries.**
- **think about the consequences of having the domain boundary in the middle of the cells vs coinciding with the cell boundary.**
- **decide what to communicate and how**
- **try to overlap computation and communication**
- **Once you understand the machinery, you can proceed further**
 - you should get a 2 domain case working
 - draw up a plan on how to go beyond 2 domains

phase 3

Evaluation

Evaluation

- **presentation of your work (probably via ms teams)**
 - explain your approach
 - evaluate the performance of your approach (measurements)
 - (I will give you some tools later)
 - I will ask questions to see if you understand what you did, and to challenge your approach
- **deliverables**
 - the presentation (.pptx or .pdf)
 - the code
 - both should be in a your github repository