UNIVERSITY OF LEEDS

Molecular Simulation – Phys3190

Session 0 – Introduction to Molecular Dynamics

BENJAMIN S. HANSON

1

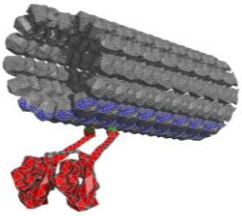
The Admin Stuff

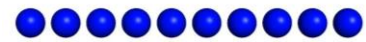
But it is pretty important

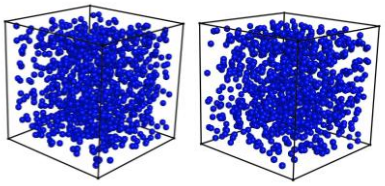
2

Who Am I?

Name: Ben Hanson
Field: Computational Biophysics
Office: Bragg Building 1.35b
Email: b.s.hanson@leeds.ac.uk





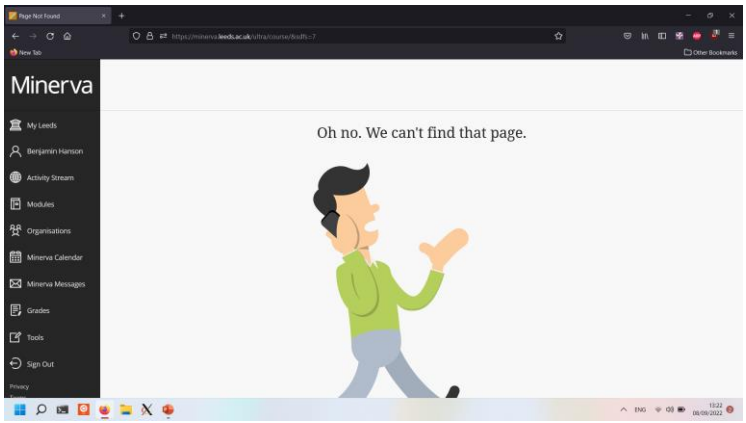


Minerva

What is it?

Minerva

What is it?



Who do I email?

- Content issue – Me (b.s.hanson@leeds.ac.uk)
- Technical issue – A friend first, then me, then IT
- Login issue – Turn it off and on again, clear browser cache, then me, then IT

School Support

Who is it?

Academic Stuff:

- Lecture Questions – Lecturer / Module Leader
- Admin Questions – Student Education Service
- Other questions – Academic Personal Tutor

Feel uncomfortable with any of these? Talk to any staff member you want and they will usually be happy to help ☺

Pastoral Stuff:

- Most questions – Student Education Service
- Want someone you know? – Academic Tutor

General Information:

- The desk in the Parkinson Building Foyer

https://students.leeds.ac.uk/

https://students.leeds.ac.uk/info/10700/support_and_wellbeing/804/helpful_support_contacts

This Module

What do I need to do?

This is **Molecular Simulation: Theory & Practice,** PHYS3190

Assignments:

1. Exercise Sheets (30% of total grade)
   - One a week, numerical and graphing questions around analysis of MD simulations: Marked by an autograder and manually by me!

2. Grant Exercise (30% of total grade)
   - One written exercise summarising all you have learned Marked by self-assessment against a mark scheme that I'll provide

This Module

What do I need to do?

This is **Molecular Simulation: Theory & Practice,** PHYS3190

Non-contact time:

1. Watch some YouTube videos
   - https://www.youtube.com/playlist?list=PLm8ZSArAXicIWTHEWgHG5mDr8YbrdcN1K – This video series is amazing, it's everything!

2. Read some papers (I've added a list to Minerva)
3. Practice!
   - Test the things we talk about!
   - Write your own code / run your own simulations to ask questions and solve problems that you set yourself

4. Go for a walk! Not even kidding.

We expect University to be a full working week (i.e. 37.5 hours). Split this however you like.

The Admin Stuff

The Really Important Stuff

- You **can** work together on the weekly exercise sheets (ideally, complete exercise sheets independently though!)
- You **must not** work together on the grant exercise
- You **must not** use any outside assistance on assessments

- Backup your data – data loss is not an accepted excuse for late submission

The Admin Stuff

But it is pretty important

Done!

Let's do some molecular simulation! ☺

Introduction

Monte Carlo and Molecular Dynamics

11

Introduction

A Link Between the Methods

Molecular Dynamics

Monte Carlo

Introduction

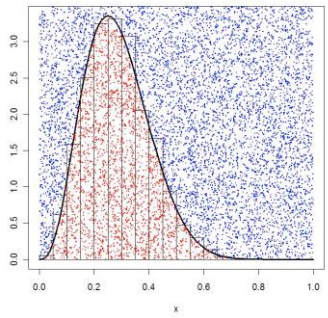
Monte Carlo and Molecular Dynamics

13

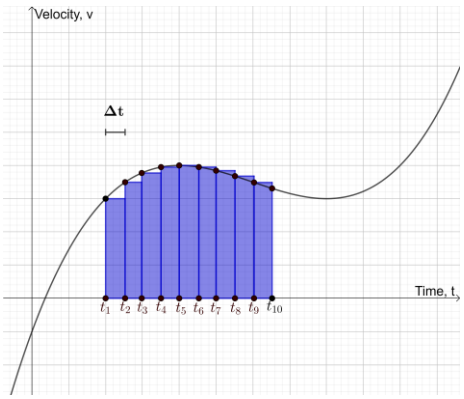
Introduction

A Link Between the Methods

Monte Carlo



Molecular Dynamics

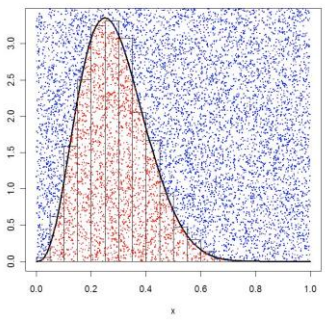


14

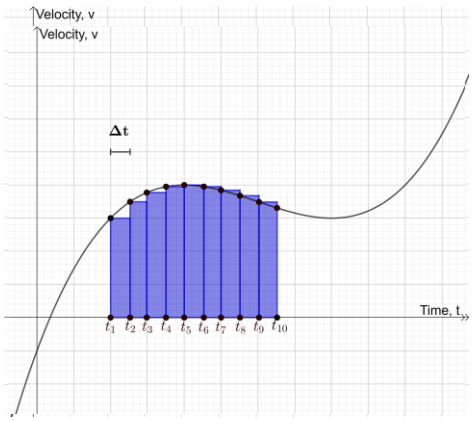
Introduction

A Link Between the Methods

Monte Carlo



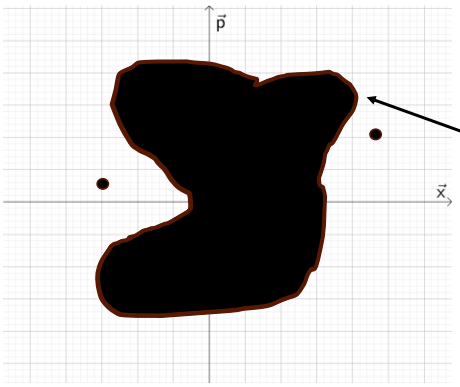
Molecular Dynamics



Introduction

A Link Between the Methods

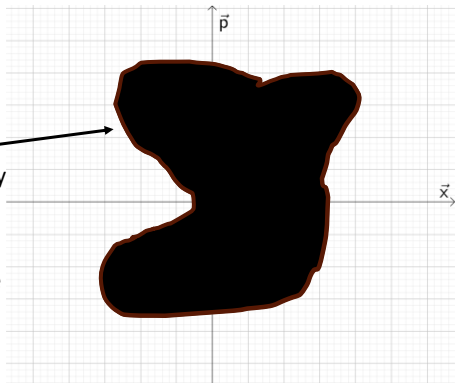
Monte Carlo (Metropolis)



Free energy surface

A long time later…

Molecular Dynamics



Introduction

A Link Between the Methods

Monte Carlo (Metropolis)

Pros:

- Fast exploration of phase space
- Quick estimation of macroscopic variables

Cons:

- Non-physical model (only converges in long-time limit)
- No understanding of microstate structure

Converge because:

1. Ergodic hypothesis
2. Equal a priori probabilities

Linking these formally is a big research area!

Molecular Dynamics

Pros:

- Logical exploration of phase space
- Microstates explicitly measured

Cons:

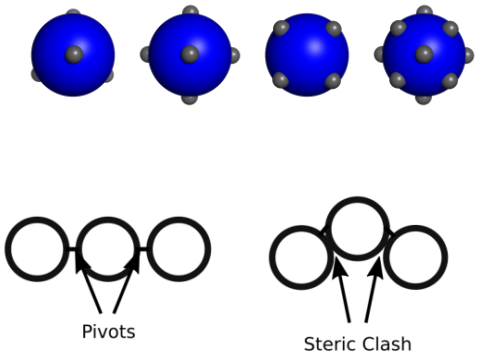
- Really fricking slow (and I mean seriously)
- Macroscopic variables always approximations (upper or lower limits)

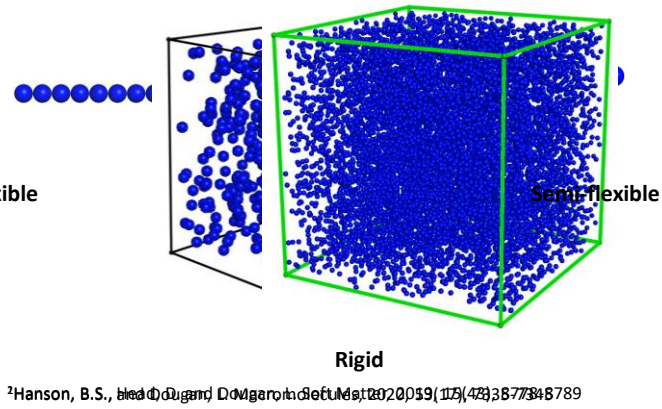
Introduction

How does molecular dynamics work?

Introduction

BioNet (an MD package I wrote)

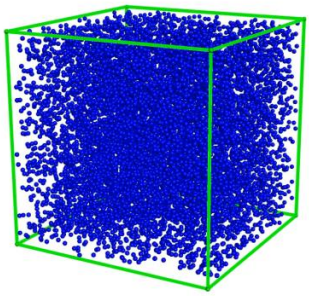




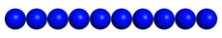
[1]Hanson, B.S., Head, D and Dougan, L. Soft Matter, 2019, 15(43), 8778-8789
[2]Hanson, B.S., and Dougan, L. Macromolecules, 2020, 53(17), 7335-7345

Introduction

Questions

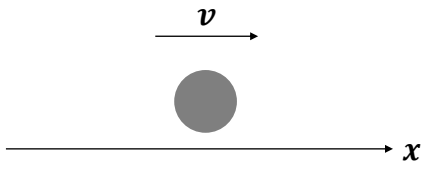


1. What physics is happening in these simulations, if any?
2. How do we translate physics into computer language?
3. How can things like fractals emerge from these simulations?
4. Why do we even need to do simulations in the first place?



A Dynamic Model

Beads & Diffusion



$$F = ma$$

$$m\frac{d^2x}{dt^2} + \lambda\frac{dx}{dt} = 0$$

Stochastic function

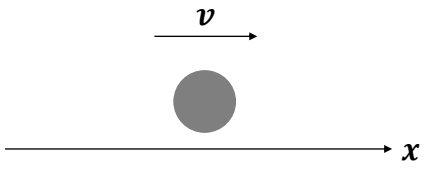
$$m\frac{d^2x}{dt^2} + \lambda\frac{dx}{dt} = F_n(t)$$

Langevin Equation

$\boldsymbol{Drag, \lambda}$ $\boldsymbol{Mass, m}$

A Dynamic Model

Beads & Diffusion



$$m\frac{d^2x}{dt^2} + \lambda\frac{dx}{dt} = F_n(t)$$

Brownian Motion Equation

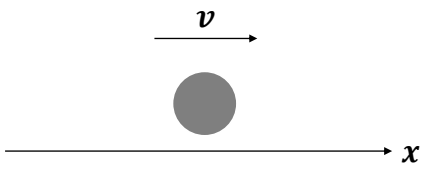
$$\int_0^t \frac{dx}{dt'}dt' = \frac{1}{\lambda}\int_0^t F_n(t')dt'$$

$$x(t) = x(0) + \frac{1}{\lambda}\int_0^t F_n(t')dt'$$

A Dynamic Model

Beads & Diffusion

$$x(t) = x(0) + \frac{1}{\lambda}\int_0^t F_n(t')dt'$$



$$\langle x(t)\rangle = x(0) + \frac{1}{\lambda}\int_0^t \langle F_n(t')\rangle dt'$$

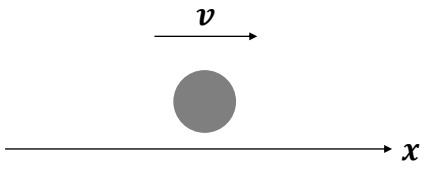
$\langle F_n(t')\rangle = 0$ — Symmetric function

$$\langle \boldsymbol{x(t)}\rangle = \boldsymbol{x(0)}$$

A Dynamic Model

Beads & Diffusion

$$x(t) = x(0) + \frac{1}{\lambda}\int_0^t F_n(t')dt'$$



$$\langle x(t) - x(0)\rangle = \frac{1}{\lambda}\int_0^t \langle F_n(t')\rangle dt'$$

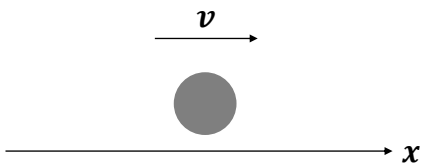
$\langle F_n(t')\rangle = 0$ — Symmetric function

$$\langle \boldsymbol{x(t) - x(0)}\rangle = \boldsymbol{0}$$

A Dynamic Model

Beads & Diffusion

$$x(t) = x(0) + \frac{1}{\lambda}\int_0^t F_n(t')dt'$$



$$\langle (x(t) - x(0))^2 \rangle = \frac{1}{\lambda^2}\iint_0^t \langle F_n(t')F_n(t'')\rangle dt'dt''$$

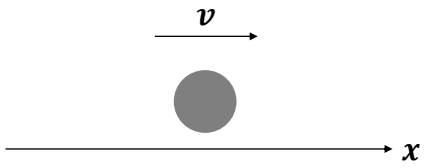
$$\langle F_n(t')F_n(t'')\rangle = 2k_BT\lambda\delta(t' - t'')$$

Fluctuation-Dissipation Theorem. Extremely profound!

A Dynamic Model

Beads & Diffusion

$$x(t) = x(0) + \frac{1}{\lambda}\int_0^t F_n(t')dt'$$



$$\langle (x(t) - x(0))^2 \rangle = \frac{2k_BT}{\lambda}\iint_0^t \delta(t' - t'')dt'dt''$$

$$\langle (x(t) - x(0))^2 \rangle = \frac{2k_BT}{\lambda}\int_0^t 1\, dt''$$

$$\langle (x(t) - x(0))^2 \rangle = \frac{2k_BT}{\lambda}t$$

$$D = \frac{k_BT}{\lambda}$$

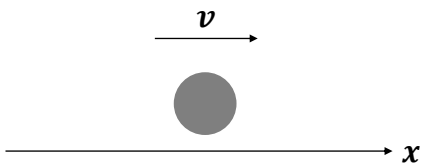
$$\boldsymbol{\langle (x(t) - x(0))^2 \rangle = 2Dt}$$

Einstein relation. Also extremely profound!

A Dynamic Model

Beads & Diffusion

$$\langle (\boldsymbol{x(t)} - \boldsymbol{x(0)})^2 \rangle = \boldsymbol{2Dt}$$



By giving the stochastic thermal noise function the following properties:

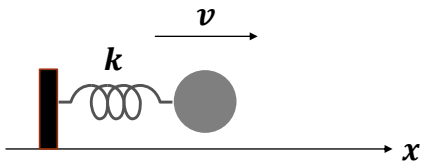
$$\langle F_n(t') \rangle = 0$$
$$\langle F_n(t') F_n(t'') \rangle = 2k_B T \lambda \delta(t' - t'')$$

We obtain the correct theoretical and experimental form of diffusion!

A Dynamic Model

Beads & Springs

$$\lambda \frac{dx}{dt} + kx = F_n(t)$$



$$\tau = \frac{\lambda}{k} \qquad \frac{d}{dt}\left[x exp\left(\frac{t}{\tau}\right)\right] = \frac{F_n(t)}{\lambda} exp\left(\frac{t}{\tau}\right)$$

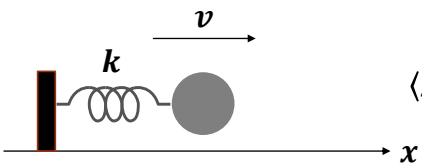
$$\left[x(t') exp\left(\frac{t'}{\tau}\right)\right]_0^t = \frac{1}{\lambda}\int_0^t F_n(t') exp\left(\frac{t'}{\tau}\right) dt'$$

$$x(t) = x(0) exp\left(-\frac{t}{\tau}\right) + \frac{1}{\lambda}\int_0^t F_n(t') exp\left(\frac{t' - t}{\tau}\right) dt'$$

A Dynamic Model

Beads & Springs

$$x(t) = x(0)exp\left(-\frac{t}{\tau}\right) + \frac{1}{\lambda}\int_0^t F_n(t')exp\left(\frac{t'-t}{\tau}\right)dt'$$



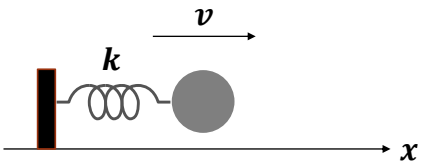
$$\langle x(t)\rangle = \langle x(0)\rangle exp\left(-\frac{t}{\tau}\right) + \frac{1}{\lambda}\int_0^t \langle F_n(t')\rangle exp\left(\frac{t'-t}{\tau}\right)dt'$$

$$\langle x(t)\rangle = x(0)exp\left(-\frac{t}{\tau}\right)$$

A Dynamic Model

Beads & Springs

$$x(t) = x(0)exp\left(-\frac{t}{\tau}\right) + \frac{1}{\lambda}\int_0^t F_n(t')exp\left(\frac{t'-t}{\tau}\right)dt'$$



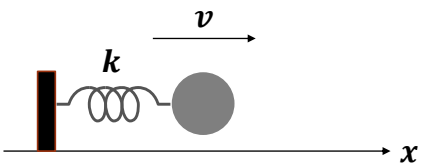
$$\langle x(t)^2\rangle = \langle x(0)^2\rangle exp\left(-\frac{2t}{\tau}\right) + \frac{1}{\lambda^2}\iint_0^t \langle F_n(t')F_n(t'')\rangle exp\left(\frac{(t'+t''-2t)}{\tau}\right)dt'dt''$$

$$= x(0)^2 exp\left(-\frac{2t}{\tau}\right) + \frac{2k_BT}{\lambda}\iint_0^t \delta(t'-t'')exp\left(\frac{t'+t''-2t}{\tau}\right)dt'dt''$$

$$= x(0)^2 exp\left(-\frac{2t}{\tau}\right) + \frac{2k_BT}{\lambda}\int_0^t exp\left(\frac{2(t''-t)}{\tau}\right)dt''$$

A Dynamic Model

Beads & Springs



$$x(t) = x(0)exp\left(-\frac{t}{\tau}\right) + \frac{1}{\lambda}\int_0^t F_n(t')exp\left(\frac{t'-t}{\tau}\right)dt'$$

$$\langle x(t)^2\rangle = x(0)^2 exp\left(-\frac{2t}{\tau}\right) + \frac{2k_BT}{\lambda}\frac{\tau}{2}\left[exp\left(\frac{2(t''-t)}{\tau}\right)\right]_0^t$$

$$\langle x(t)^2\rangle = x(0)^2 exp\left(-\frac{2t}{\tau}\right) + \frac{k_BT}{k}\left(1 - exp\left(-\frac{2t}{\tau}\right)\right)$$

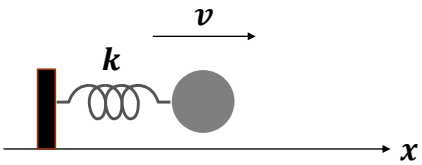
$$\langle x(t)^2\rangle = \frac{k_BT}{k} + \left(x(0)^2 - \frac{k_BT}{k}\right)exp\left(-\frac{2t}{\tau}\right)$$

$t \to \infty$ Equipartition!

$$\boldsymbol{U = \frac{1}{2}k\langle x(t)^2\rangle = \frac{1}{2}k_BT + \left(\frac{1}{2}kx(0)^2 - \frac{1}{2}k_BT\right)exp\left(-\frac{2t}{\tau}\right)}$$

A Dynamic Model

Beads & Springs



$$\boldsymbol{U = \frac{1}{2}k\langle x(t)^2\rangle = \frac{1}{2}k_BT + \left(\frac{1}{2}kx(0)^2 - \frac{1}{2}k_BT\right)exp\left(-\frac{2t}{\tau}\right)}$$

By giving the stochastic thermal noise function the following properties:

$$\langle F_n(t')\rangle = 0$$

$$\langle F_n(t')F_n(t'')\rangle = 2k_BT\lambda\delta(t'-t'')$$

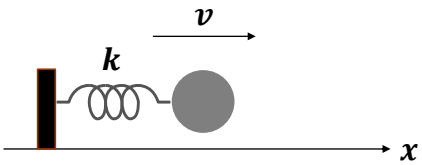
We obtain the correct theoretical and experimental structure of energy!

Introduction

Why do we need computer simulations?

A Dynamic Simulation

Complicated Forces



In general, $\lambda \frac{dx}{dt} = F_n(t) + F_{ex}(t)$

$F_{ex}(t) - Springs, Steric, Hydro, Electro, Everything!$

Anything more complicated than linear or more than 1 particle?
Can't analytically solve it, sorry! Too difficult ☹

$$\frac{x(t+\Delta t) - x(t)}{\Delta t} = \frac{1}{\lambda}(F_n(t) + F_{ex}(t))$$

$$\boldsymbol{x(t+\Delta t) = x(t) + \frac{\Delta t}{\lambda}(F_n(t) + F_{ex}(t))}$$

Thanks! Any questions?

CCPBioSim is the UK network / consortium for biomolecular simulation. They do many awesome things, and also work with the UK national supercomputing facilities
https://www.ccpbiosim.ac.uk/

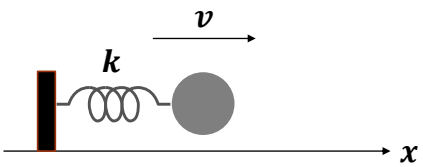
CECAM is the European equivalent, and their head office is in Lausanne, on the shore of Lake Geneva in Switzerland. Get yourself out there! ☺
https://www.cecam.org

Numerical Instability

What do we need to consider in computer simulations?

Numerical Instability

Beads & Springs



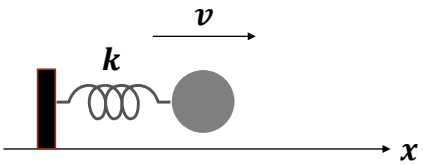
$$\lambda\frac{dx}{dt} + kx = F_n(t)$$

$$\frac{x(t+\Delta t) - x(t)}{\Delta t} + \frac{1}{\tau}x(t) = \frac{1}{\lambda}F_n(t)$$

$$x(t+\Delta t) = \left(1 - \frac{\Delta t}{\tau}\right)x(t) + \frac{\Delta t}{\lambda}F_n(t)$$

Numerical Instability

Beads & Springs



$$\lambda\frac{dx}{dt} + kx = F_n(t)$$

$$x(t+\Delta t) = \left(1 - \frac{\Delta t}{\tau}\right)x(t) + \frac{\Delta t}{\lambda}F_n(t)$$

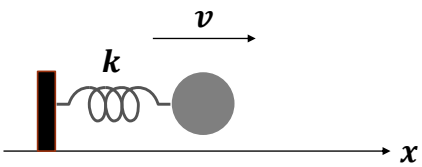
$$x(t+2\Delta t) = \left(1 - \frac{\Delta t}{\tau}\right)x(t+\Delta t) + \frac{\Delta t}{\lambda}F_n(t+\Delta t)$$

$$= \left(1 - \frac{\Delta t}{\tau}\right)^2 x(t) + \left(1 - \frac{\Delta t}{\tau}\right)\frac{\Delta t}{\lambda}F_n(t) + \frac{\Delta t}{\lambda}F_n(t+\Delta t)$$

$$x(t+3\Delta t) = \left(1 - \frac{\Delta t}{\tau}\right)^3 x(t) + \left(1 - \frac{\Delta t}{\tau}\right)^2\frac{\Delta t}{\lambda}F_n(t) + \left(1 - \frac{\Delta t}{\tau}\right)\frac{\Delta t}{\lambda}F_n(t+\Delta t) + \frac{\Delta t}{\lambda}F_n(t+2\Delta t)$$

Numerical Instability

Beads & Springs



$$\lambda\frac{dx}{dt}+kx=F_n(t)$$

$$x(t+N\Delta t)=\left(1-\frac{\Delta t}{\tau}\right)^N x(t)+\sum_{n=0}^{N-1}\left(1-\frac{\Delta t}{\tau}\right)^n F_n(t+(N-1-n)\Delta t)$$

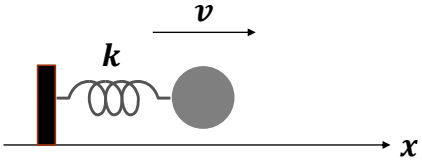
$$\boldsymbol{x(N\Delta t)=\left(1-\frac{\Delta t}{\tau}\right)^N x(0)+\sum_{n=0}^{N-1}\left(1-\frac{\Delta t}{\tau}\right)^n F_n((N-1-n)\Delta t)}$$

Numerical equivalent of the continuous solution.

Numerical Instability

Beads & Springs

$$x(N\Delta t)=\left(1-\frac{\Delta t}{\tau}\right)^N x(0)+\sum_{n=0}^{N-1}\left(1-\frac{\Delta t}{\tau}\right)^n F_n((N-1-n)\Delta t)$$



$$\langle x(N\Delta t)\rangle=\left(1-\frac{\Delta t}{\tau}\right)^N\langle x(0)\rangle+\sum_{n=0}^{N-1}\left(1-\frac{\Delta t}{\tau}\right)^n\langle F_n((N-1-n)\Delta t)\rangle$$

$$\boldsymbol{\langle x(N\Delta t)\rangle=\left(1-\frac{\Delta t}{\tau}\right)^N x(0)}$$

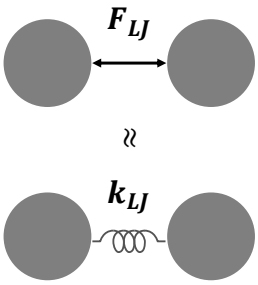
$$\boldsymbol{\rightarrow \Delta t \ll \tau}$$

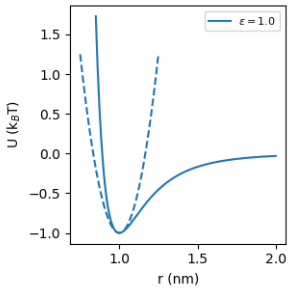
Can also do $\langle x(N\Delta t)^2\rangle$, with discrete $\langle F_n(m\Delta t)F_n(n\Delta t)\rangle = 2k_BT\lambda\Delta t$ and solving a geometric series sum.
It turns out that for numerical stability, $\rightarrow \boldsymbol{\Delta t \ll \tau}$, but for numerical accuracy, $\rightarrow \boldsymbol{\Delta t \ll \tau/2}$. You can even work out the exact level of accuracy you have for these simple models. I'll leave that to you if you're interested…or it's in my thesis!

Numerical Instability

Lennard-Jones Linearisation

We can linearise any force to see how check unstable it (potentially) is





$$U = \epsilon\left(\left(\frac{r_0}{r}\right)^{12} - 2\left(\frac{r_0}{r}\right)^{6}\right)$$

$$k_{LJ} = \left.\frac{d^2U}{dx^2}\right|_{r=r_0}$$

$$\boldsymbol{k_{LJ} = 72\frac{\epsilon}{r_0^2}}$$

$$\boldsymbol{U_{LJ}(r) \approx \frac{1}{2}k_{LJ}(r - r_0)^2 - \epsilon = 36\epsilon\left(\frac{r}{r_0} - 1\right)^2 - \epsilon}$$

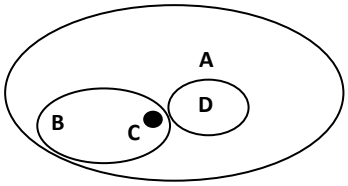
Philosophy

How do simulations relate to science more generally?

Philosophy

The Scientific Method

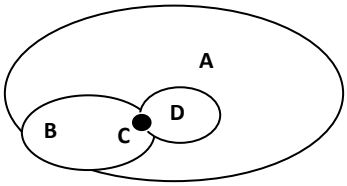
Theoretical Science



- C in B
- All Bs are As
- Therefore, C in A

Assumptions:
- Maths is fundamental
- Everything can be deduced with certainty

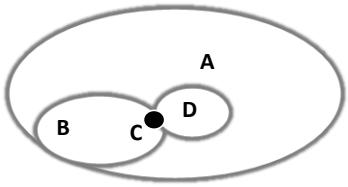
Computational Science



- Assume some experimental state to be true with 100% certainty.
- Numerically derive from there.

Assumptions:
- Either could be fundamental
- Based on inferences, things can be deduced…but only to the level of accuracy of the initial inferences.

Experimental Science



- Measure a set of unbiased data
- All Bs are probably As. Some Bs may be Ds
- C in B
- C is probably in A, and maybe in D

Assumptions:
- (Observable) data is fundamental
- Everything can be inferred with error (which tends to zero at infinite data)

Philosophy

Newton & Einstein

- Newton said $F_{12} = \frac{Gm_1m_2}{r^2}$. He said this was true for **all** massive bodies. In other words, he assumed that his derivation was true *for everything* i.e. for the set of all bodies in the Universe.
- Einstein **observed** that this is not true if $v_1 - v_2 \sim c$.
- Newton derived (theoretical), and compared to data (experiment) but his error was that he assumed his derivation was true for all data, when in fact he only had access to a subset of all data. His initial assumptions were partially incorrect, and so his derivation had a systematic error i.e. didn't account for fast objects.
- Einstein also derived. We may find one day that he too did not account for all the data. But for now, it's the best we've got.

Philosophy

Conclusions

- All simulations assume some state of the universe, some equation, or both, to be 100% true, and use computers to numerically derive from there. Numerical derivation is exactly as true as analytical mathematics in the limit of $\Delta t \to 0$
    - MD says if F=ma and atoms are point-like particles with some (measured) charge subject to thermal noise, and covalent bonds are linear springs with some (measured) stiffness and angular stiffness, and VdW works as Lennard-Jones potential with some cutoff, this is what will happen.
    - BioNet (my code) says if F=ma, and proteins are spheres which interact via surface steric interactions and are subjected to thermal noise, this is what will happen.
- All dynamic simulations numerically solve some equation of motion (continuous-space, discrete-time)
- All kinetic simulations use "if" statements to make decisions (discrete-space, discrete-time)
- The numerics introduce systematic error due to integration resolution. This can be reduced together with the timestep, but remember! The stronger the force, the worse the error and instabilities will be, and the more time-limited the simulation will be.
- **Mind-blowing part:** The universe is potentially a simulation with a tiny timestep (plank time), where the rules of the simulation are the rules of the universe itself i.e. physical laws. Experimental science tries to infer these rules. Is maths limited by this, or is it valid outside of the universe? I wonder… (talk to me about this stuff, I find it fascinating!)