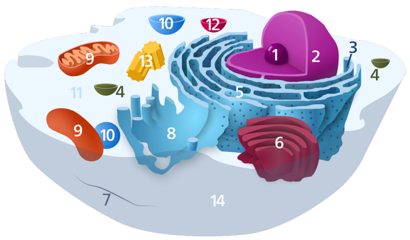
Take-home message

More coarse-graining?

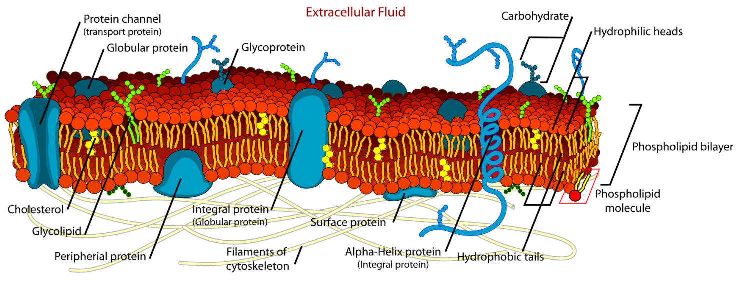
When do we stop?

We throw away time …

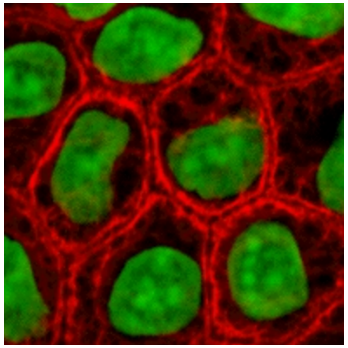
Cellular cytoskeleton components



Wikipedia



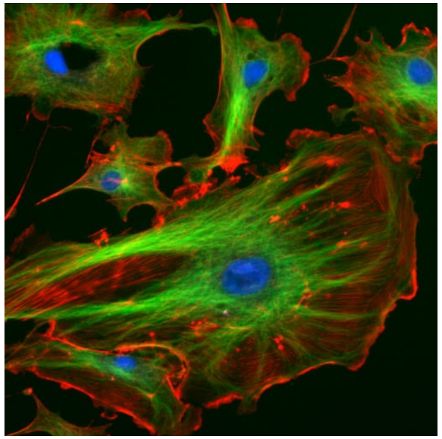
Wikipedia



Wikipedia

Keratin intermediate filaments

http://rsb.info.nih.gov/ij/images/



Actin filaments
Microtubules
Nucleus

Filament type	Monomer protein	Filament width	Persistence length
Actin filaments	Actin (43 kDa)	7 nm	10 μm
Intermediate filaments	Vimentin, keratin, etc.	10 nm	0.3 - 1 μm
Microtubules	Tubulin (50kDa)	25 nm	100 - 10,000 μm
Plasma membrane-associated cytoskeleton	Actin, spectrin, ankyrin, etc	Hex width ~ 70 μm	

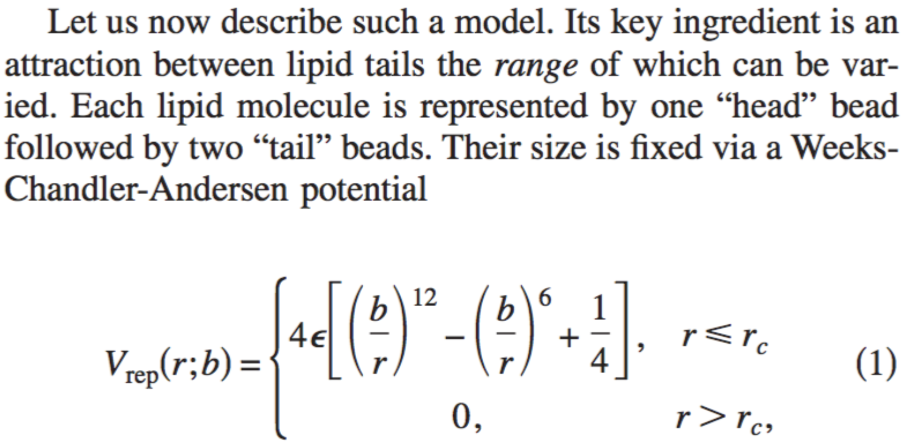
Ananthakrishnan et al. Int. J. Biol. Sci. 3:303 (2007)

Actin is very abundant in eukayrotes (~ 5%), and forms long filaments used in muscle, cell crawling, the *cell cortex* attached to the plasma membrane. About 50% of actin in a cell is unpolymerized and occurs as monomeric actin or actin-protein complexes. A dynamic equilibrium between monomeric actin and the polymers is used by the cell to drive surface movement. Many proteins binds to actin filaments and control their growth, disassembly, cross-linking and network properties.

Solvent-Free Molecular Dynamics

Recall that a single dendritic spine contains ~10^9 molecules, most of them water.

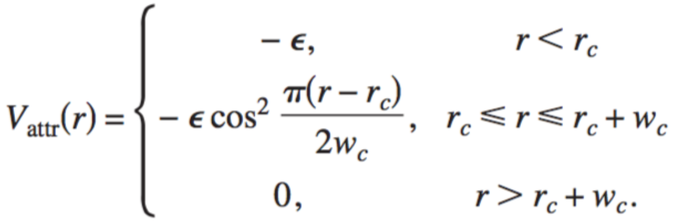
Even coarse-grained MD and DPD retain water particles. We could speed up a simulation a lot if we could remove the water: this is the aim of *solvent-free* simulations. The solvent is replaced by (complex) long-range forces between the particles of interest. This has obvious problems for membranes as they only form in the presence of a suitable solvent. As the inventors of a code called *Espresso*, Cooke et al. say:



Let us now describe such a model. Its key ingredient is an attraction between lipid tails the *range* of which can be varied. Each lipid molecule is represented by one "head" bead followed by two "tail" beads. Their size is fixed via a Weeks-Chandler-Andersen potential

$$V_{\text{rep}}(r;b) = \begin{cases} 4\epsilon\left[\left(\frac{b}{r}\right)^{12} - \left(\frac{b}{r}\right)^{6} + \frac{1}{4}\right], & r \leqslant r_c \\ 0, & r > r_c, \end{cases} \qquad (1)$$

www.espresso.mpg.de



$$V_{\text{attr}}(r) = \begin{cases} -\epsilon, & r < r_c \\ -\epsilon \cos^2 \frac{\pi(r-r_c)}{2w_c}, & r_c \leqslant r \leqslant r_c + w_c \\ 0, & r > r_c + w_c. \end{cases}$$

The idea is to mimic the hydrophobic effect by a long-range (compared to the particle size) attraction. This attraction is tuned to reproduce the membrane's material properties. Such models have a long history as seen in the references, but we don't discuss them further here.

O. Farago, J. Chem. Phys. 119:596 (2003)
G. Brannigan and F. L. H. Brown J. Chem. Phys. 120: 1059 (2004)
Z. Wang and D. Frankel J. Chem. Phys. 122:234711 (2005)
I. R. Cooke, K. Kremer and M. Deserno, PRE 72:011506 (2005)

Brownian Dynamics

Brownian dynamics rests on two assumptions (recall Lecture 2 for Langevin's explanation of Brown's observations on pollen grains in water):

- particles of interest are larger than surrounding solvent particles
- effect of solvent particles has two parts: a viscous drag and a continual rapidly-fluctuating force that is uncorrelated in time and space

This allows us to replace the solvent by effective forces on the particles of interest, so no HD forces, and no Navier-Stokes: all motion is diffusive. It may never reach states accessible to HD-preserving methods or may take a long time.

$m.dv/dt = F$ MD = Newton's 2nd law with complex force field

$m.dv/dt = F^C + F^D + F^R$ DPD = Newton's 2nd law with simpler force field

$0 = -\gamma.v + F^C + \sigma.\zeta(t)$ BD = Drop accel., $x(t)$ is updated from drag and noise

BD simulations of filament growth

Assumptions of Brownian dynamics:

- objects of interest are larger/more massive than particles in the surrounding fluid medium
- fluid imposes continuous random force on objects
- all motion is damped by the fluid, so Force $\propto$ velocity not acceleration

With these assumptions, the discretized **vector** equation of motion of a Brownian particle is:

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \frac{\Delta t}{k_B T}\sum_j D_{ij}\cdot \mathbf{F}_{ij} + \mathbf{R}_i(\Delta t) \qquad D_{ij} = D\delta_{ij} = \frac{k_B T}{6\pi\eta r_{\mathrm{pa}}}\delta_{ij}$$

where F_{ij} is the applied force on the particle, $\mathbf{R}_i$ is a random displacement (i.e. $R(\Delta t) \sim \text{Force}.\sqrt{\Delta t}$), and the particle has a radius r_{pa}, and a diffusion coefficient D that (in the absence of hydrodynamic forces) is just Stokes-Einstein. The random term obeys:

$$\langle \mathbf{R}_i \rangle = 0 \qquad \langle \mathbf{R}_i(\Delta t).\mathbf{R}_i(\Delta t) \rangle = 6.D.\Delta t$$

At each time-step, we sample the random displacement vector, calculate the net external force and update the particle's position. If the external force is zero, the particle performs a purely diffusive motion in space, that is just a random walk in 3D.

DL Ermak, JA McCammon, J. Chem. Phys. 69:1352 (1978)

More coarse-graining?

Sometimes, even throwing away the solvent is not enough to be able to simulate some systems - we need to use larger units than single particles or atoms in rigid molecules (recall liquid crystals).

This is the motivation for *Monte Carlo* simulations: instead of following the trajectories of all the particles moving in space and time, we average over everything the whole molecule (or other aggregate) could possibly do.

Thermodynamics is based on two ideas:

if something can happen… it will (i.e., a system explores all of its available microstates)

whatever can happen in the most ways is what is observed in equilibrium (i.e., that macrostate with the most microstates is the observed equilibrium state)

Can we explore states of a system without any equations of motion?

Random walks in phase space

So far, we have used the Langevin equation to model the *actual kinetics* of a small particle (small means $k_B T$ is significant). We could, in principle, do the experiment and compare with the theory.

But RWs are more useful that this: we can apply them to *physically unrealisable* situations.

Consider trying to predict the motion of many interacting rigid particles in a dense fluid.

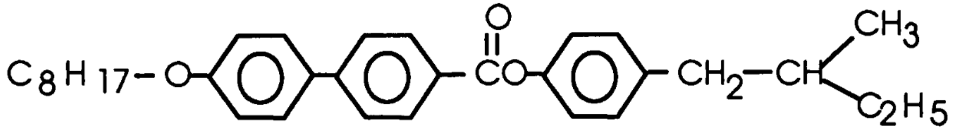
We can do Molecular Dynamics (MD) simulations, which integrate Newton's laws for atoms interacting via complex force fields. But many problems involve rigid molecules or high potential barriers, e.g.,

Hard disks or spheres
Liquid crystal molecules
Rigid proteins
Molecules with many tightly connected degrees of freedom (e.g., a billiard ball)

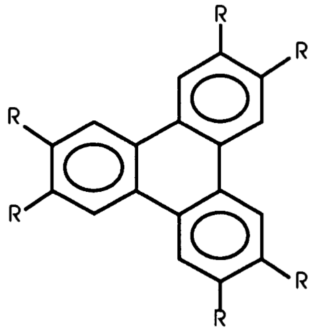
To simulate a system of such molecules using MD requires integrating EOMs for many atoms, even though each molecule has a much simpler behaviour as a rigid rod or disk.

Liquid crystals

Consider liquid crystals (LC) that are rod-like or disk-like; as a function of temperature or density they take up different phases that are crucial for their technological function, e.g., LCD displays.



8 OSI



Hexasubstituted triphenylene

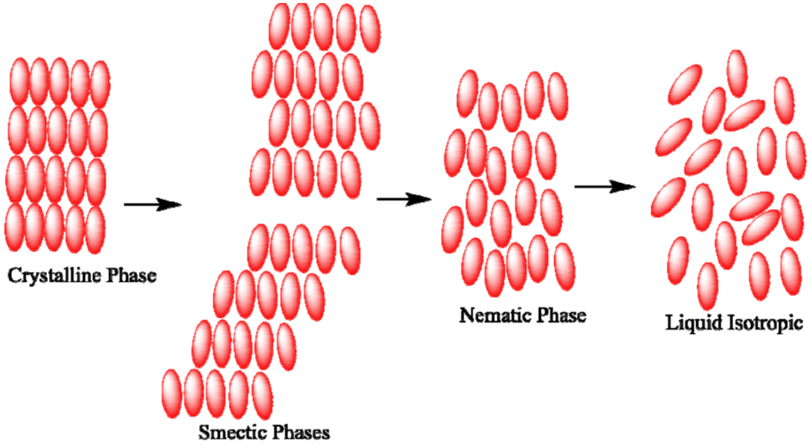
We want to answer questions like: how do the molecules arrange themselves in space as temperature or density are changed?

Can we predict properties of such systems without keeping all the atomic coordinates?

Some LCs are rigid rods

Rigid rods are a simple model of complex molecules, e.g., viruses, bacteria, liquid crystals, actin filaments, proteins, etc. These are too complex to simulate using MD.

Consider the problem of N rigid rods moving freely (but without intersecting) in a fixed volume under thermal motion, what is their equilibrium state as a function of density ?



M.A. Qaddoura and K. D. Belfield, *Int. J. Mol. Sci.* **2009**, *10*(11), 4772-4788

Intuitively, we expect that at low densities, they move and are oriented randomly. But as the density is increased, they start to "notice" each other's presence and become ordered: this is the Isotropic-Nematic (I-N) phase transition: unlike the steam-water-ice transitions it is driven by density not temperature.

Hamiltonian

The Hamiltonian of a system is its total energy (kinetic plus potential) as a function of its d.o.f. (often position and momentum coordinates):

$H(\{x\}) = H(x_1, x_2, x_3, \ldots x_N; p_1, p_2, p_3, \ldots p_N) = K(\{p\}) + V(\{x\})$

In classical mechanics, the Hamiltonian defines the equation of motion of a mechanical system, which in turn determines the trajectory of the system through phase space.

$$dp_i/dt = - \partial H/dq_i$$

$$dq_i/dt = \partial H/dp_i$$

But for many interesting problems, instead of integrating these EOM for the system, we use the Hamiltonian to define the energy of a state of the system and *construct* a trajectory through phase space (that may or may not be a physically-realisable one) corresponding to a physically-meaningful thermodynamic ensemble, and calculate the observable properties of the system as averages over this trajectory.

Phase space

Given a set of N d.o.f {x_i}, the *Phase Space* of a system is the N-dimensional space where each coordinate axis is labelled by one *coordinate* or degree of freedom (d.o.f). Each point in this space represents a *state* of the system, i.e., it is an assignment of a value to each d.o.f.

Blackboard: a point mass in 1d has two coordinates (x,v). The XY plane is the phase space of this particle, its *(position, velocity)* is its state and this can be any point in the plane. If the system's total energy is constant, its phase space trajectory is restricted to a subset of the plane, e.g., SHM of a pendulum has an ellipse for its trajectory in phase space and has constant total energy.

Note that phase space does not *exist* but is a mathematical construct to help calculations

(what does it mean here to say it doesn't exist?)

As a system evolves in time, it moves through its phase space subject to *external constraints* imposed on it - this defines the *ensemble*. Monte Carlo simulations are a means of calculating observables by averaging over a system's trajectory in phase space.

There are many different ***ensembles***, which just means there are many different sets of constraints that we can impose on a system's motion through its phase space.

Ensembles, energy and moves

We have to define the thermodynamic ensemble of our system before we can calculate anything. Typically, the mass or number of particles is constant as well as:

Microcanonical (N,V,E):	Energy = constant
Canonical (N,V,T):	Temperature = constant
Grand canonical (μ,V,T):	Chemical potential = constant

Many MC simulations are done in the canonical ensemble with either (N,V,T) or (N, p, T) constant as these correspond to many physical systems.

- Hamiltonian = PE of N particles in a volume V at constant temperature T
- PE is independent of the velocities, so the KE contribution cancels in ensemble averages
- Particles make *moves* that change their coordinates, but these moves do not necessarily represent physical motion through real space
- Instead of an EOM we calculate observables as averages over their values obtained as the system moves through its phase space.

MC replaces the need to integrate Newton's laws by a need to evaluate a high-dimensional integral; instead of a complex force calculation we need an efficient quadratures scheme: the Metropolis algorithm.

Monte Carlo simulation

MC simulations calculate high-dimensional integrals of the form

$$\langle A \rangle = 1/Z \int d\mathbf{x}\, A(\{\mathbf{x}\})\, e^{-\beta H(\{\mathbf{x}\})}$$

$$Z = \int d\mathbf{x}\, e^{-\beta H(\{\mathbf{x}\})}$$

where H is the Hamiltonian of the system, Z is the Partition Function, and A is a function of all the coordinate $\{\mathbf{x}\}$ that represents an observable.

We cannot select points at random in the phase space for evaluating these integrals because many of them will have large values of the energy, and the Boltzmann weight, $e^{-\beta H(\{\mathbf{x}\})}$, will kill off their contribution to the integral.

We also cannot calculate the integrals using quadratures as their dimension is too high:

e.g., N particles in a box of size L in 3d have 3N d.o.f, and a uniform grid will have $(L/\Delta L)^{3N}$ points, and for $N \sim 1000$ and $\Delta L/L \sim 0.01$ (a small system) this is impossible.

# of grid points at which we must evaluate the integrand $\sim 100^{3000}$

N. Metropolis et al. Equation of State Calculations Using Fast Computing Machines, J. Chem. Phys. 21:1087 (1953)

Chapter 1, Monte Carlo Methods in Statistical Physics, Ed. K Binder, Topics in Current Physics, Vol 7, Springer-Verlag, 1986

Metropolis MC 1) *Moves*

Consider a set of disks in a 2D box interacting via a potential $V(r_{ij})$ that depends only on the magnitude of the their relative separation $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ (i.e., a central force).

Define a small square centred on each particle of side length 2Δ, and randomly choose a small displacement within this box (dx, dy) such that $-\Delta \leq dx \leq \Delta$ (and for dy) and set:

$X \rightarrow X + dx$
$Y \rightarrow Y + dy$

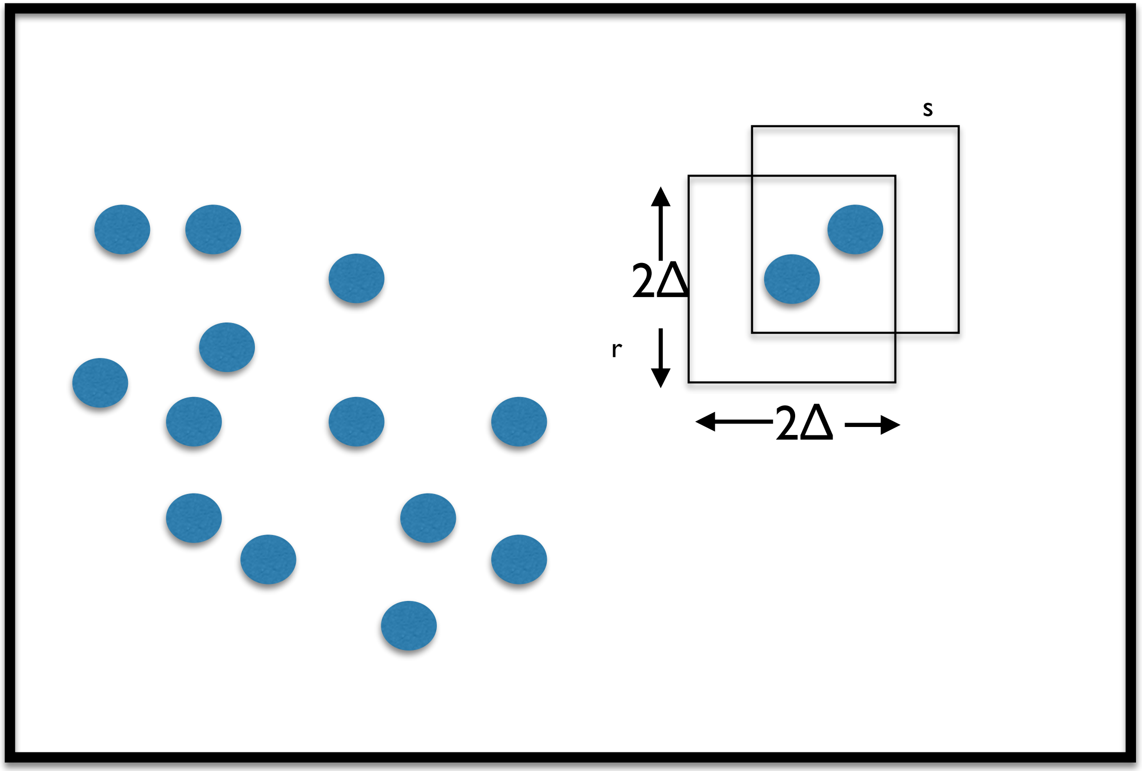
Then $r \leftarrow (X,Y)$
$s \leftarrow (X + dx, Y + dy)$

Let P_{rs} be the *a priori* probability of a move from state r to state s.

We must have

$$P_{rs} = P_{sr}$$

which is true here because it is equally likely to choose a positive displacement as a negative one.



Metropolis MC 2) *Importance sampling*

Now consider an *ensemble* of many copies of the system, and let N_r systems in the ensemble be in state r, and N_s in state s.

Let the two states r, s have energies $H(r) > H(s)$

i.e., state r has a higher energy than state s. We expect that in equilibrium $N_r < N_s$ as it has higher energy, and in the canonical ensemble, $N_s \sim e^{-\beta H(s)}$

Metropolis MC importance sampling has the rules:

1) a move from a state of higher energy to one of lower energy is accepted with prob. 1
2) a move from an initial state s with lower energy to a final one r with a higher energy is accepted with a probability proportional to $e^{-\beta(H(r) - H(s))}$

How many systems in the ensemble move from state r to s and vice versa?

r → s $N_r . P_{rs}$ as all moves from higher to lower energy are accepted

s → r $N_s . P_{sr} . \exp(-\beta(H(r) - H(s))$ as moves from lower to higher energy are accepted with prob. proportional to the Boltzmann factor

Detailed balance

The **net number of systems** moving from state **s** to state **r** is (recall $P_{rs} = P_{sr}$ and $H(r) > H(s)$)

$$N_s \,.\, P_{sr} \,.\, \exp(-\beta(H(r) - H(s)) - N_r \,.\, P_{rs}$$

or,

$$N_s \,.\, P_{rs} \,.\, (\exp(- \beta(H(r) - H(s)) - N_r / N_s)$$

So, if N_r / N_s, is bigger than the exponential there is a net movement from state r to state s, and if N_r / N_s is smaller than the exponential, there is a net movement from s to r, so eventually

$$\exp(- \beta(H(r) - H(s)) - N_r / N_s = 0$$

i.e.,

$$N_r / N_s = \exp(-\beta H(r)) / \exp(-\beta H(s))$$

Which can be written:

$$P_{eq}(r).P(r \rightarrow s) = P_{eq}(s).P(s \rightarrow r)$$

which is the condition of *detailed balance*. It guarantees that our moves will eventually lead us into equilibrium in the Canonical ensemble.

Metropolis MC recipe

- Define the d.o.f (=coordinates) of the system and its Hamiltonian (total energy)
- Put the system into an initial state (either randomly chosen or carefully constructed)
- Select a coordinate sequentially or at random
- Calculate the energy of the system in the current state H(r)
- Make an unbiased random move of the selected coordinate, and calculate new energy H(s), and the energy difference $\Delta E = H(s) - H(r)$
- Select a random number ζ uniformly distributed on (0,1), and apply the test

```
if ( $\Delta E < 0$ or $\zeta < e^{-\beta \Delta E}$)
    accept the move, system is now in state s
else
    reject the move, return system to state r
```

- Repeat for every coordinate (one *sweep*), and then repeat for many sweeps.
- Sample observables every so often, and calculate ensemble averages as simple averages of the sampled values.

Important: A *move* does not have to be physically-realisable, but must satisfy detailed balance.

Correlation times

Configurations generated by the Metropolis algorithm are **strongly** correlated as typically only a single d.o.f is changed at each attempt, and many of these fail.

Consider a set of disks moving in 3d space. Each move is only a small change in the position of each disk, so an observable like the average orientation will not change much between moves nor between successive sweeps.

In order to get meaningful ensemble averages, we have to take independent samples, and allow correlations to die away between samples. We use the same two-time correlation function as in MD to determine how many sweeps between samples are needed.

$$C_2(\tau) = (\langle O(t+\tau).O(t)\rangle - \langle O(t)\rangle^2) \; / \; (\langle O(t)^2\rangle - \langle O(t)\rangle^2)$$

The "time" variable is just the number of sweeps and does NOT represent the evolution of the system in real time.

Time and realism in MC simulations

In the LC example, each disk made moves that are physically possible: rotations and translations.

But we could have included moves that destroy a disk at one place and reinsert it elsewhere. There is no concept of "time" in these changes of configuration, nor must moves be continuous in space.

Recall that we are moving through *phase space* not real space. Any configuration that is physically possible can change to any other possible configuration - with a certain probability. All that matters is that we satisfy detailed balance and have a reasonable success rate so that we generate a trajectory along which we can measure accurate ensemble averages.

In polymer simulations, typical unrealistic moves include breaking a chain and reattaching a piece to another polymer for example.

However, we can sometimes interpret the moves dynamically, and define a Monte Carlo *time* from the rate, e.g., disks move in space.

This motion will NOT be the same as that generated by solving Newton's laws for the moving rods but in equilibrium the states will be sampled correctly, i.e., with the probability given by the Boltzmann factor.

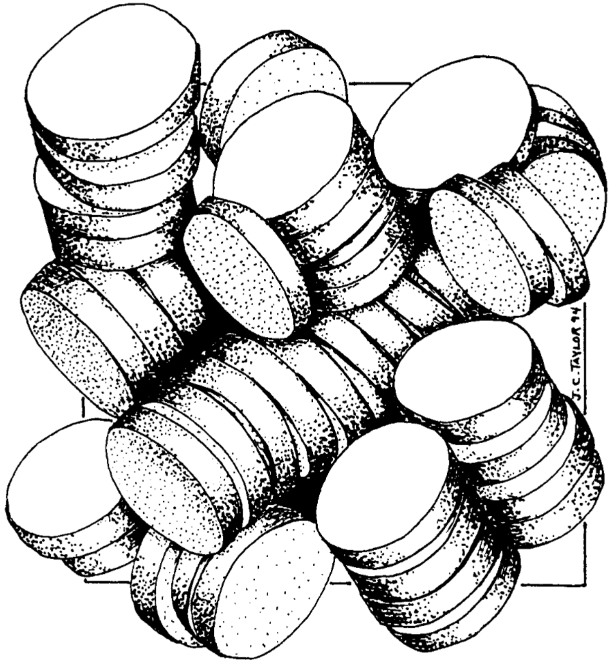
Discotic liquid crystals

Consider the example of a 3D box containing N hard, rigid LC molecules, e.g., disks.

They have no long-range interactions (uncharged) but do have (short-range) steric repulsion.

All disks diffuse around randomly in space under random, thermal motion.

This corresponds to: H = 0 unless two disks intersect when H = ∞. So any configuration of the disks is allowed except those that have an intersection.



This example is artificially simple because there is no energy in the system: disks either don't overlap with energy zero, or they overlap with infinite energy (i.e., overlaps are forbidden).

Question. If there is no temperature, what does the Boltzmann factor $\exp(-V(x)/k_B T)$ do in the averages <A>?

LC moves

Each disk has 5 d.o.f:

3 translational for the CM (x, y, z)
2 orientational (θ, φ) for its normal vector's orientation in a fixed coordinate system.

We implement diffusion by making small changes in each d.o.f that are **a)** small enough that some succeed, **b)** large enough to ensure that system moves "efficiently" through its phase space.

Moves MUST be uncorrelated and symmetric otherwise the equilibrium state will NOT be the desired thermodynamic ensemble but some other, non-equilibrium, steady state.

3 translational for the CM (x, y, z)
2 orientational (θ, φ) for its normal vector's orientation in a fixed coordinate system.

$x \longrightarrow x + dx$ (similarly for y and z) where $-\Delta \leq dx \leq \Delta$
$\theta \longrightarrow \theta + d\theta$ where $-1 \leq \cos\theta \leq 1$ (Warning - don't use $-\Delta\theta \leq d\theta \leq \Delta\theta$)
$\varphi \longrightarrow \varphi + d\varphi$ where $-\Delta\varphi \leq d\varphi \leq \Delta\varphi$ and

and θ, φ are periodic in (0,π) and (0, 2π) respectively. Note that the ranges are chosen so that the success rate is around 0.1 - 0.5 of all attempts for each type of move. This ensures that the system moves through its phase space reasonably fast.

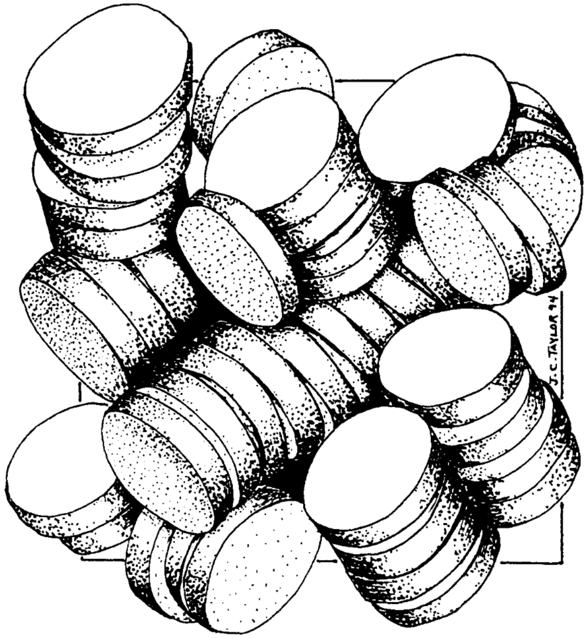
I-N Phase Transition

As a function of density disks/rods take up different phases, viz, isotropic, nematic, ...

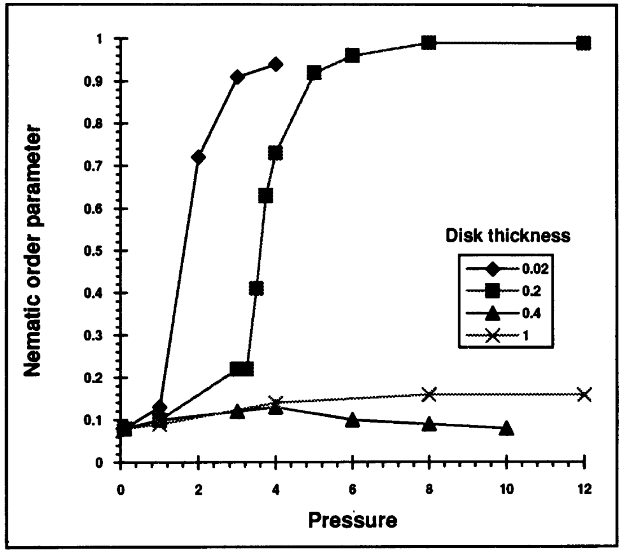
At low density, disks/rods are randomly oriented - *isotropic* phase

At high density, they align in a preferred direction - the *nematic* phase

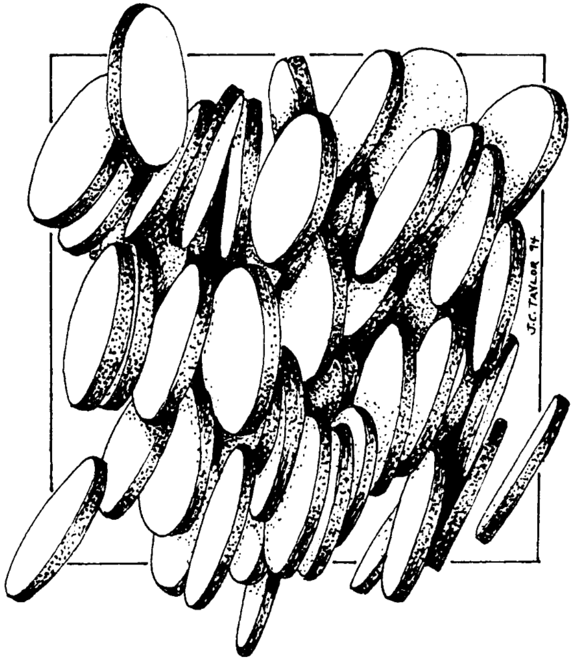
How does the molecular shape influence the **I-N transition**?



$T/R = 0.4$
$NR^3/V = 0.51$



Nematic order parameter for disks of different thicknesses
$T/R = 0.02\ 0.2\ 0.4\ 1.0$ in 3d under an applied pressure



$T/R = 0.2$
$NR^3/V = 0.57$

Summary

- Mechanical simulations yield (expensive) real-time trajectories
- If you don't need this detail, Monte Carlo simulations yield equilibrium properties of a system typically must faster than MD
- MC can have unrealistic moves that speed up motion through phase space