A brief interlude on Monte Carlo

Averages by integration: $\langle A \rangle = \int d\vec{r}\, d\vec{p}\; A(\vec{r},\vec{p})\, P(\vec{r},\vec{p})$

or from states: $= \sum_{\{s\}} A_{state}\, P_{state}$

In general: $P_{state} = e^{-\beta E_{state}} \cdot \frac{1}{Q}$

$P(\vec{r},\vec{p}) = e^{-\beta \mathcal{H}(\vec{r},\vec{p})} \cdot \frac{1}{Q}$

If the property we are interested in depends only on configurational information: $A = A(\vec{r})$

$$P(\vec{r}) = \frac{e^{-\beta U(\vec{r})}}{Z}$$ ← configurational partition function

$$Z = \int d\vec{r}\, e^{-\beta U(\vec{r})}$$ ← what is Z for Ising model?

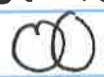
A brute force way to do averaging:

$$\langle A \rangle = \frac{\sum_{i=1}^{N_{trial}} A(\vec{r}_i)\, e^{-\beta U(\vec{r}_i)}}{\sum_{i=1}^{N_{trial}} e^{-\beta U(\vec{r}_i)}}$$

each trial (i) corresponds to randomly choosing all $3N$ positions for all N atoms (e.g. $\vec{r}_i$)

Many of these trials will be unimportant parts of space, i.e. where $e^{-\beta U(\vec{r}_i)} \approx 0$

overlapping atoms



unrealistic bond lengths

Sampling via brute force is inefficient. How do we fix this?

Metropolis MC

Suppose we have a simple 2 level system

— E_1 ← excited state

— E_0 ← ground state

Also, suppose we know that $e^{-\beta E_0} \approx 2e^{-\beta E_1}$ (the energies are such that the ground state is twice as likely as the excited state.)

So: $$P_0 = \frac{e^{-\beta E_0}}{e^{-\beta E_0} + e^{-\beta E_1}} = \frac{2e^{-\beta E_1}}{2e^{-\beta E_1} + e^{-\beta E_1}} = \frac{2}{3}$$

$$P_1 = \frac{e^{-\beta E_1}}{e^{-\beta E_0} + e^{-\beta E_1}} = \frac{e^{-\beta E_1}}{2e^{-\beta E_1} + e^{-\beta E_1}} = \frac{1}{3}$$

The central idea is to pick states so that we visit state 0 twice as often as state 1.

We use a transmission matrix π

$$\begin{pmatrix} P_0(2) \\ P_1(2) \end{pmatrix} = \pi \cdot \begin{pmatrix} P_0(1) \\ P_1(1) \end{pmatrix}$$

↑ probabilities after 2nd step

← probabilities in 1st step

After n steps:

$$\begin{pmatrix} P_0(n) \\ P_1(n) \end{pmatrix} = \pi^{n-1} \begin{pmatrix} P_0(1) \\ P_1(1) \end{pmatrix}$$

In this particular case π is the 2x2 matrix:

$$\pi = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{pmatrix}$$

Let's try an experiment. If we start our system in the ground state $\begin{pmatrix} P_0(1) \\ P_1(1) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \vec{P}(1)$

$$\vec{P}(2) = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

$$\vec{P}(3) = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 3/4 \\ 1/4 \end{pmatrix}$$

$$\vec{P}(4) = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 3/4 \\ 1/4 \end{pmatrix} = \begin{pmatrix} 0.625 \\ 0.375 \end{pmatrix}$$

$$\vec{P}(\infty) = \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}$$

Suppose we had started our experiment in the excited state: $\vec{P}(1) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$$\vec{P}(2) = \begin{pmatrix} 1/2 & 1 \\ 1/2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

← same as $\vec{P}(1)$ from previous experiment

$$\vdots$$

$$\vec{P}(\infty) = \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}$$

What is special about $\underline{\pi}$?

$$\vec{P}(\infty) = \underline{\pi} \cdot \vec{P}(\infty)$$

← large N limit must approach equilibrium

$$\begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix} = \begin{pmatrix} 1/2 & 1 \\ 1/2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}$$ ✓

In an ensemble at equilibrium, the transition matrix must return the equilibrium probabilities for state j from the equilibrium probabilities of the other states:

$$\sum_i P_i(\infty)\, \pi_{ij} = P_j(\infty)$$

We can separate the weights or the equilibrium population information from the purely random part:

$$\pi_{mn} = \alpha_{mn} P_{mn}$$

← probability of going from state m to n

π → transition matrix

P ← purely random stochastic matrix

P_{mn} contains equilibrium information about the relative state probabilities ρ_m & ρ_n

α_{mn} contains information about how the states are chosen (e.g. if we are in state m what probability is there of choosing state n)

The Metropolis Rules

1) $\pi_{mn} = \alpha_{mn}$ if $\rho_n \geq \rho_m$

2) $\pi_{mn} = \alpha_{mn}\left(\frac{\rho_n}{\rho_m}\right)$ if $\rho_n < \rho_m$

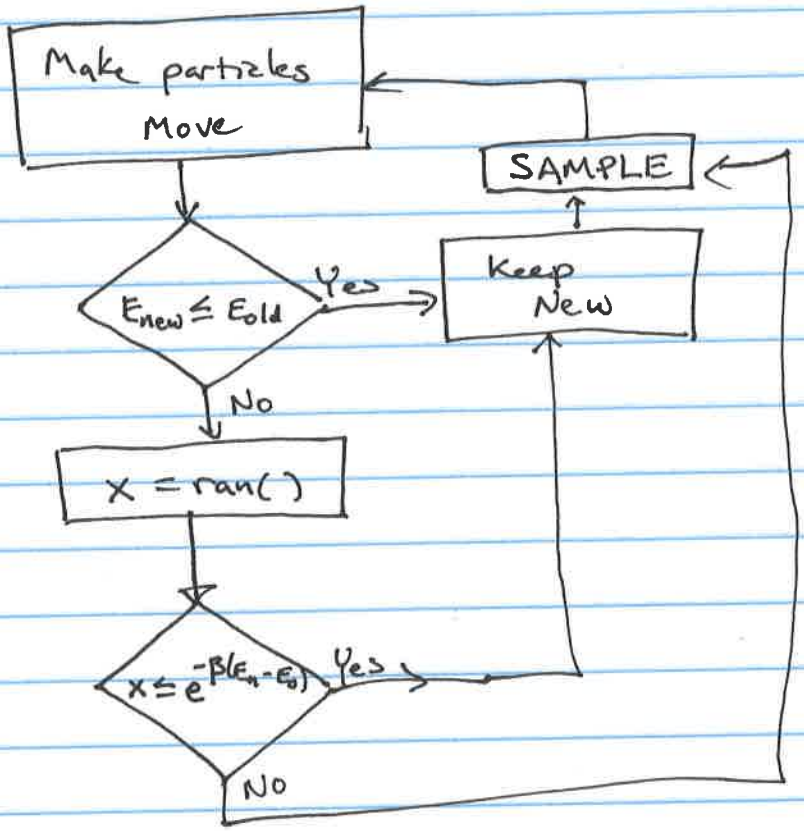
3) $\pi_{mm} = 1 - \sum_{m' \neq m} \pi_{m'm}$

1) If we pick (randomly) a new state n that is more probable than our current state m, then transition probability is just the probability of ~~keeping~~ picking it in the first place. (e.g. keep the new state!)

2) If we pick (randomly) a new state that is less probable, then we might keep it, but the probability of keeping it depends on the relative probabilities of the 2 states.

3) If we pick the same state randomly, we need to figure out the probabilities of all other transitions. In systems with many states, this rarely happens.

MC Flowchart



Some subtlety: If we make big moves, we'll probably reject most of them. This is inefficient sampling. Efficient sampling has $P_{accept} \approx \frac{1}{2}$.

Make Moves : $x_{new} = x_{old} + (2\xi - 1)\delta r_{max}$

ξ : random on $[0,1]$

δr_{max} : max motion along 1 direction

If P_{accept} is too small after 1000 steps, adjust $\delta r_{max} \downarrow$

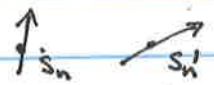
If P_{accept} is too large after 1000 steps, adjust $\delta r_{max} \uparrow$

Heisenberg & X-Y models

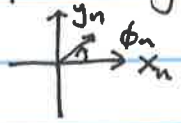
$$\mathcal{H} = -\frac{1}{2} \sum_{n,n'} J_{nn'} \vec{S}_n \cdot \vec{S}_{n'}$$

$\vec{S}_n$ is a vector of length 1 for all n

$$\vec{S}_n = \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}, \qquad x_n^2 + y_n^2 + z_n^2 = 1$$



In the X-Y model, only 1 angle describes the orientation of each vector:



$$\vec{S}_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} \cos\phi_n \\ \sin\phi_n \end{pmatrix}, \quad \text{so} \quad \vec{S}_n \cdot \vec{S}_{n'} = \cos\phi_n \cos\phi_{n'} + \sin\phi_n \sin\phi_{n'} = \cos(\phi_n - \phi_{n'})$$

So the X-Y Hamiltonian can be written:

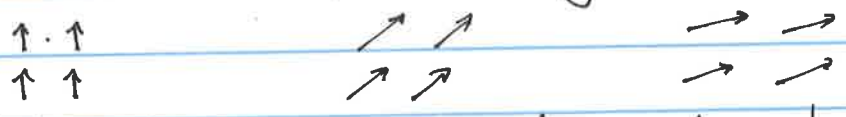
$$\mathcal{H} = -\frac{1}{2} \sum_{n,n'} J_{nn'} \vec{S}_n \cdot \vec{S}_{n'} = -\frac{1}{2} \sum_{n,n'} J_{nn'} \cos(\phi_n - \phi_{n'})$$

What <u>symmetries</u> exist in this model?

That is, in the Ising model, we can flip <u>all</u> spins and not change the energy

$\sigma_n \to -\sigma_n$ is an <u>inversion</u> symmetry

What can we do to a low-energy Heisenberg or X-Y model & not change the energy?



In the 2D X-Y model, there is no spontaneously broken symmetry (No 2D magnets in X-Y model).

In 3D, there is. The theorem that tells us this is called the Mermin-Wagner theorem

The Nearest-Neighbor XY model in 2D:

$$\mathcal{H} = -J/2 \sum_{n,n'} \vec{S}_n \cdot \vec{S}_{n'} = -\frac{J}{2} \sum_{n,n'} \cos(\phi_n - \phi_{n'})$$

Lowest energy = all ϕ_n values identical

A low energy excitation would be a block wave

In a big system, this excitation can be made arbitrarily long in wavelength and thus arbitrarily close in energy to the ground state.

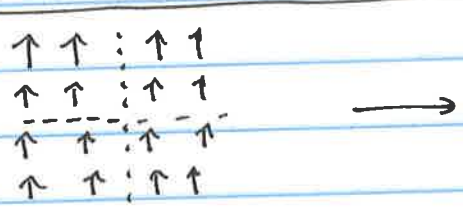
What does this do to $\langle \vec{S} \rangle$?

Following the argument of the Ising model,

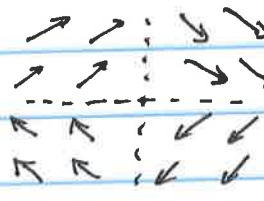
↑↑↑↓↓↓↓ $\Delta E = 2J$, but many places to make this error,

ΔE of a block wave is small, but there are a huge number of low-E block waves to make! $\therefore \langle \vec{S} \rangle = 0$

The Kosterlitz-Thouless transition



perfectly aligned



a vortex with a core size of a

each quadrant is rotated by 90° relative to neighbors

What's the energy required to make a vortex like this?

Energy along 1 of the lines (length L):

$$-J\sum_{n}\sum_{n'}\vec{s}_n\cdot\vec{s}_{n'} = \begin{cases} -JL & \text{if perfectly aligned} \\ 0 & \text{if 90° rotated} \end{cases}$$

$$\Delta E = 2JL \quad \leftarrow \text{a high energy vortex}$$

Cheaper vortices can be made by gently approaching from far away

← we recover alignment far from vortex core!

Kosterlitz & Thouless were able to show that vortices can be created as cheaply as $x \ln(L/a)$

(x: J dependent; L: system size; a: vortex core radius)

There are also $\left(\frac{L}{a}\right)^2$ places to put a vortex.

So, entropically, $S = 2k_B \ln(L/a)$

And, energetically, $E = x \ln(L/a)$

So the free energy for vortex formation:

$$A = E - TS = (x - 2k_B T)\ln\left(\frac{L}{a}\right)$$

When $A > 0$, the creation of vortices is thermodynamically favored. When $A < 0$, the elimination of vortices is favored

$$\boxed{T_c = \frac{x}{2k_B}}$$

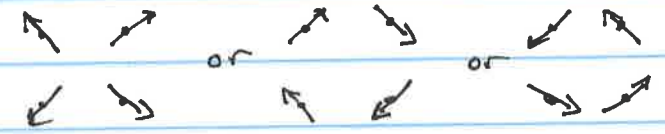
So, what does this mean?

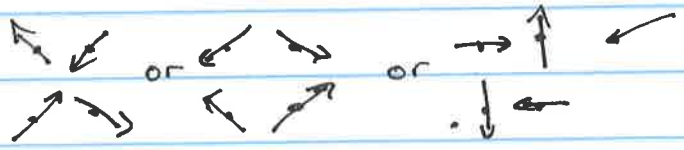
At hig T, vortices are easily created and destroyed, the system looks random.

At low T, vortices are Stable and even though the system is not a spontaneous magnet, there is long-range order.

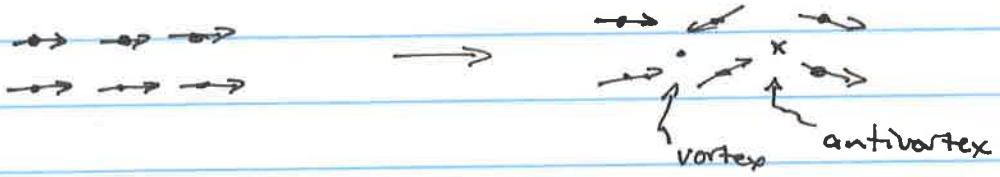
This is an order-disorder phase transition!

The X-Y model also exhibits excitations called anti-vortices!

Vortices:  vortex patterns don't change if we rotate around the vortex.

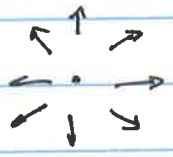
Anti vortices:  anti vortex patterns change sign for 90° rotations of all spins.

Vortices are born together with antivortex siblings:

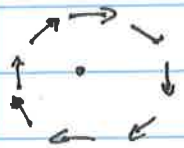


Consider a ring of XY spins around some point

$\phi(\vec{r}) = \phi_0$



$\phi(\vec{r}) = \theta$ (spin orientation is the same as spatial angular coordinate)



$\phi(\vec{r}) = \theta + \frac{\pi}{2}$

$\phi(\vec{r})$ describes the state of the system around 1 point and we may use it to find the energy of the system:

$$\mathcal{H} = -J \sum_n \sum_{n'} \vec{S}_n \cdot \vec{S}_{n'} = -J \sum_n \sum_{n'} \cos(\phi_n - \phi_{n'})$$

If $\phi_n \approx \phi_{n'}$ then $\cos(\phi_n - \phi_{n'}) \approx 1 - \frac{\delta\phi^2}{2}$

$$\mathcal{H} = \sum_n \sum_{n'} -J\left(1 - \frac{\delta\phi^2}{2}\right)$$

$$= \underbrace{-JN2d}_{\#\text{ of Nearest Neighbors}} + \frac{J}{2} \sum_n \sum_{n'} \delta\phi_{nn'}^2$$

If ρ is the density of spins, the distance between spins is $\frac{1}{\rho}$, and we can estimate

$$\delta\phi = \frac{1}{\rho} \frac{d\phi}{dx}$$

We can estimate the energy of a spin configuration as:

$$\mathcal{H} = \frac{J}{2\rho^2} \int dx \int dy \left(\left(\frac{\partial \phi}{\partial x}\right)^2 + \left(\frac{\partial \phi}{\partial y}\right)^2 \right) - JN2d$$

Long-Range Order in the XY model

Consider this correlation function:

$$\langle(\phi_n - \phi_{n'})^2\rangle = \langle(\Delta\phi)^2\rangle = \frac{\int\int\cdots\int d\phi_1 \cdots d\phi_N \, e^{-\beta\mathcal{H}} \, (\Delta\phi)^2}{Z}$$

If there's no long range order $\langle\phi\phi\rangle^2$ will vanish for large $n-n'$.

If there <u>is</u> a spontaneous symmetry breaking, there will be a T_c below which $\langle(\Delta\phi)^2\rangle \neq 0$.

Last time, we approxmated the X-Y Hamiltonian:

$$\mathcal{H} = \frac{-J}{2}\sum_{n,n'} \vec{S}_n \cdot \vec{S}_{n'} = \frac{-J}{2}\sum_{n,n'} \cos(\phi_n - \phi_{n'})$$

for small deviations in ϕ: $\cos(\delta\phi) \sim 1 - \frac{\delta\phi^2}{2}$, so:

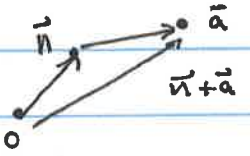
$$\mathcal{H} \approx -\frac{J}{2} + \frac{J}{4}\sum_n \sum_\alpha (\phi_n - \phi_{n+\alpha})^2$$

For a <u>block wave</u> we can have a Fourier series describe the action of that wave:

$$\phi_{\vec{k}} = \frac{1}{\sqrt{N}}\sum_{\vec{n}} \phi_{\vec{n}} \, e^{-i\vec{k}\cdot\vec{n}}$$ ← defines a wave with <u>wave vector</u> $\vec{k}$.

$$\phi_{\vec{n}} = \frac{1}{\sqrt{N}}\sum_{\vec{k}} \phi_{\vec{k}} \, e^{i\vec{k}\cdot\vec{n}}$$ ← defines a spatial arrangement under the action of many block waves (each with a different $\vec{k}$).

To describe the angle at a different location, we need to use some vector addition:



$$\phi_{\vec{n}+\vec{a}} = \frac{1}{\sqrt{N}} \sum_{\vec{k}} \phi_{\vec{k}} e^{i\vec{k}\cdot(\vec{n}+\vec{a})}$$

Let's try to put these into the Hamiltonian

$$\phi_{\vec{n}} - \phi_{\vec{n}+\vec{a}} = \frac{1}{\sqrt{N}} \sum_{\vec{k}} \phi_{\vec{k}} \left(e^{i\vec{k}\cdot\vec{n}} - e^{i\vec{k}(\vec{n}+\vec{a})} \right)$$

$$= \frac{1}{\sqrt{N}} \sum_{\vec{k}} \phi_{\vec{k}} e^{i\vec{k}\cdot\vec{n}} \left(1 - e^{i\vec{k}\cdot\vec{a}} \right)$$

A product $(\Delta\phi)^2$ will require 2 independent sums over $\vec{k}$:

$$(\phi_n - \phi_{n+a}) = (\phi_n - \phi_{n+a})^* \cdot (\phi_n - \phi_{n+a})$$

$$= \left(\frac{1}{\sqrt{N}} \sum_k \phi_k^* e^{-ikn} \left(1 - e^{-ika}\right) \right) \left(\frac{1}{\sqrt{N}} \sum_{k'} \phi_{k'} e^{ik'n} \left(1 - e^{ik'a}\right) \right)$$

$$= \frac{1}{N} \sum_k \sum_{k'} \phi_k^* \phi_{k'} e^{in(k'-k)} \left(1 - e^{-ika}\right)\left(1 - e^{ik'a}\right)$$

So:

$$\mathcal{H} \approx \frac{-J}{2} + \frac{J}{4N} \sum_n \sum_a \sum_k \sum_{k'} \phi_k^* \phi_{k'} e^{in(k'-k)} \left(1 - e^{-ika}\right)\left(1 - e^{ik'a}\right)$$

We can recognize:

$$\delta(x) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{inx} \quad \text{when} \quad -2\pi \le x \le 2\pi$$

So $$\frac{1}{N} \sum_n e^{in(k'-k)} = 2\delta_{kk'}$$

So: $\mathcal{H} = \frac{-J}{2} + \frac{J}{2}\sum_a \sum_k \phi_k^* \phi_k \left(1 - e^{-ika}\right)\left(1 - e^{ika}\right)$

$$= \frac{-J}{2} + \frac{1}{2}\sum_k |\phi_k|^2 J \sum_a |1 - e^{ika}|^2$$

$$\mathcal{H} = \frac{-J}{2} + \frac{1}{2}\sum_k J_k |\phi_k|^2$$

Here we've defined $J_k = J\sum_a |1 - e^{ika}|^2$

$$= J\sum_a (1 - \cos ka)$$

Let's make sure we understand this:

$$\mathcal{H} = \text{const} + \frac{1}{2}\sum_k J_k |\phi_k|^2$$

Each ϕ_k is <u>independent</u> (no coupling between them) so they can be treated as a <u>normally</u> distributed quantity:

$$\langle |\phi_k|^2 \rangle = \int d\phi_k |\phi_k|^2 e^{-\beta J_k \phi_k^2/2}$$

$$= \frac{k_B T}{J_k} \quad \leftarrow \text{this is a statement of } \underline{\text{equipartition}}$$

To solve the correlation function, we need:

$$\langle (\Delta\phi)^2 \rangle = \langle (\phi_0 - \phi_n)^2 \rangle = \langle \phi_0^2 \rangle + \langle \phi_n^2 \rangle - 2\langle \phi_0 \phi_n \rangle$$

We still need the <u>cross</u> term: ↑ <u>cross</u> correlation

$$\langle \phi_n \phi_{n'} \rangle = \frac{1}{N}\sum_k \sum_{k'} \langle \phi_k^* \phi_{k'} \rangle e^{-ik\cdot n} e^{ik'\cdot n'}$$

And to get the cross term, we again need:

$$\langle \phi_k^* \phi_{k'} \rangle = \delta_{kk'} \langle |\phi_k|^2 \rangle = \delta_{kk'} \frac{k_B T}{J_k}$$

$$\therefore \quad \langle \phi_n \phi_{n'} \rangle = \frac{1}{N} \sum_k \langle |\phi_k|^2 \, e^{ik(n'-n)} \rangle$$

$\therefore$

$$\langle (\Delta\phi)^2 \rangle = \langle \phi_n^2 \rangle + \langle \phi_{n'}^2 \rangle - 2\langle \phi_n \phi_{n'} \rangle$$

$$= \frac{1}{N} \sum_k \langle |\phi_k|^2 \rangle + \frac{1}{N} \sum_k \langle |\phi_k|^2 \rangle - 2\frac{1}{N} \sum_k \langle |\phi_k|^2 \cos k_n \rangle$$

$$\boxed{\langle (\Delta\phi)^2 \rangle = \frac{2}{N} \sum_k \frac{k_B T}{J_k} (1 - \cos \vec{k} \cdot \vec{n})}$$

In the limits of large systems, the black waves become continuous, so

$$\sum_k \longrightarrow \left(\frac{a}{2\pi}\right)^d \int d^d k$$

$\int d^d k$ ← d-dimensional integral

$\left(\frac{a}{2\pi}\right)^d$ ← granularity of angular change allowed due to lattice

$$d^d k = k^{d-1} \, dk \, d\hat{k} \quad \leftarrow \begin{cases} \text{in 2D: } r\,dr \\ \text{in 3D: } r^2 dr \end{cases}$$

$$\langle (\Delta\phi)^2 \rangle = \frac{2 k_B T \, a^d}{(2\pi)^d} \int d^d k \, \frac{1 - \cos(k \cdot n)}{J_k}$$

Remember, $J_k = J \sum_a (1 - \cos k a)$

These integrals can be done analytically:

In 1-D: $\langle(\Delta\phi)^2\rangle = \left(\frac{2k_BT}{Ja}\right) n$

2-D: $\langle(\Delta\phi)^2\rangle = \frac{k_BT}{J\pi} \ln\frac{n}{a}$

3D: $\langle(\Delta\phi)^2\rangle = \frac{k_BT}{\pi^2 a^2 J}\underbrace{\left(k_c a - \frac{\pi a}{4n}\right)}$

↑ note the critical crossing!