Statistical Mechanics Core: Lecture #12

October 25, 2017

Our goal is to study phenomena such as phase transitions for many-body systems of N interacting particles.

Recall that the expectation value of a quantum mechanical observable quantity Q is given by

$$\langle Q \rangle = \mathrm{Tr}(Q\rho)$$

In the canonical ensemble, we have:

$$\rho = \frac{1}{Z} e^{-\beta H}$$

$$\langle Q \rangle = \frac{1}{Z} \mathrm{Tr}(Q e^{-\beta H})$$

where:

- $\beta = 1/(k_B T)$ (we will set $k_B = 1$)
- H is the Hamiltonian
- $Z = \mathrm{Tr}(e^{-\beta H})$ is the normalizing partition function

Today, we will consider classical systems, for which

$$\langle Q \rangle = \sum_\mu Q_\mu \rho_\mu$$

$$\rho_\mu = \frac{1}{Z} e^{-\beta E_\mu}$$

$$Z = \sum_\mu e^{-\beta E_\mu}$$

where:

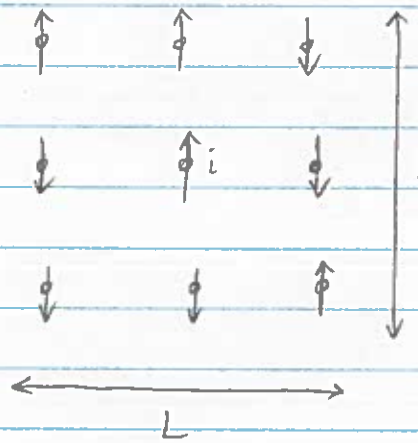
- μ represents a state of the system
- E_μ is the energy of state μ

We would like to study expectation values and, ultimately, phase transitions in the thermodynamic limit ($N \to \infty$).

For example, today we will study the expectation value of the energy:

$$\langle E \rangle = \frac{\sum_\mu E_\mu e^{-\beta E_\mu}}{\sum_\mu e^{-\beta E_\mu}}$$

Let's consider how we might evaluate the sums over μ directly for the case of a nearest-neighbour (nn) Ising model on a d-dimensional cubic lattice of length L:



$N = L^d$

On each lattice site i, there is a binary variable $\sigma_i = \pm 1$

$$H = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i$$

where:

- $\sum_{\langle ij \rangle}$ denotes a sum over nn pairs
- J is a coupling strength
- h is an external field

Exact solutions are known for:

- $d=1$ (see Tutorial 2)
- $d=2$ (Onsager, 1944)

In this case, the total number of accessible states is

$$2^N = 2^{(L^d)}$$

For $N = 25$, the fastest algorithms for computing the sums Σ'_μ require about a minute. Since the time required scales as 2^N:

N	time
25	~ 1 min.
30	~ 30 min.
35	~ 20 hours
40	~ 20 days
50	~ 60 years
80	~ 7×10^{10} years ← longer than the age of the universe!

We see that we are limited to $L = 7$ or 8 in $d = 2$

Classical Monte Carlo (MC) Methods

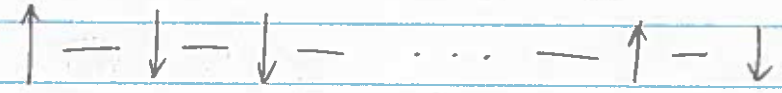
We will now introduce techniques that handle the exponential complexity of calculating expectation values $\langle Q \rangle$ by considering only M states $\mu_1, \mu_2, \ldots, \mu_M$ selected at random. We would like to get a good estimate for $\langle Q \rangle$ with $M \ll 2^N$ for large N.

For example, if we choose each of the M states μ_m $(1 \le m \le M)$ from the 2^N possible states with equal probability, then we could estimate $\langle Q \rangle$ from:

$$\langle Q \rangle \approx \underbrace{Q_M}_{\text{"estimator"}} = \frac{\sum_{m=1}^{M} Q_{\mu_m} e^{-\beta E_{\mu_m}}}{\sum_{m=1}^{M} e^{-\beta E_{\mu_m}}}$$

For $M \ll 2^N$, this estimator gives an accurate estimate of $\langle Q \rangle$ when $T \to \infty$ $(\beta \to 0)$. However, it fails at low temperatures when only a small subset of states have high probability in the original sums $\sum_{\mu}$.

Let's examine the estimator E_M for the energy of the 1d nn Ising model



$$H = -J \sum_{i=1}^{N-1} \sigma_i \sigma_{i+1} \qquad \text{(no magnetic field, open boundary conditions)}$$

From Tutorial 2:

$$Z = 2\,[2\cosh(\beta J)]^{N-1}$$

$$\Rightarrow \quad E = -\frac{\partial \log Z}{\partial \beta} = -(N-1)\,J \cdot \tanh(\beta J)$$

See ising1D_flatSampling.py: Study E_M vs. T for various M (# of samples) and N (# of spins)

More generally, if the states μ_m are taken from a probability distribution W_μ, then the estimator for $\langle Q \rangle$ is:

$$Q_M = \frac{\sum_{m=1}^{M} Q_{\mu_m} W_{\mu_m}^{-1} e^{-\beta E_{\mu_m}}}{\sum_{m=1}^{M} W_{\mu_m}^{-1} e^{-\beta E_{\mu_m}}}$$

Importance Sampling

We would like to sample the more "important" (highly probable) states more frequently.

In particular, if we sample from the Boltzmann distribution such that $W_\mu = P_\mu = e^{-\beta E_\mu}/Z$, then:

$$Q_M = \frac{1}{M} \sum_{m=1}^{M} Q_{\mu_m}$$

But <u>how</u> do we choose the states according to this distribution P_μ?

Markov Chain Monte Carlo

Idea: Given a state μ_m, move to the next state μ_{m+1} according to transition probability $T(\mu_m \to \mu_{m+1})$, where:

- $T(\mu_m \to \mu_{m+1})$ depends only on the states μ_m and μ_{m+1} (not on $\mu_1, \mu_2, \ldots, \mu_{m-1}$)
- the transition probabilities have no time dependence
- $\sum_\nu T(\mu \to \nu) = 1$

The resulting set $\{\mu_1, \mu_2, \ldots, \mu_M\}$ is called a Markov chain of states.

In order to reach states from the target probability distribution W_μ, we must design an algorithm that satisfies the following two conditions:

(1) Ergodicity: given two states μ and ν, it must be possible to transition from μ to ν with non-zero probability through some chain of intermediate states

$$\mu \longrightarrow \cdots \longrightarrow \nu$$

(2) Detailed Balance (DB): the rates of transition into and out of any state μ must be equal (otherwise W_μ would change with time).

$$\underbrace{\sum_\nu W_\nu T(\nu \to \mu)}_{\text{rate of transition into } \mu} - \underbrace{\sum_\nu W_\mu T(\mu \to \nu)}_{\text{rate of transition out of } \mu} = 0$$

One possible way to satisfy this condition is:

$$\boxed{W_\mu T(\mu \to \nu) = W_\nu T(\nu \to \mu)} \quad \text{DB}$$

So when $W_\mu = e^{-\beta E_\mu}/Z$:

$$\frac{T(\mu \to \nu)}{T(\nu \to \mu)} = e^{-\beta(E_\nu - E_\mu)}$$

We have freedom in how we choose the transition probabilities. Before considering possible algorithms, let's introduce the notation

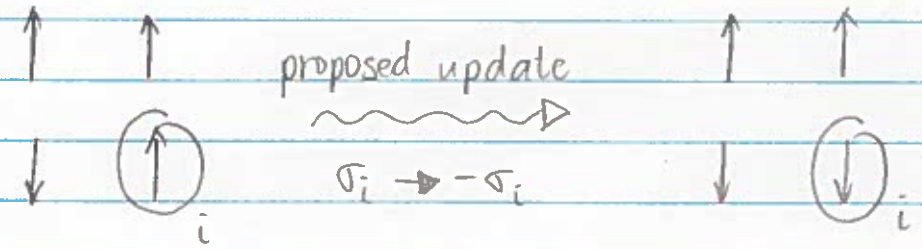
$$T(\mu \to \nu) \equiv \underbrace{g(\mu \to \nu)}_{\text{selection probability}} \underbrace{A(\mu \to \nu)}_{\text{acceptance probability}}$$

- $g(\mu \to \nu)$: probability of proposing a move to state ν given initial state μ
- $A(\mu \to \nu)$: probability of accepting the proposed move $\mu \to \nu$.

We want the acceptance probabilities to be as close to 1 as possible so that we sample as many different states as possible.

Single-Spin-Flip Algorithms for the Ising Model

Given a state μ_m of N spins at step m of the Markov chain, propose a new state μ_{m+1} that differs from μ_m by a single spin flip of the spin at some lattice site i



The probability of choosing spin i is $\frac{1}{N}$.

The selection probabilities are:

$$g(\mu \to \nu) = \begin{cases} \frac{1}{N} & \text{if } \mu \text{ and } \nu \text{ differ by a single-spin flip} \\ 0 & \text{if } \mu \text{ and } \nu \text{ differ by more than a single-spin flip} \end{cases}$$

By DB, for a single-spin flip:

$$e^{-\beta(E_\nu - E_\mu)} = \frac{A(\mu \to \nu) \cdot \frac{1}{N}}{A(\nu \to \mu) \cdot \frac{1}{N}} = \frac{A(\mu \to \nu)}{A(\nu \to \mu)}$$

Let's consider two options for choosing the acceptance probabilities.

Option #1: $A(\mu \to \nu) = A_0 e^{-\frac{1}{2}\beta(E_\nu - E_\mu)}$

We are free to choose the constant A_0 as long as $A(\mu \to \nu) \leq 1$ for all possible single-spin flips. Since we want to maximize $A(\mu \to \nu)$ and the minimum energy difference $E_\nu - E_\mu$ for a single-spin flip on a d-dimensional lattice is $-4Jd$, choose:

$$A_0 = e^{-2\beta J d}$$

$$A(\mu \to \nu) = e^{-\frac{1}{2}\beta(E_\nu - E_\mu + 4Jd)}$$

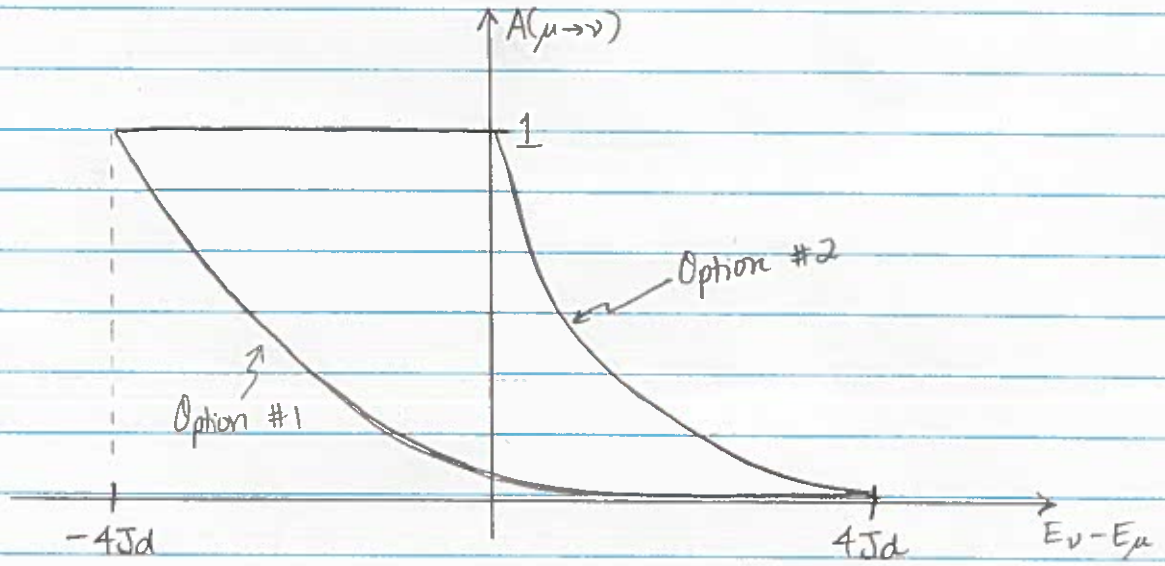
Option #2: Metropolis Algorithm (1953)

For a general proposed move $\mu \to \nu$ and target distribution W_μ:

$$A(\mu \to \nu) = \min\left(1, \frac{g(\nu \to \mu)}{g(\mu \to \nu)} \frac{W_\nu}{W_\mu}\right)$$

For the single-spin flip ($g(\mu \to \nu) = \frac{1}{N}$) and $W_\mu = e^{-\beta E_\mu}/Z$:

$$A(\mu \to \nu) = \begin{cases} e^{-\beta(E_\nu - E_\mu)} & \text{if } E_\nu - E_\mu > 0 \\ 1 & \text{otherwise} \end{cases}$$



Single-spin-flip Metropolis Algorithm (d-dim. Ising Model)

1) Generate a random initial state μ_1

REPEAT

2) Choose a site i of the lattice at random.

3) Calculate the energy difference ΔE associated with flipping $\sigma_i \to -\sigma_i$.

4) Generate a random number $r \in [0,1)$ from a uniform distribution.

5) If $\Delta E \leq 0$ OR $r < e^{-\beta \Delta E}$, accept the flip to get μ_{m+1}. Otherwise, reject the flip ($\mu_{m+1} = \mu_m$).

Statistical Mechanics Core: Lecture #13

October 26, 2017

Main reference: Newman & Barkema, "Monte Carlo Methods in Statistical Physics", chapters 1-4, 8.3

Review:

We started with the goal of studying phase transitions for interacting systems characterized by Hamiltonian H.

We focus on the classical nearest-neighbour (nn) Ising model:

$$H = -J\sum_{\langle ij\rangle} \sigma_i \sigma_j - h\sum_i \sigma_i , \qquad (\sigma_i = \pm 1)$$

Yesterday, we saw that we can estimate observable quantities

$$\langle Q \rangle = \frac{\sum_\mu Q_\mu e^{-\beta E_\mu}}{\sum_\mu e^{-\beta E_\mu}}$$

using Markov chain Monte Carlo methods to generate a Markov chain $\{\mu_1, \mu_2, \ldots, \mu_M\}$ of states sampled from the distribution $p_\mu = e^{-\beta E_\mu}/Z$ and estimate $\langle Q \rangle$ from:

$$Q_M = \frac{1}{M}\sum_{m=1}^{M} Q_{\mu_m}$$

Specifically, we studied the single-spin-flip Metropolis algorithm, where we get μ_{m+1} from μ_m by choosing a random lattice site i at random and accepting the move $\sigma_i \to -\sigma_i$ with probability $A(\mu,\nu) = \min\left(1, e^{-\beta(E_\nu - E_\mu)}\right)$.

Outline for today:

- Data Analysis practices
  - ↳ Equilibration
  - ↳ Measurement correlations
- Code: Monte Carlo simulation of the 2D Ising model
- Cluster algorithms
- Specific heat and susceptibility
- Finite-size scaling
- XY model

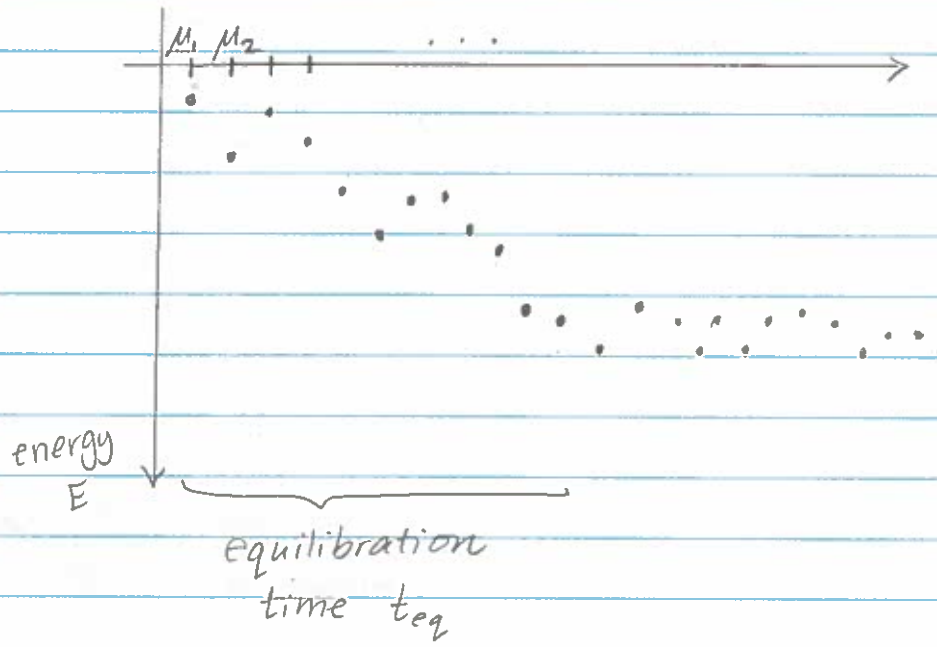
Equilibration

Recall that we can start our sampling from any state μ_1. However, we would like to (eventually) generate samples from the distribution $p_\mu = e^{-\beta E_\mu}/Z$.

In many cases, μ_1 and the subsequent $\mu_2, \mu_3, \ldots$ might have very low p_{μ_m} such that the probability of finding the system in state μ does not follow the desired distribution p_μ.

(For discussion of why we eventually approach the correct distribution, see Section 2.2.3 of Newman & Barkema.)

For example, if we plot the energy of our samples as a function of "Markov time", we might find that:



We don't want to use $\mu_1, \mu_2, \ldots, \mu_{t_{eq}}$ when we calculate our estimators, so:

$$Q_M = \frac{1}{M - t_{eq}} \sum_{m=t_{eq}+1}^{M} Q_{\mu_m}$$

Measurement

When calculating Q_M, we should ideally use sampled states μ_m that are statistically independent.

The Markov time required to get independent samples depends on both the algorithm and the temperature.

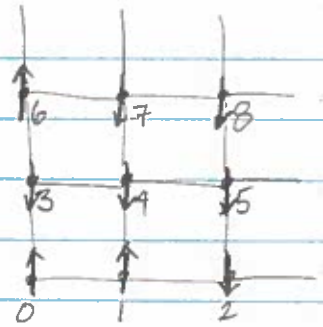
For single-spin-flip algorithms, μ_m and μ_{m+1} will be highly correlated.

As a result, we usually do a "sweep" of $O(N)$ proposed updates before performing a measurement so that μ_m and μ_{m+1} will actually be separated by many single-spin flips.

(For explanation of how to determine the "autocorrelation time" required to generate independent samples, see Section 3.3.1 of Newman & Barkema.)

Code for today's tutorial: ising2d_mc.py

Ising model on an $L \times L$ lattice with periodic boundary conditions and $h = 0$



$L = 3$

$N = 3 \times 3$

$$H = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j$$

We use an array called "spins" to store the present state:

spins = [+1, +1, -1, -1, -1, -1, +1, -1, -1]

In order to calculate the energy of a given configuration, we need to know the nearest neighbours of each lattice site i. We store this information in a two-dimensional array called "neighbours".

For example, when $L = 3$:

neighbours[3] = [4, 6, 5, 0]

right, up, left, down

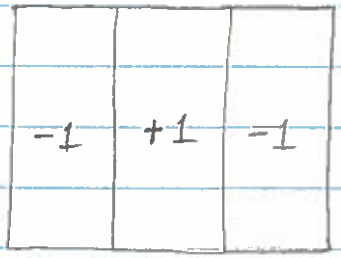
This code implements the single-spin-flip Metropolis algorithm to sample states μ at different temperatures.

The single-spin-flip algorithm has some problems:

(1) It takes a long time to generate independent samples (especially near T_c)

(2) It can get stuck in local energy minima at low temperatures, where it is very unlikely to accept moves that raise the energy.

For example:



There is no possible single-spin flip that lowers the energy of this configuration, but it is not the ground state.

Cluster Algorithms

Idea: Probabilistically build clusters of either all up or all down spins and propose flipping the entire cluster.

↳ Swendsen and Wang (1987)
↳ Wolff (1989)

Specific heat and susceptibility

Recall that, in the vicinity of a continuous phase transition:

$C_v/N \sim |t|^{-\alpha}$ specific heat per spin

$\chi/N \sim |t|^{-\gamma}$ susceptibility per spin

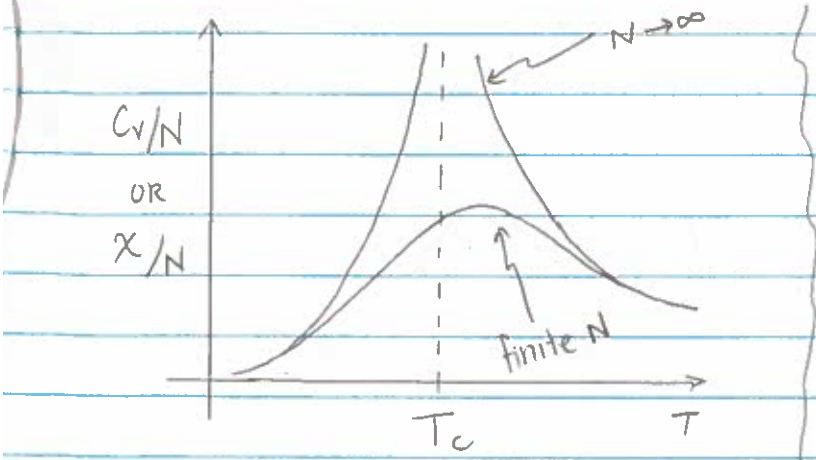
$\xi \sim |t|^{-\nu}$ correlation length

where: • α, γ and ν are called critical exponents.

• $t = \dfrac{T - T_c}{T_c}$

For the 2D Ising model C_v/N and χ/N both diverge at T_c.

However, on a finite-size lattice, C_v/N and χ/N will peak near T_c but will not diverge



One can show that:

$$C_v = \frac{\partial \langle E \rangle}{\partial T} = \frac{\langle E^2 \rangle - \langle E \rangle^2}{T^2}$$

$$\chi = \left.\frac{\partial \langle M \rangle}{\partial h}\right|_{h=0}$$

$$= \frac{\langle M^2 \rangle - \langle M \rangle^2}{T}$$

In today's tutorial, we will plot C_V/N and χ/N vs. T and get a rough estimate for T_c from the location of the peak. This estimate gets better as N increases.

Finite-size scaling

↳ technique for calculating T_c and critical exponents from calculations on finite lattices

Note that since $\frac{\chi}{N} \sim |t|^{-\gamma}$ and $\xi \sim |t|^{-\nu}$:

$$\frac{\chi}{N} \sim \xi^{\gamma/\nu}$$

On a finite system, correlations cannot exceed $O(L)$ and so:

$$\frac{\chi}{N} \sim \begin{cases} \xi^{\gamma/\nu} & \text{when } \xi \ll L \\ L^{\gamma/\nu} & \text{when } \xi \gg L \end{cases}$$

We can express this behaviour as:

$$\frac{\chi}{N} = \xi^{\gamma/\nu} f\left(\frac{L}{\xi}\right)$$

where $f\left(\frac{L}{\xi}\right) \sim \begin{cases} \text{constant} & \text{when } \xi \ll L \\ \left(\frac{L}{\xi}\right)^{\gamma/\nu} & \text{when } \xi \gg L \end{cases}$

Let's eliminate ξ from this expression:

$$\frac{\chi}{N} = L^{\gamma/\nu} \left(\frac{L}{\xi}\right)^{-\gamma/\nu} f\left(\frac{L}{\xi}\right)$$

$$= L^{\gamma/\nu} \left(L|t|^{\nu}\right)^{-\gamma/\nu} f\left(L|t|^{\nu}\right)$$

$$\equiv L^{\gamma/\nu} \tilde{f}\left(L^{1/\nu}|t|\right)$$

where $\tilde{f}(x) \equiv x^{-\gamma} f(x^{\nu})$.

$$\Rightarrow \frac{\chi}{N} L^{-\gamma/\nu} = \tilde{f}\left(L^{1/\nu}|t|\right)$$

We see that plots of $\frac{\chi}{N} L^{-\gamma/\nu}$ vs. $L^{1/\nu}|t|$ for various system sizes should all collapse onto the same curve in the vicinity of T_c.

From this collapse, one can estimate γ, ν and T_c.

Similarly, one can extract other critical exponents:

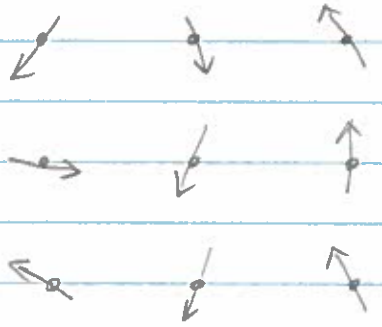
$$\frac{C_v}{N} L^{-\alpha/\nu} = \tilde{f}_c\left(L^{1/\nu}|t|\right)$$

$$\frac{M}{N} L^{\beta/\nu} = \tilde{f}_M\left(L^{1/\nu}|t|\right) \qquad \left(\frac{M}{N} \sim |t|^{\beta}\right)$$

XY Model

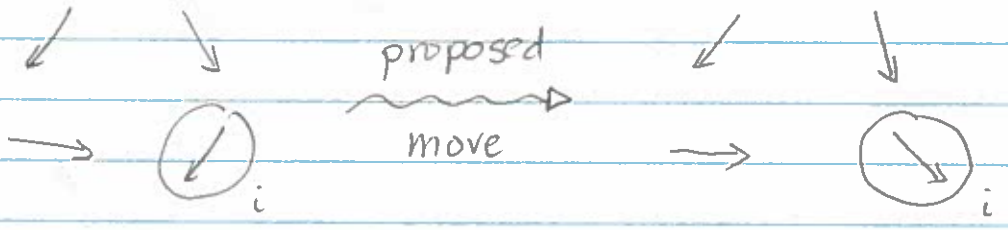
$$H = -J \sum_{\langle ij \rangle} \vec{S}_i \cdot \vec{S}_j$$

$$= -J \sum_{\langle ij \rangle} \cos(\theta_i - \theta_j), \quad \theta \in [0, 2\pi)$$



In 2D, there is no phase transition to a long-range ordered phase (by Mermin-Wagner theorem). Instead, there is a topological Kosterlitz-Thouless transition.

Single-spin updates are very similar to the Ising model:



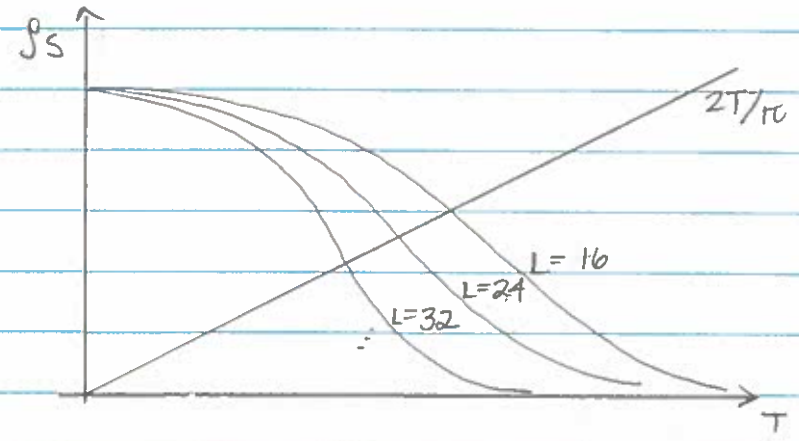
Propose the move $\theta_i \to \theta_i + \Delta\theta$ where $\Delta\theta \in [0, 2\pi)$ is selected from a uniform distribution

At the transition, it is known that

$$\underbrace{\rho_s(T_{KT})}_{\text{spin stiffness}} = \frac{2T_{KT}}{\pi}$$

(Nelson and Kosterlitz, 1977)

Plots of $\rho_s(T)$ vs. T should intersect with the line $\frac{2T}{\pi}$ at $T = T_{KT}$ when $L \to \infty$



Study the crossing points as a function of L and use finite-size scaling to extrapolate to $L \to \infty$.
(See arXiv:1302.2900)