

Selected Topics in Computational Quantum Physics

量子物理计算方法选讲

Shuo Yang (杨硕)

Department of Physics, Tsinghua University

Email: shuoyang@tsinghua.edu.cn

WeChat: condmat-ys

Quantum Monte Carlo

- classical Monte Carlo
 - importance sampling, Markov chain, detailed balance, Metropolis method
- quantum Monte Carlo for spin and bosonic systems
- quantum Monte Carlo for fermionic systems

Monte Carlo method

- a broad class of computational algorithms that rely on **repeated random sampling** to obtain numerical results
- **essential idea:** use randomness to solve problems that might be deterministic in principle
- the modern version of the Markov Chain Monte Carlo method was invented in the late 1940s by Stanislaw Ulam, while he was working on **nuclear weapons projects** at the Los Alamos National Laboratory
Monte Carlo methods were central to the simulations required for the **Manhattan Project**
- name from the Monte Carlo Casino in Monaco



Example

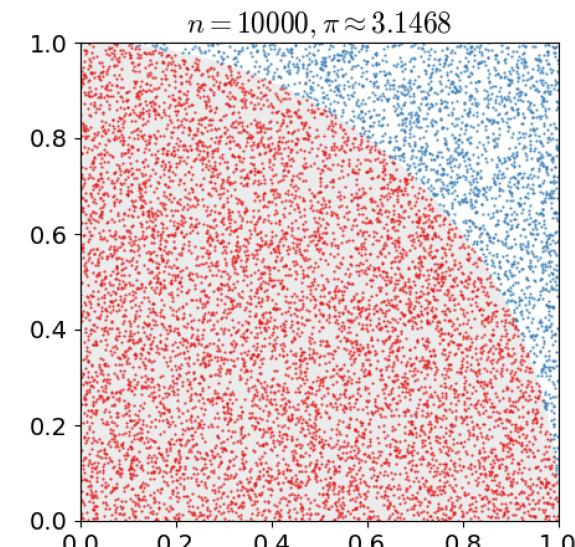
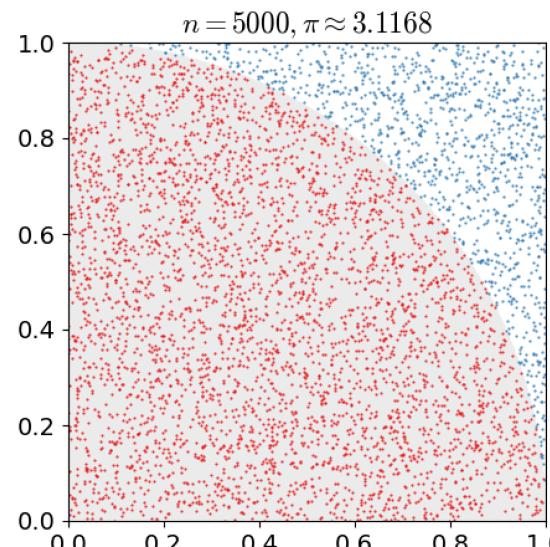
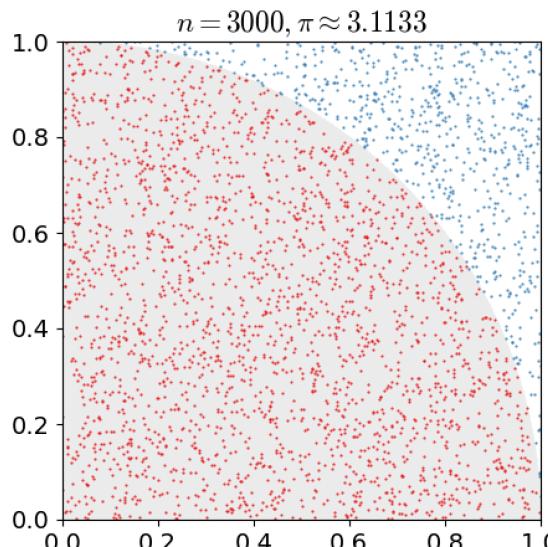
- estimate the value of π using a Monte Carlo method

draw a square, then inscribe a quadrant within it

uniformly scatter a given number of points over the square

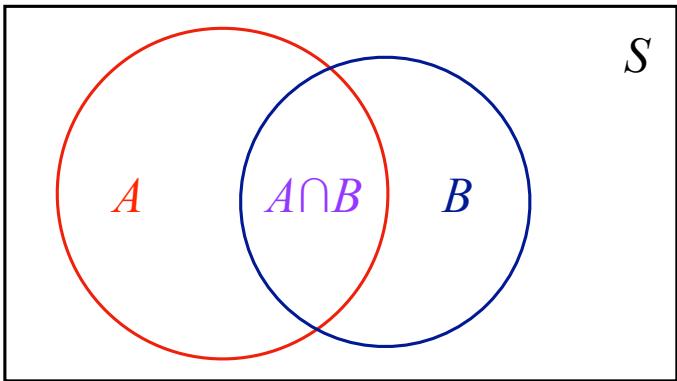
count the number of points inside the quadrant

the ratio of the inside-count and the total-sample-count is an estimate of the ratio of the two areas, $\pi/4$, multiply the result by 4 to estimate π



Probability

- consider a set S



$$0 \leq P(X) \leq 1$$

$$P(\emptyset) = 0, P(S) = 1$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- conditional probability of A given B , assuming $P(B) > 0$

$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$

- independence

if A is independent of B , $P(A | B) = P(A)$

equivalently, $P(A \cap B) = P(A) P(B)$

- Bayes' theorem, gives the relationship between $P(A | B)$ and $P(B | A)$

$$P(A | B) = P(B | A) \frac{P(A)}{P(B)}$$

because $P(B) P(A | B) = P(A) P(B | A) = P(A \cap B)$

Random variable

- a random variable has a whole set of values and it could take on any of those values, randomly
- a random variable follows a probability density function $P(x)$, which specifies the probability of its values
- discrete variable, e.g., $X = \{x_1, x_2, x_3\}$ with probability $P(X) = \left\{ \frac{2}{8}, \frac{5}{8}, \frac{1}{8} \right\}$
the cumulate distribution function is defined as

$$F_X(x) = P(X \leq x) \quad P(a < X \leq b) = F_X(b) - F_X(a)$$

$$\text{e.g., } F_X(x_1) = \frac{2}{8}, F_X(x_2) = \frac{2}{8} + \frac{5}{8} = \frac{7}{8}, F_X(x_3) = \frac{2}{8} + \frac{5}{8} + \frac{1}{8} = 1$$

- continuous variable, $P(x)dx$ gives probability that x falls between x and $x+dx$
probability that $a \leq x \leq b$ is given by $\int_a^b P(x) dx$
the cumulate distribution function is defined as

$$F(x) = \int_{-\infty}^x P(t) dt$$

Statistic of a random variable

- discrete variable X with probability $x_1 \mapsto p_1, x_2 \mapsto p_2, \dots, x_n \mapsto p_n$

average

$$\mu = \sum_{i=1}^n p_i x_i = \langle x \rangle$$

variable

$$\text{Var}(x) = \sum_{i=1}^n p_i (x_i - \mu)^2 = \sum_{i=1}^n p_i x_i^2 - \mu^2 = \langle x^2 \rangle - \langle x \rangle^2$$

- continuous variable

average

$$\mu = \int x P(x) dx = \int x dF(x)$$

variable

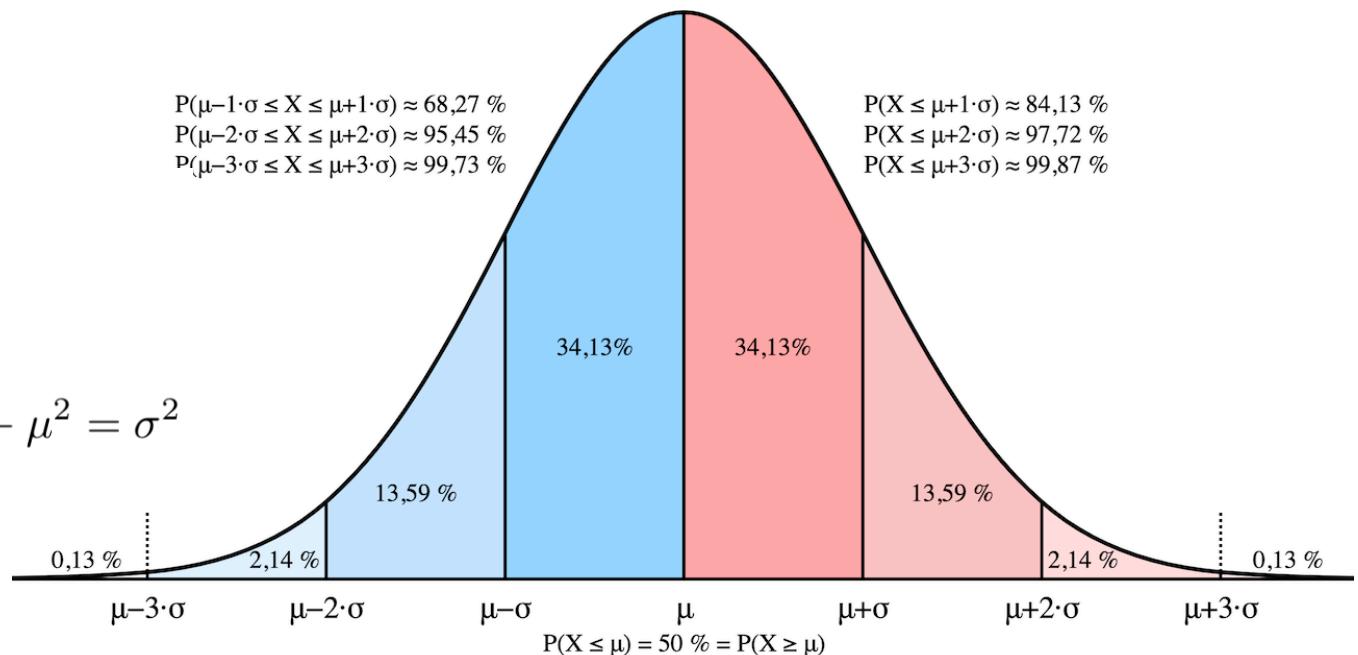
$$\text{Var}(x) = \sigma^2 = \int (x - \mu)^2 P(x) dx = \int x^2 dF(x) - \mu^2 = \langle x^2 \rangle - \langle x \rangle^2$$

- normal distribution

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$E(x) = \mu$$

$$\text{Var}(x) = \int_{-\infty}^{\infty} x^2 P(x) dx - \mu^2 = \sigma^2$$

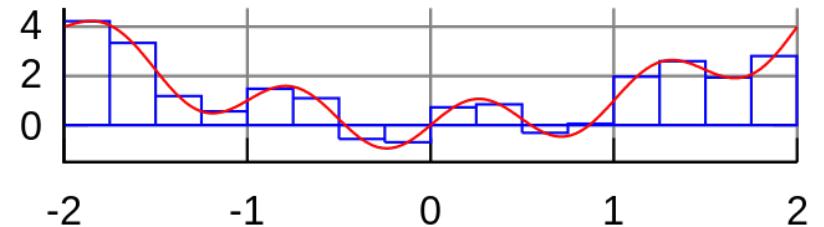


Numerical Integration

- basic problem: compute an approximate solution to a definite integral
- rectangle approach

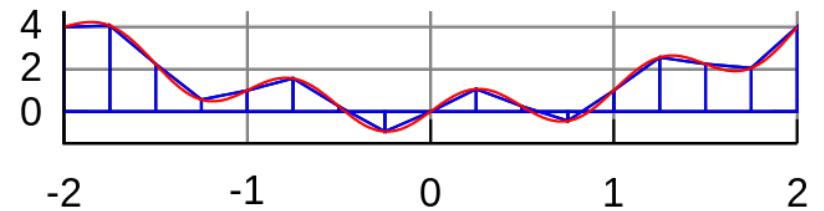
$$\int_a^b f(x) dx = \lim_{\Delta x \rightarrow 0} \sum_{i=1}^N f[(x_{i-1} + x_i)/2] \Delta x$$

$$x_i = a + i\Delta x \quad \Delta x = (b - a)/N$$



- trapezoidal approach

$$\int_a^b f(x) dx = \lim_{\Delta x \rightarrow 0} \sum_{i=1}^N [f(x_{i-1}) + f(x_i)] \Delta x / 2$$



- naive Monte Carlo approach

we sample the points not on regular grids, but randomly (uniformly distributed)

$$\int_a^b f(x) dx \approx (b - a) \frac{1}{N} \sum_{i=1}^N f(x_i) = (b - a) \langle f \rangle$$

Monte Carlo error decreases with sample size N as

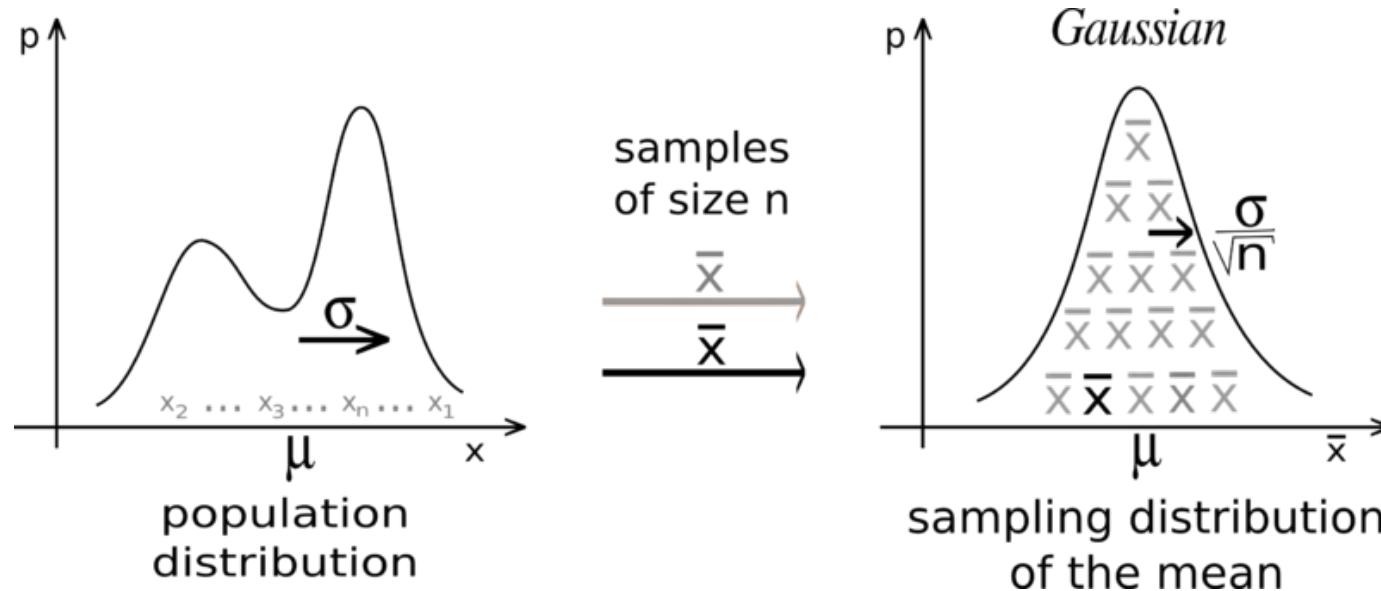
$$\varepsilon \propto \frac{1}{\sqrt{N}}$$

why?

Central limit theorem (CLT)

- what is the central limit theorem?

CLT states that **the distribution of sample means** approaches a **normal distribution** as the sample size gets larger, regardless of the population distribution shape



- sample sizes ≥ 30 are considered sufficient for the CLT to hold
- a sufficiently large sample can **predict** the parameters of a population such as the mean and standard deviation

Importance sampling

- the Monte Carlo error σ/\sqrt{n} can be reduced by reducing the effective variance σ^2 using a method called importance sampling
- concentrate the sampling on regions where $f(x)$ is large, using a statistical weight $\omega(x) > 0$ with

$$\int_a^b \omega(x) dx = 1$$

the sampling then becomes

$$\langle f \rangle = \int \frac{f(x)}{\omega(x)} \omega(x) dx \simeq \frac{1}{N} \sum_{i=1}^N \frac{f(x_i)}{\omega(x_i)}$$

so that the error estimate changes according to

$$\text{Var} \langle f \rangle \rightarrow \text{Var} \left\langle \frac{f}{\omega} \right\rangle = \frac{1}{N} \text{Var} \frac{f}{\omega}$$

in order to keep the variance small, the distribution ω should be similar to f

- some random variables have more impact than others
if these important values are emphasized by sampling more frequently,
then the estimator variance can be reduced

Importance sampling

- importance sampling enters the simulation of a physical system naturally through the partition function

$$Z = \sum_n e^{-\beta E_n}$$

the canonical average is

$$\langle A \rangle = \frac{\sum_n e^{-\beta E_n} A_n}{Z}$$

- we choose a subset of microstates $\{x\}_i$ with probability distribution

$$\omega(\{x\}_i) = e^{-\beta E(\{x\}_i)}$$

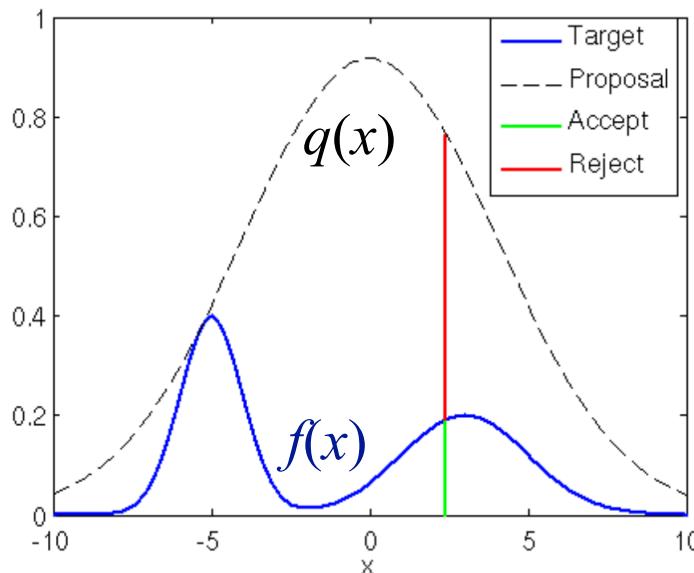
then the average becomes

$$\langle A \rangle = \frac{\frac{1}{N} \sum_{i=1}^N \frac{e^{-\beta E(\{x\}_i)}}{\omega(\{x\}_i)} A(\{x\}_i)}{\frac{1}{N} \sum_{i=1}^N \frac{e^{-\beta E(\{x\}_i)}}{\omega(\{x\}_i)}} \simeq \frac{1}{N} \sum_{i=1}^N A(\{x\}_i)$$

- the next question is how to generate random variable $\{x\}_i$ with the desired probability distribution $\omega(\{x\}_i)$?

Acceptance-rejection sampling

- suppose that we want to sample from a distribution $f(x)$ that is difficult or impossible to sample from directly, but instead have a simpler distribution $q(x)$ from which sampling is easy



- we sample from $q(x)$ and apply some rejection/acceptance criterion such that the samples that are accepted are distributed according to $f(x)$
- in high dimensions, suitable $f(x)$ is hard to find
a lot of rejections can take place before a useful sample is generated, thus making the algorithm inefficient and impractical

Markov chain

- Markov process

a stochastic process is a sequence of random variables $x_0, x_1, \dots, x_n, \dots$,
the probability of moving to the next state depends only on the present
state and not on the previous states

$$P(x_{n+1} | x_0, x_1, \dots, x_n) = P(x_{n+1} | x_n)$$

- Markov chain

we use a Markov process repeatedly to generate a Markov chain of states

Markov chain is completely characterized by an initial probability distribution $P(x_0)$ and the transition probability $W(x_n \rightarrow x_{n+1})$

the probability that x_0, x_1, \dots, x_n appears is

$$P(x_0) W(x_0 \rightarrow x_1) W(x_1 \rightarrow x_2) \dots W(x_{n-1} \rightarrow x_n)$$

Equilibrium of Markov process

- the Markov process is chosen specially so that starting from any state when it is run for long enough it will eventually produce a succession of states with some probability distribution (coming to equilibrium)
- the rate at which the system makes transitions into and out of any state x must be equal (equilibrium condition)

$$\sum_y P(x) W(x \rightarrow y) = \sum_y P(y) W(y \rightarrow x)$$

- conditions to satisfy
 1. the transition probability satisfy the constraint $\sum_y W(x \rightarrow y) = 1$ since the Markov process must generate some state y when in the state x the equilibrium condition is simplify to

$$P(x) = \sum_y P(y) W(y \rightarrow x)$$

$P(x)$ is invariant with respect to $W(y \rightarrow x)$

Equilibrium of Markov process

2. ergodicity

it should be possible to reach any state from any other state, if we run it for long enough

3. detailed balance

$$P(x)W(x \rightarrow y) = P(y)W(y \rightarrow x)$$

on average the system should go from x to y just as often as it goes from y to x

- any set of transition probabilities which satisfy this condition also satisfy the equilibrium condition

to prove it we sum over y

$$\sum_y P(x)W(x \rightarrow y) = \sum_y P(y)W(y \rightarrow x)$$

- Monte Carlo sampling do not deal with probability $P(x)$ directly
it generates next sample y from the current x using the transition probability $W(x \rightarrow y)$
our task is to find suitable W

Metropolis (-Hasting) algorithm

- the algorithm was named after Nicholas Metropolis

THE JOURNAL OF CHEMICAL PHYSICS

VOLUME 21, NUMBER 6

JUNE, 1953

Equation of State Calculations by Fast Computing Machines

NICHOLAS METROPOLIS, ARIANNA W. ROSENBLUTH, MARSHALL N. ROSENBLUTH, AND AUGUSTA H. TELLER,
Los Alamos Scientific Laboratory, Los Alamos, New Mexico

AND

EDWARD TELLER,* *Department of Physics, University of Chicago, Chicago, Illinois*

(Received March 6, 1953)

A general method, suitable for fast computing machines, for investigating such properties as equations of state for substances consisting of interacting individual molecules is described. The method consists of a modified Monte Carlo integration over configuration space. Results for the two-dimensional rigid-sphere system have been obtained on the Los Alamos MANIAC and are presented here. These results are compared to the free volume equation of state and to a four-term virial coefficient expansion.

- this paper proposed the algorithm for the case of symmetrical proposal distributions, and W. K. Hastings extended it to the more general case in 1970
- Metropolis algorithm for Monte Carlo is among the "Top Ten Algorithms of the 20th Century"



- Metropolis algorithm for Monte Carlo
- simplex method for linear programming
- Krylov subspace iteration methods
- the decompositional approach to matrix computations
- the Fortran optimizing compiler
- QR algorithm for computing eigenvalues
- quicksort algorithm for sorting
- fast Fourier transform
- integer relation detection
- fast multipole method

Two parts of transition probability

- we break the **transition probability** down into two parts

$$W(x \rightarrow y) = Q(x \rightarrow y) A(x \rightarrow y)$$

- $Q(x \rightarrow y)$ is the **selection probability** $Q(x \rightarrow y) = P(y | x)$
it is the probability, given an initial state x , that our algorithm will generate a new target state y
- $A(x \rightarrow y)$ is the **acceptance probability**
if we start from x and our algorithm generates a new state y from it, we can accept or reject state y with acceptance ratio $A(x \rightarrow y) \in (0,1)$
- in **Metropolis** algorithm, $Q(x \rightarrow y) = Q(y \rightarrow x)$, Q is symmetric such as Gaussian distribution centered at x
in **Metropolis-Hastings** algorithm, $Q(x \rightarrow y) \neq Q(y \rightarrow x)$

Metropolis algorithm

- for each iteration t
 1. generate: randomly generate a candidate state y according to $Q(x \rightarrow y)$
 2. calculate: calculate the **acceptance probability**

$$A(x \rightarrow y) = \min \left\{ 1, \frac{P(y)}{P(x)} \frac{Q(y \rightarrow x)}{Q(x \rightarrow y)} \right\} = \min \left\{ 1, \frac{P(y)}{P(x)} \right\}$$

- 3. accept or reject:
 - generate a uniform random number $u \in (0,1)$
 - if $u \leq A(x \rightarrow y)$, **accept** the new state and set $x_{t+1} = y$
 - if $u > A(x \rightarrow y)$, **reject** the new state and copy the old state forward $x_{t+1} = x$
- 4. increment: set $t = t + 1$

- the choice of $A(x \rightarrow y)$ ensures the **detailed balance condition is satisfied**

$$\begin{aligned} P(x) W(x \rightarrow y) &= P(x) Q(x \rightarrow y) A(x \rightarrow y) = P(x) Q(y \rightarrow x) \min \left\{ 1, \frac{P(y)}{P(x)} \right\} \\ &= P(y) Q(y \rightarrow x) \min \left\{ \frac{P(x)}{P(y)}, 1 \right\} = P(y) Q(y \rightarrow x) A(y \rightarrow x) \\ &= P(y) W(y \rightarrow x) \end{aligned}$$

Metropolis-Hastings algorithm

- in the general case, $Q(x \rightarrow y) \neq Q(y \rightarrow x)$
we use a more general acceptance probability

$$A(x \rightarrow y) = \min \left\{ 1, \frac{P(y)}{P(x)} \frac{Q(y \rightarrow x)}{Q(x \rightarrow y)} \right\}$$

- the choice of $A(x \rightarrow y)$ ensures the **detailed balance condition** is satisfied

$$\begin{aligned} & P(x) W(x \rightarrow y) \\ &= P(x) Q(x \rightarrow y) A(x \rightarrow y) = P(x) Q(x \rightarrow y) \min \left\{ 1, \frac{P(y)}{P(x)} \frac{Q(y \rightarrow x)}{Q(x \rightarrow y)} \right\} \\ &= \min \{P(x) Q(x \rightarrow y), P(y) Q(y \rightarrow x)\} = P(y) Q(y \rightarrow x) \min \left\{ \frac{P(x)}{P(y)} \frac{Q(x \rightarrow y)}{Q(y \rightarrow x)}, 1 \right\} \\ &= P(y) Q(y \rightarrow x) A(y \rightarrow x) = P(y) W(y \rightarrow x) \end{aligned}$$

- if we construct the Markov chain according to the acceptance probability $A(x \rightarrow y)$, we will eventually produce a succession of states with the desired probability distribution $P(x)$

An alternative derivation of the acceptance probability

- start from the detailed balance condition

$$\begin{aligned} P(x)W(x \rightarrow y) &= P(y)W(y \rightarrow x) \\ \Rightarrow P(x)Q(x \rightarrow y)A(x \rightarrow y) &= P(y)Q(y \rightarrow x)A(y \rightarrow x) \\ \Rightarrow \frac{A(x \rightarrow y)}{A(y \rightarrow x)} &= \frac{P(y)Q(y \rightarrow x)}{P(x)Q(x \rightarrow y)} \end{aligned}$$

- we make the larger one of $A(x \rightarrow y)$ and $A(y \rightarrow x)$ to be 1
then we choose the acceptance probability as

$$A(x \rightarrow y) = \min \left\{ 1, \frac{A(x \rightarrow y)}{A(y \rightarrow x)} \right\} = \min \left\{ 1, \frac{P(y)}{P(x)} \frac{Q(y \rightarrow x)}{Q(x \rightarrow y)} \right\}$$

- it can be verified that

$$A(x \rightarrow y) = \begin{cases} 1 & \text{when } 1 = A(x \rightarrow y) > A(y \rightarrow x) \\ A(x \rightarrow y) & \text{when } 1 = A(y \rightarrow x) > A(x \rightarrow y) \end{cases}$$

Classical Ising model

- the Hamiltonian takes the form

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j + H \sum_i \sigma_i$$

where H is an external field

- the partition function is $Z = \sum_{\{\sigma_i\}} e^{-\beta E(\{\sigma_i\})}$

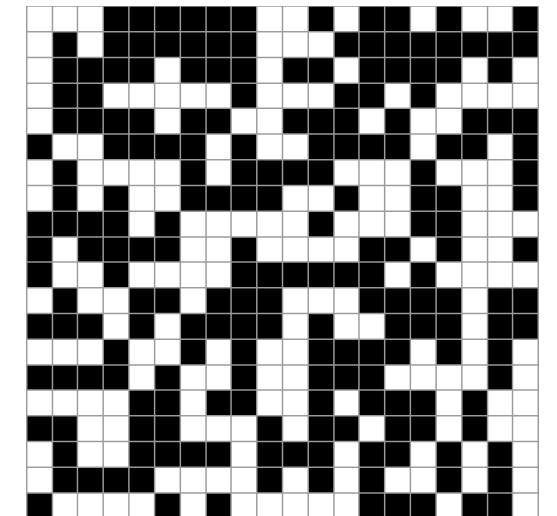
the average energy is $\langle E \rangle = \frac{1}{Z} \sum_{\{\sigma_i\}} E(\{\sigma_i\}) e^{-\beta E(\{\sigma_i\})}$

- the specific heat is calculated from the energy fluctuations

$$C = \frac{1}{kT^2} \left(\langle E^2 \rangle - \langle E \rangle^2 \right)$$

this is because $C = \frac{\partial \langle E \rangle}{\partial T} = \frac{\partial \langle E \rangle}{\partial \beta} \frac{\partial \beta}{\partial T} = -\frac{1}{kT^2} \frac{\partial \langle E \rangle}{\partial \beta}$

$$\begin{aligned} \frac{\partial \langle E \rangle}{\partial \beta} &= \frac{1}{Z^2} \left[\begin{aligned} &\left(- \sum_{\{\sigma_i\}} [E(\{\sigma_i\})]^2 e^{-\beta E(\{\sigma_i\})} \right) Z \\ &- \left(\sum_{\{\sigma_i\}} E(\{\sigma_i\}) e^{-\beta E(\{\sigma_i\})} \right) \left(- \sum_{\{\sigma_i\}} E(\{\sigma_i\}) e^{-\beta E(\{\sigma_i\})} \right) \end{aligned} \right] \\ &= \langle E \rangle^2 - \langle E^2 \rangle \end{aligned}$$



Classical Ising model

- similarly, the average magnetization is

$$\langle M \rangle = \frac{1}{Z} \sum_{\{\sigma_i\}} m(\{\sigma_i\}) e^{-\beta E(\{\sigma_i\})}$$

- the magnetic susceptibility is

$$\begin{aligned}\chi &= \frac{\partial \langle M \rangle}{\partial H} = \frac{1}{Z^2} \left[- \left(\sum_{\{\sigma_i\}} m(\{\sigma_i\}) e^{-\beta E(\{\sigma_i\})} \right) \left(-\beta \sum_{\{\sigma_i\}} \frac{\partial E(\{\sigma_i\})}{\partial H} e^{-\beta E(\{\sigma_i\})} \right) \right] \\ &= \beta \frac{1}{Z^2} \left[\left(\sum_{\{\sigma_i\}} [m(\{\sigma_i\})]^2 e^{-\beta E(\{\sigma_i\})} \right) Z - \left(\sum_{\{\sigma_i\}} m(\{\sigma_i\}) e^{-\beta E(\{\sigma_i\})} \right)^2 \right] \\ &= \frac{1}{kT} \left(\langle M^2 \rangle - \langle M \rangle^2 \right)\end{aligned}$$

- the classical Ising model undergoes a phase transition between an ordered and a disordered phase in two dimensions or more
- in 2D, $T_c = \frac{2J}{\ln(1 + \sqrt{2})} \simeq 2.269J$

Monte Carlo simulation

- (1) **initialization**: choose an initial configuration for the spins
 - (2) **trial move**: we propose a move by randomly choosing a spin and flip it
 - (3) **acceptation**:

if we have flipped the j^{th} spin $\{\sigma_{\text{old}}\} = \{\sigma_1, \sigma_2, \dots, \sigma_j, \dots, \sigma_N\}$
 $\{\sigma_{\text{trial}}\} = \{\sigma_1, \sigma_2, \dots, -\sigma_j, \dots, \sigma_N\}$

we evaluate the number

$$\alpha = \frac{P(\{\sigma_{\text{trial}}\})}{P(\{\sigma_{\text{old}}\})} = \exp[-\beta(E(\{\sigma_{\text{trial}}\}) - E(\{\sigma_{\text{old}}\}))] = \exp[-\beta\Delta E]$$

$$\Delta E = 2\sigma_j \left(J \sum \sigma_{\text{near}} - H \right)$$

then we generate a uniform random number $r \in (0,1)$

if $r < \alpha$, accept the move, defining $\{\sigma_{\text{new}}\} = \{\sigma_{\text{trial}}\}$

if $r > \alpha$, reject the move, defining $\{\sigma_{\text{new}}\} = \{\sigma_{\text{old}}\}$

- (4) **iteration**: then we use σ_{new} as the new starting point and go back to (2)

we proceed with the process for at least 10^5 - 10^6 Monte Carlo steps

Commonly adopted empirical strategies

- equilibration

we discard the first steps of the random walk, when the distribution of the sampled Markov chain has not yet reached its limit $P(x)$

- sparse averaging

to reduce statistical correlations, we wait for a given number of steps between two measurements

the waiting time can be adjusted empirically or evaluated computing an autocorrelation time

$$A(j) = \frac{\langle \mathcal{O}_i \mathcal{O}_{i+j} \rangle - \langle \mathcal{O}_i \rangle \langle \mathcal{O}_i \rangle}{\langle \mathcal{O}_i^2 \rangle - \langle \mathcal{O}_i \rangle \langle \mathcal{O}_i \rangle}$$

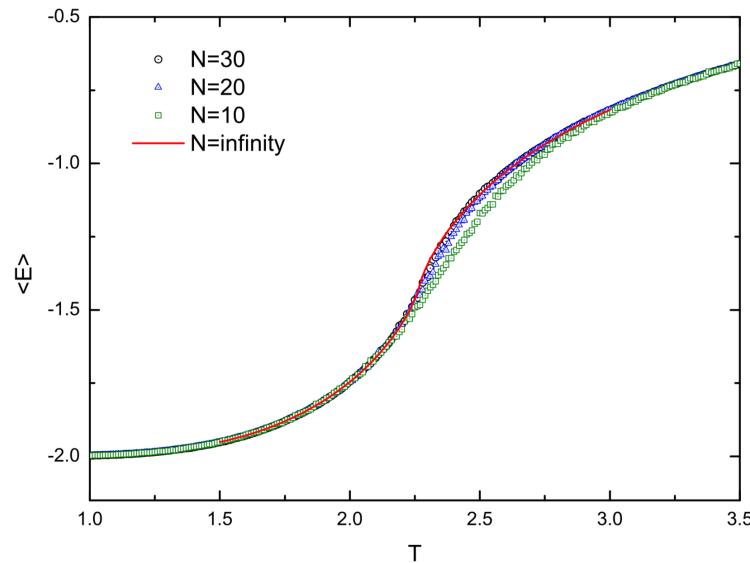
- data blocking: perform many different independent simulations

we divide the simulation in large blocks, for each block we provide an estimation

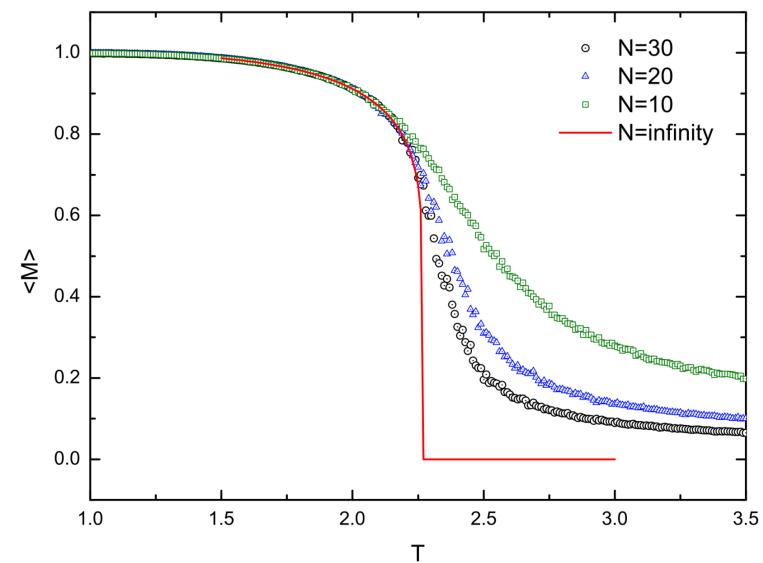
if the blocks are large enough, we may rely on the central limit theorem and interpret each estimation as a realization of a normal random variable

Results

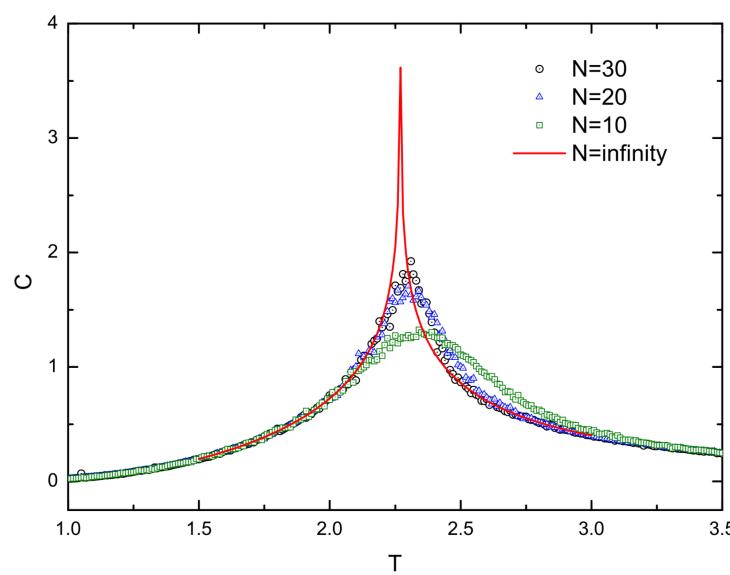
- average energy



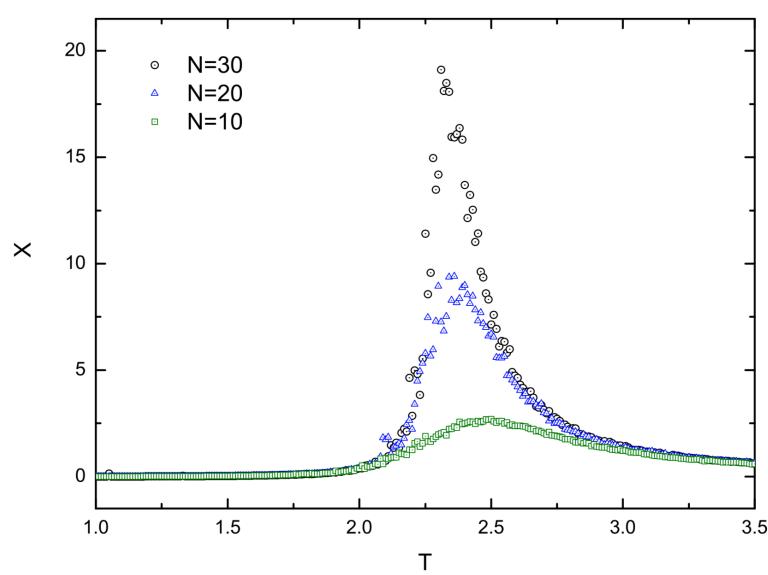
- average magnetization



- specific heat



- magnetic susceptibility



Critical exponent

- we want to describe the behavior of a physical quantity f in terms of a power law around the critical temperature

we introduce the reduced temperature $t = \frac{T - T_c}{T_c}$

- the **critical exponents** of the classical Ising model are

$$M \propto (-t)^\beta, t \rightarrow -0$$

$$\chi \propto t^{-\gamma}, t \rightarrow +0$$

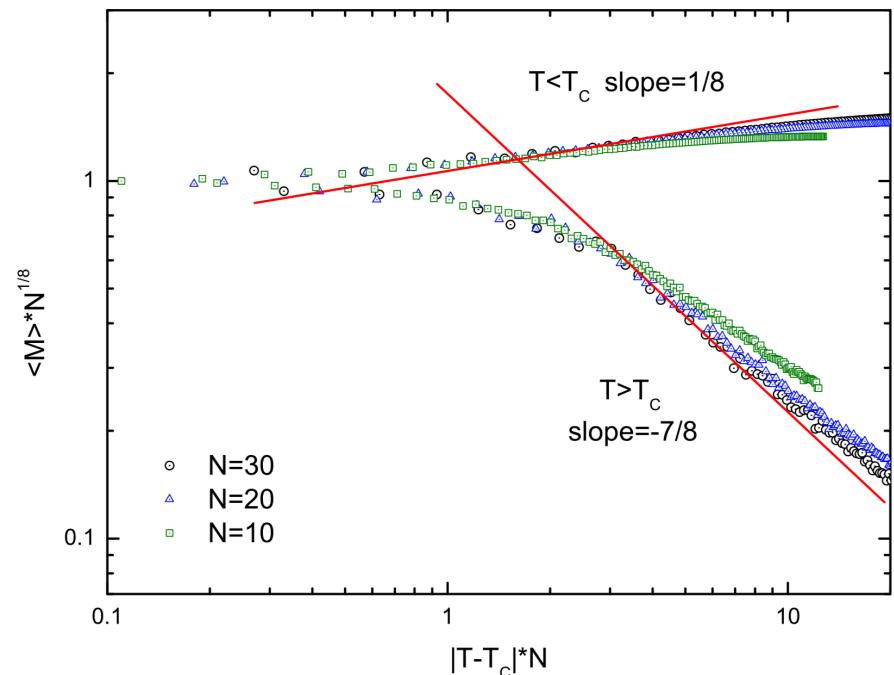
$$\chi \propto (-t)^{-\gamma'}, t \rightarrow -0$$

$$M \propto H^{1/\delta}$$

- in 2D, exact solution gives $\beta = 1/8$

- numerical results agree well with the exact solution

- demo: <http://physics.weber.edu/schroeder/software/demos/IsingModel.html>



Homework rules

- this homework has only one task
please submit through <http://learn.tsinghua.edu.cn>
- please submit the source code and a detailed note including:
 - (1) a brief summary of theory and algorithm
 - (2) the structure and technical details of the code
 - (3) problems encountered and solutions
 - (4) summary of results and your understanding
- deadline is Dec. 15, 23:00
if you submit it after the deadline, you will have fewer points
- if you are an expert in classical Monte Carlo, this homework may be waived
please contact me in person and submit something you have done before

Homework

- consider the 2D classical Ising model on a $L \times L$ square lattice with periodic boundary conditions in both directions

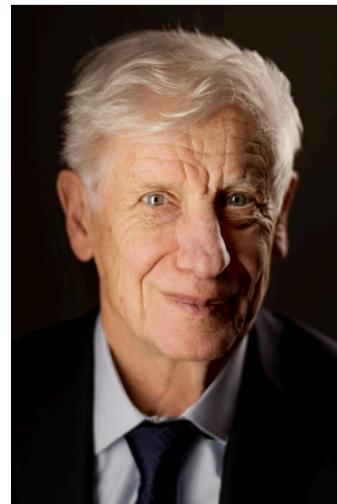
$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j + H \sum_i \sigma_i \quad \sigma_i = \pm 1, H \rightarrow 0$$

- write a code for the Metropolis algorithm
- plot the average energy, average magnetization, specific heat, and magnetic susceptibility as functions of temperature T
(see the figures on Page 26)
- explore the scaling behavior near the critical point $T_c = \frac{2J}{\ln(1 + \sqrt{2})} \simeq 2.269J$
compare with exact results
(see the figure on Page 27)

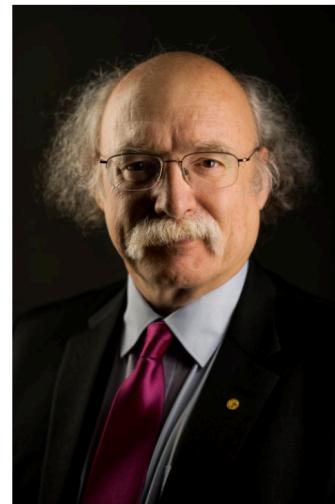
Grading of this homework

- academic integrity [+50]
completing assignments independently, creating and expressing your own ideas, DON'T copy answers from others or allow others to copy your answers
- the source code can be executed [+10] and can provide correct results [+10]
the source code has high readability [+10]
- there is a detailed note file to explain the basic principle, source code and results [+10]
the note file is well-written and easy to understand [+10]
- fail to plot figures [-10 for each]
- time-dependent score: after DDL: [-1 per day]
- self-motivation score: new ideas or extra results [+10]

Nobel prize in physics 2016



© Nobel Media AB. Photo: A.
Mahmoud
David J. Thouless
Prize share: 1/2



© Nobel Media AB. Photo: A.
Mahmoud
**F. Duncan M.
Haldane**
Prize share: 1/4



© Nobel Media AB. Photo: A.
Mahmoud
J. Michael Kosterlitz
Prize share: 1/4

- for theoretical discoveries of topological phase transitions and topological phases of matter

This year's Laureates opened the door on an unknown world where matter can assume strange states. They have used advanced mathematical methods to study unusual phases, or states, of matter, such as superconductors, superfluids or thin magnetic films. Thanks to their pioneering work, the hunt is now on for new and exotic phases of matter. Many people are hopeful of future applications in both materials science and electronics.

Berezinskii–Kosterlitz–Thouless transition

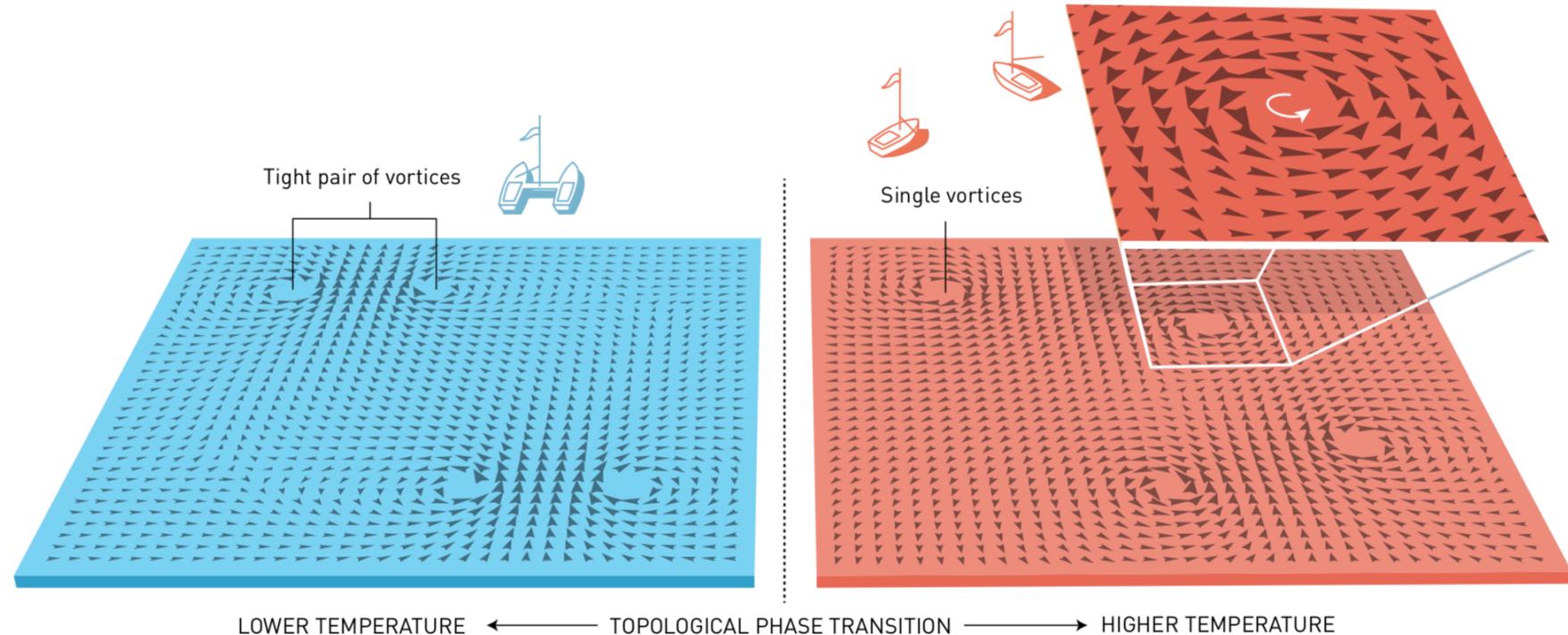


Fig. 2 Phase transition. This occurs when phases of matter transition between each other, such as when ice melts and becomes water. Using topology, Kosterlitz and Thouless described a topological phase transition in a thin layer of very cold matter. In the cold, vortex pairs form and then suddenly separate at the temperature of the phase transition. This was one of the twentieth century's most important discoveries in the physics of condensed matter.

- the leading role in a **topological transition** is played by small **vortices** in the flat material
- at low temperatures they form tight pairs
 - when the temperature rises, a phase transition takes place: the vortices suddenly move away from each other and sail off in the material on their own

Two dimensional XY model

- planar rotors of unit length arranged on a two dimensional square lattice

$$H = -J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j = -J \sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j)$$

summation over all nearest neighbor sites, θ_i denotes the angle of the rotor on site i with respect to x axis $\theta_i \in [0, 2\pi]$

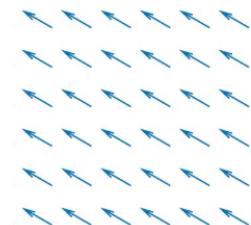
- H is invariant under global continuous rotation of all spins

has $U(1)$ gauge symmetry

the ordered state: all spins are aligned parallel



- the magnetization $S = \langle \vec{\mathbf{S}} \rangle$ can take any value between 1 and -1



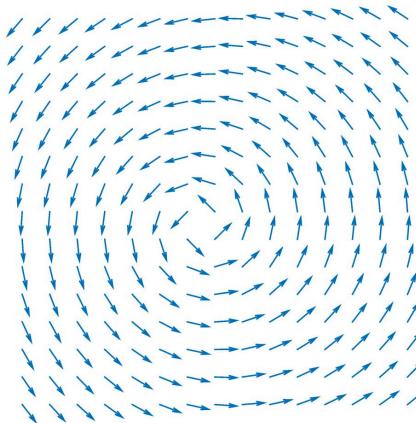
- different states can have the same energy

all of them related to each other by a global rotation of all spins

the number of ground states is infinite

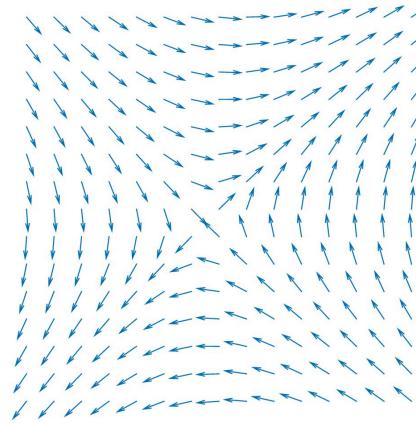
Vortex and anti-vortex

- vortex



$$n = +1$$

- anti-vortex

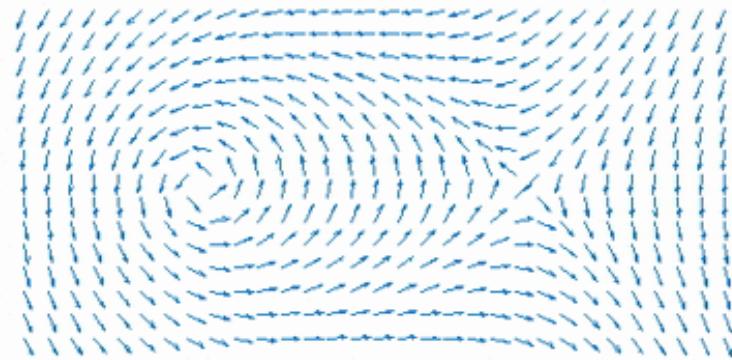
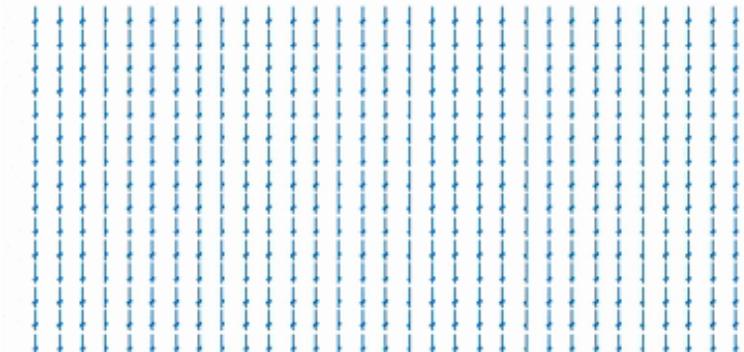


$$n = -1$$

- winding number

$$\oint \nabla \theta(\mathbf{r}) \cdot d\mathbf{l} = 2\pi n$$

- vortex and anti-vortex are created from the vacuum in pairs and they can destroy each other when brought together



Monte Carlo simulation of XY model

- the Metropolis algorithm is used to sample the configurations of the system in equilibrium with a heat bath at temperature T
- it can generate a sequence of configurations $\{\theta_i\}$ with probability distribution $\omega(\{\theta_i\}) = e^{-\beta E(\{\theta_i\})}$, where $E(\{\theta_i\}) = -J \sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j)$
- a demo can be downloaded from <https://www.compadre.org/OSP/document/ServeFile.cfm?ID=8642&DocID=1048#Doc1048>
- typical configurations showing the spins and vortices

