

Computational Statistical Physics

Part I: Statistical Physics and Phase Transitions

Lucas Böttcher

August 2, 2019

ETH Zürich

Institute for Theoretical Physics

HIT K 11.3

Wolfgang-Pauli-Strasse 27

8093 Zürich

ETH zürich

402-0812-00L

FS 2019

Dates and Further Information

- Lectures: Friday 10.45–12.30 in HIT H51
- Exercices: Friday 8.45–10.30 in HIT F21
- Oral exams take place in August 2019
- Webpage: <http://www.itp.phys.ethz.ch/education/computational-statphys.html>

Teaching Assistants

- Maximilian Holst: holstm@student.ethz.ch
- Giovanni Balduzzi: gbalduzz@itp.phys.ethz.ch

Course Offered in (Selection)

- Mathematics, Computer Science (Master)
- Physics (Master)
- Material Science (Spezialvorlesung, Master)
- Civil Engineering (Spezialvorlesung)
- Biomedical Engineering (Master)

Learning Objectives

Students are able to

- *distinguish between different computational methods to study statistical physics models.*
- *apply appropriate computational techniques to efficiently simulate a given equilibrium or non-equilibrium system.*
- *differentiate between different simulation methods to study particle and molecular interactions.*
- *apply appropriate molecular dynamics modeling concepts.*

Phase Transitions

- 22.02. Introduction to statistical physics and the Ising model
- 01.03. Monte Carlo methods
- 08.03. Finite size methods and cluster algorithms
- 15.03. Histogram methods and renormalization group
- 22.03. Renormalization group and Boltzmann machines
- 29.03. Non-equilibrium phase transitions and summary

Molecular Dynamics

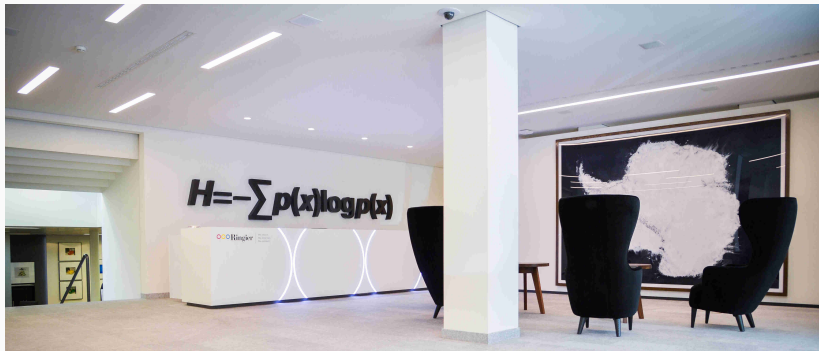
- 05.04. Molecular Dynamics, Verlet and Leapfrog methods
- 12.04. Optimization, Composed Particle Systems
- 19.04. and 26.04. ETH vacations
- 03.05. Arbitrary Shapes, Long-range Potentials
- 10.05. Nosé-Hoover thermostat, stochastic method, constant pressure ensemble

Event-driven Dynamics

- 17.05. Event driven, inelastic collisions, friction
- 24.05. Contact dynamics, Particles in Fluids, (if time allows: ab initio MD, Car-Parrinello)
- 31.05. shifted to 24.05

Classical Statistical Mechanics

Classical Statistical Mechanics



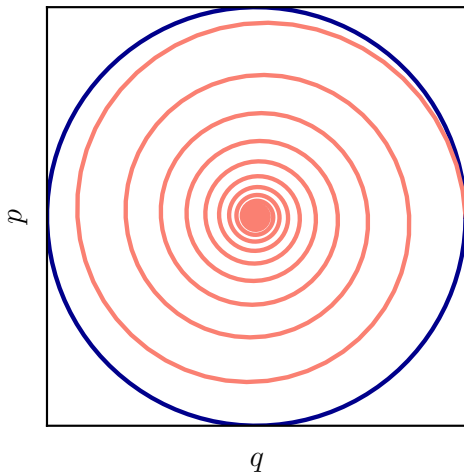
Phase Space

Let us consider a classical physical system with N particles whose canonical coordinates and the corresponding conjugate momenta are given by $q_1, \dots, q_{3N}$ and $p_1, \dots, p_{3N}$, respectively. The $6N$ -dimensional space Γ defined by the last set of coordinates defines the *phase space*.



Ludwig Boltzmann
(1844-1906)

Phase Space



Phase spaces of an undamped (blue) and a damped (orange) harmonic oscillator.

Ensemble Average

The assumption that all states in an ensemble are reached by the time evolution of the corresponding system is referred to as *ergodicity hypothesis*. We define the *ensemble average* of a quantity $Q(p, q)$ as

$$\langle Q \rangle = \frac{\int Q(p, q) \rho(p, q) \, dp dq}{\int \rho(p, q) \, dp dq}, \quad (1)$$

where $\rho(p, q)$ denotes the *phase space density* and $dp dq$ is a shorthand notation for $dp^{3N} dq^{3N}$.

Hamiltonian

The dynamics of the considered N particles is described by their *Hamiltonian* $\mathcal{H}(p, q)$, i.e., the equations of motions are

$$\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} \quad \text{and} \quad \dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i} \quad (i = 1, \dots, 3N). \quad (2)$$

Moreover, the temporal evolution of a phase space element of volume V and boundary ∂V is given by

$$\frac{\partial}{\partial t} \int_V \rho \, \mathrm{d}V + \int_{\partial V} \rho v \, \mathrm{d}A = 0, \quad (3)$$

where $v = (\dot{p}_1, \dots, \dot{p}_{3N}, \dot{q}_1, \dots, \dot{q}_{3N})$ is a generalized velocity vector.

Liouville Theorem

Applying the divergence theorem to Eq. (3), we find that ρ satisfies the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0, \quad (4)$$

where $\nabla = (\partial/\partial p_1, \dots, \partial/\partial p_{3N}, \partial/\partial q_1, \dots, \partial/\partial q_{3N})$. We can further simplify Eq. (4) since

$$\nabla \cdot v = \sum_{i=1}^{3N} \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) = \sum_{i=1}^{3N} \underbrace{\left(\frac{\partial}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} \right)}_{=0} = 0. \quad (5)$$

Liouville Theorem

Rewriting Eq. (4) using Poisson brackets¹ yields *Liouville's Theorem*

$$\frac{\partial \rho}{\partial t} = \{\mathcal{H}, \rho\} \quad (7)$$

which describes the time evolution of the phase space density ρ .

¹The Poisson bracket is defined as

$$\{u, v\} = \sum_i \left(\frac{\partial u}{\partial q_i} \frac{\partial v}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial v}{\partial q_i} \right) = -\{v, u\}. \quad (6)$$

Thermal Equilibrium

In *thermal equilibrium*, the system reaches a steady state in which the distribution of the configurations is constant and time-independent, i.e., $\partial\rho/\partial t = 0$. Liouville's theorem leads to the following condition

$$v \cdot \nabla \rho = \{\mathcal{H}, \rho\} = 0. \quad (8)$$

The last equation is satisfied if ρ depends on quantities which are conserved during the time evolution of the system. We then use such a phase space density ρ to replace the *time average*

$$\langle Q \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T Q(p(t), q(t)) \, dt. \quad (9)$$

by its ensemble average as defined by Eq. (1).

Ensemble Average

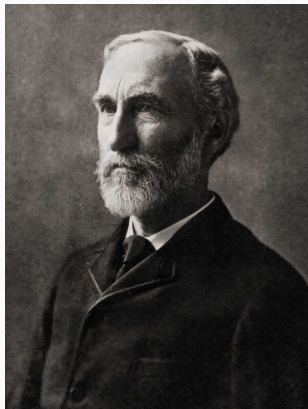
In the subsequent sections, we also consider discrete configurations X . In this case, we define the ensemble average as

$$\langle Q \rangle = \frac{1}{\Omega} \sum_X Q(X) \rho(X), \quad (10)$$

where Ω is the normalizing volume such that $\Omega^{-1} \sum_X \rho(X) = 1$. With the help of ensemble averages, systems can be described by means of some macroscopic quantities, such as temperature, energy and pressure.

Ensembles

- Microcanonical ensemble:
constant E, V, N
- Canonical ensemble: constant
 T, V, N
- Canonical pressure ensemble:
constant T, p, N
- Grandcanonical ensemble:
constant T, V, μ



Josiah W. Gibbs
(1839-1903)

Microcanonical Ensemble

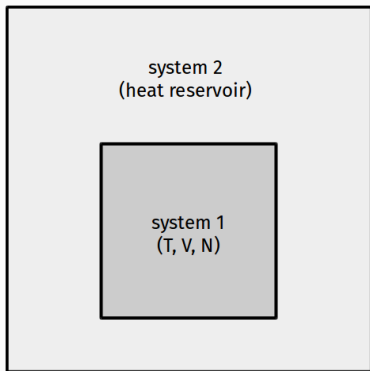
The *microcanonical ensemble* is defined by a constant number of particles, volume and energy. Thus, any configuration X of the system has the same energy $E(X) = \text{const.}$ The phase space density is also constant and given by

$$\rho(X) = \frac{1}{Z_{\text{mc}}} \delta(\mathcal{H}(X) - E), \quad (11)$$

with Z_{mc} being the *partition function* of the microcanonical ensemble

$$Z_{\text{mc}} = \sum_X \delta(\mathcal{H}(X) - E).$$

Canonical Ensemble



In a canonical ensemble setup, the system we study (system 1) is coupled to a heat reservoir (system 2) that guarantees a constant temperature.

Canonical Ensemble

At a given temperature T , the probability for a system to be in a certain configuration X with energy $E(X)$ is given by

$$\rho(X) = \frac{1}{Z_T} \exp \left[-\frac{E(X)}{k_B T} \right], \quad (12)$$

with

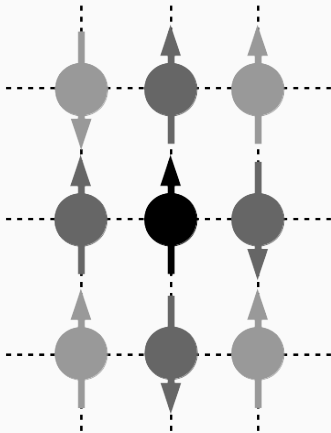
$$Z_T = \sum_X \exp \left[-\frac{E(X)}{k_B T} \right] \quad (13)$$

being the partition function of the canonical ensemble. According to the prior definition in Eq. (10), the ensemble average of a quantity Q is then given by

$$\langle Q \rangle = \frac{1}{Z_T} \sum_X Q(X) e^{-\frac{E(X)}{k_B T}}. \quad (14)$$

Ising Model

Ising Model



An illustration of the interaction of a magnetic dipole (black) with its nearest neighbors (dark grey) on a two-dimensional lattice.

Ising Model

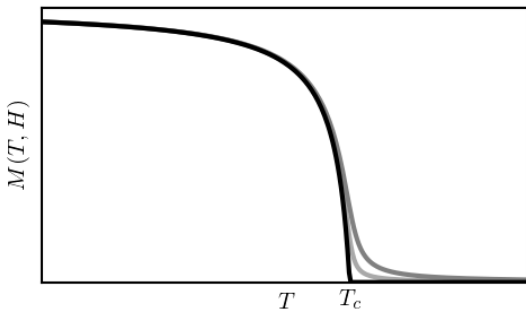
We consider a two-dimensional lattice with sites $\sigma_i \in \{1, -1\}$ which only interact with their nearest neighbors. Their interaction is described by the Hamiltonian

$$\mathcal{H}(\{\sigma\}) = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - H \sum_{i=1}^N \sigma_i, \quad (15)$$

where the first term denotes the interaction between all nearest neighbors represented by a sum over $\langle i, j \rangle$, and the second one the interaction of each site with an external magnetic field H .

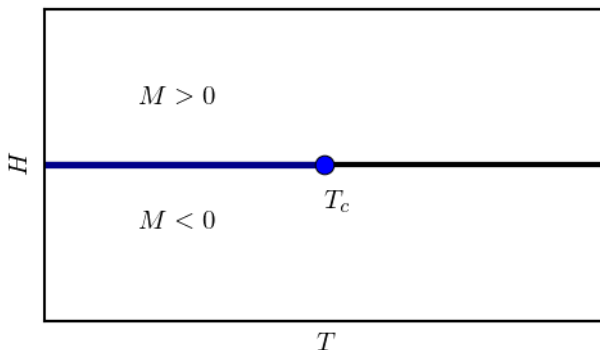
- Ferromagnetic case: $J > 0$ (parallel spins)
- Antiferromagnetic case: $J < 0$ (anti-parallel spins)
- No interaction: $J = 0$

Phase Transition in the Ferromagnetic Case



The magnetization $M(T, H)$ for different fields $H \geq 0$ as a function of T . For $T \leq T_c$, the system is characterized by a ferromagnetic phase. The black solid line represents the spontaneous magnetization $M_S(T)$ for $H = 0$ and should be interpreted in the sense that $\lim_{H \rightarrow 0^+} M(T, H)$.

Phase Transition in the Ferromagnetic Case



The first-order transition as a consequence of a sign change of the external field.

Order Parameter

The magnetization is defined as

$$M(T, H) = \left\langle \frac{1}{N} \sum_{i=1}^N \sigma_i \right\rangle, \quad (16)$$

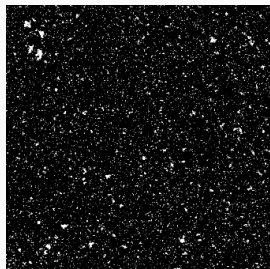
and corresponds to the ensemble average of the mean value of all spins.

On average, $M(T)$ vanishes since for every configuration there exists one of opposite sign which neutralizes the other one. As a consequence, we define the *order parameter* of the Ising model as

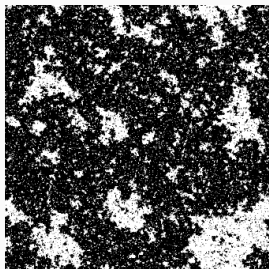
$$M_S(T) = \lim_{H \rightarrow 0^+} \left\langle \frac{1}{N} \sum_{i=1}^N \sigma_i \right\rangle \quad (17)$$

and refer to it as the *spontaneous magnetization*.

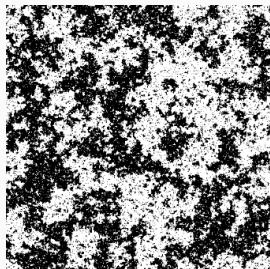
Magnetic Domains



$$T < T_c$$



$$T = T_c$$



$$T > T_c$$

The simulations have been performed on a square lattice with
 512×512 sites using
<https://mattbierbaum.github.io/ising.js/>.

Critical Exponents I

In the vicinity of the critical temperature for $T < T_c$, the spontaneous magnetization scales as

$$M_S(T) \propto (T_c - T)^\beta. \quad (18)$$

For $T = T_c$ and $H \rightarrow 0$, we find the following scaling

$$M(T = T_c, H) \propto H^{1/\delta}. \quad (19)$$

The exponents β and δ are so-called *critical exponents* and characterize together with other exponents the underlying phase transition.

$$\text{2D: } \beta = 1/8, \delta = 15$$

$$\text{3D: } \beta \approx 0.326, \delta \approx 4.790$$

Fluctuations

The magnetic susceptibility is defined as the change of the magnetization M in response to an applied magnetic field H , i.e.,

$$\chi(T) = \frac{\partial M(T, H)}{\partial H}. \quad (20)$$

We now use the definition of the spontaneous magnetization given by Eq. (17) and plug it into Eq. (20) leading to

$$\begin{aligned} \chi(T) &= \lim_{H \rightarrow 0^+} \frac{\partial \langle M(T, H) \rangle}{\partial H} \\ &= \lim_{H \rightarrow 0^+} \frac{\partial}{\partial H} \frac{1}{N} \frac{\sum_{\{\sigma\}} \sum_{i=1}^N \sigma_i \exp\left(\frac{E_0 + H \sum_{i=1}^N \sigma_i}{k_B T}\right)}{\underbrace{\sum_{\{\sigma\}} \exp\left(\frac{E_0 + H \sum_{i=1}^N \sigma_i}{k_B T}\right)}_{=Z_T(\mathcal{H})}}. \end{aligned}$$

Fluctuations

Using the product rule yields

$$\begin{aligned}\chi(T) &= \lim_{H \rightarrow 0^+} \frac{1}{Nk_B T} \frac{\sum_{\{\sigma\}} \left(\sum_{i=1}^N \sigma_i \right)^2 \exp \left(\frac{E_0 + H \sum_{i=1}^N \sigma_i}{k_B T} \right)}{Z_T(\mathcal{H})} \\ &\quad - \frac{1}{Nk_B T} \frac{\left[\sum_{\{\sigma\}} \sum_{i=1}^N \sigma_i \exp \left(\frac{E_0 + H \sum_{i=1}^N \sigma_i}{k_B T} \right) \right]^2}{[Z_T(\mathcal{H})]^2} \\ &= \frac{N}{k_B T} \left[\langle M_S(T)^2 \rangle - \langle M_S(T) \rangle^2 \right] \geq 0.\end{aligned}\tag{21}$$

The last equation defines the fluctuation-dissipation theorem for the magnetic susceptibility. Analogously, the specific heat is connected to energy fluctuations by

$$C(T) = \frac{\partial \langle E \rangle}{\partial T} = \frac{1}{k_B T^2} \left[\langle E(T)^2 \rangle - \langle E(T) \rangle^2 \right]. \tag{22}$$

Critical Exponents II

Similarly to the power-law scaling of the spontaneous magnetization defined in Eq. (18), we find for the magnetic susceptibility and the specific heat in the vicinity of T_c

$$\chi(T) \propto |T_c - T|^{-\gamma}, \quad (23)$$

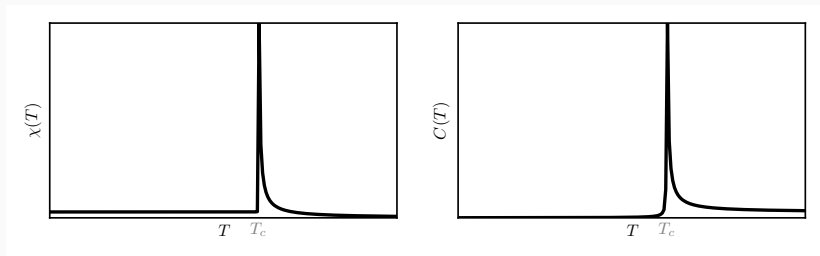
$$C(T) \propto |T_c - T|^{-\alpha}. \quad (24)$$

The corresponding exponents are in

$$\text{2D: } \gamma = 7/4, \alpha = 0$$

$$\text{3D: } \gamma \approx 1.24, \alpha \approx 0.11$$

Critical Exponents II



Susceptibility and specific heat as a function of temperature for the three dimensional Ising model. Both quantities diverge at the critical temperature T_c in the thermodynamic limit.

Correlation Length

The correlation function is defined by

$$G(r_1, r_2; T, H) = \langle \sigma_1 \sigma_2 \rangle - \langle \sigma_1 \rangle \langle \sigma_2 \rangle, \quad (25)$$

where the vectors r_1 and r_2 pointing in the direction of lattice sites 1 and 2.

Correlation Length

The correlation function is defined by

$$G(r_1, r_2; T, H) = \langle \sigma_1 \sigma_2 \rangle - \langle \sigma_1 \rangle \langle \sigma_2 \rangle, \quad (26)$$

where the vectors r_1 and r_2 pointing in the direction of lattice sites 1 and 2. If the system is translational and rotational invariant, the correlation function only depends on $r = |r_1 - r_2|$. At the critical point, the correlation function decays as

$$G(r; T_c, 0) \propto r^{-d+2-\eta}, \quad (27)$$

where η is another critical exponent and d the dimension of the system.

$$\text{2D: } \eta = 1/4$$

$$\text{3D: } \eta \approx 0.036$$

Correlation Length

For temperatures away from the critical temperature, the correlation function exhibits an exponential decay

$$G(r; T, 0) \propto r^{-\vartheta} e^{-r/\xi}, \quad (28)$$

where ξ defines the *correlation length*. The exponent ϑ equals 2 above and 1/2 below the transition point. In the vicinity of T_c , the correlation length ξ diverges since

$$\xi(T) \propto |T - T_c|^{-\nu}. \quad (29)$$

$$\text{2D: } \nu = 1$$

$$\text{3D: } \nu \approx 0.63$$

Critical Exponents and Universality

The aforementioned six critical exponents are connected by four scaling laws

$$\alpha + 2\beta + \gamma = 2 \quad (\text{Rushbrooke}), \quad (30)$$

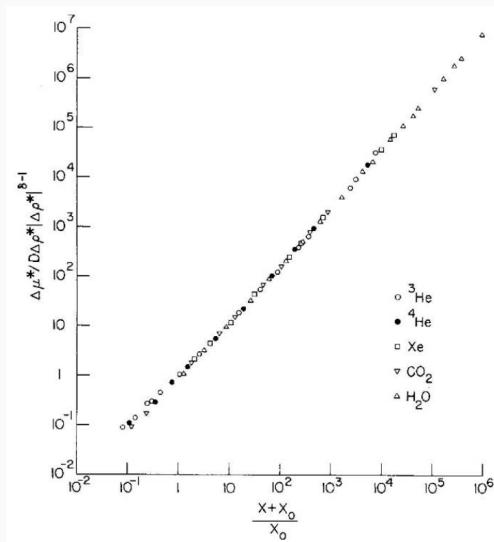
$$\gamma = \beta(\delta - 1) \quad (\text{Widom}), \quad (31)$$

$$\gamma = (2 - \eta)\nu \quad (\text{Fisher}), \quad (32)$$

$$2 - \alpha = d\nu \quad (\text{Josephson}), \quad (33)$$

which have been derived in the context of the phenomenological scaling theory for ferromagnetic systems [1, 2]. Due to these relations, the number of independent exponents reduces to two.

Critical Exponents and Universality



Universal scaling for five different gases. The figure is taken from Ref. [3].

Critical Exponents and Universality

The critical exponents of the Ising model in two and three dimensions [4].

Exponent	$d = 2$	$d = 3$
α	0	0.110(1)
β	1/8	0.3265(3)
γ	7/4	1.2372(5)
δ	15	4.789(2)
η	1/4	0.0364(5)
ν	1	0.6301(4)

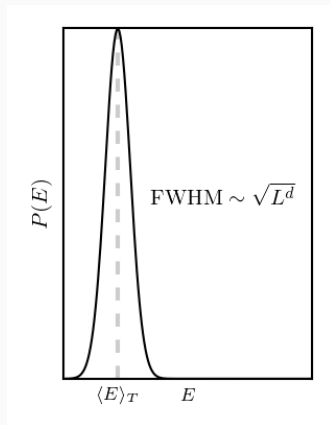
Monte Carlo Methods

Monte Carlo Methods

The main steps of the Monte Carlo sampling are

1. Choose randomly a new configuration in phase space based on a Markov chain.
2. Accept or reject the new configuration, depending on the strategy used (e.g., Metropolis or Glauber dynamics).
3. Compute the physical quantity and add it to the averaging procedure.
4. Repeat the previous steps.

Markov Chains



Example of an energy distribution with a system size L dependence of the distribution width which scales as $\propto \sqrt{L^d}$ where d is the system dimension.

Markov Chains

In terms of a Markov chain, the transition probability from one state to another is given by the probability of a new state to be proposed (T) and the probability of this state to be accepted (A). Specifically, $T(X \rightarrow Y)$ is the probability that a new configuration Y is proposed, starting from configuration X . For the thermodynamic systems we consider, the transition probability fulfills three conditions:

1. *Ergodicity*: any configuration in the phase space must be reachable within a finite number of steps,
2. *Normalization*: $\sum_Y T(X \rightarrow Y) = 1$,
3. *Reversibility*: $T(X \rightarrow Y) = T(Y \rightarrow X)$.

Markov Chains

Once a configuration is proposed, we can accept the new configuration with probability $A(X \rightarrow Y)$ or reject it with probability $1 - A(X \rightarrow Y)$. The *probability of the Markov chain* is then given by

$$W(X \rightarrow Y) = T(X \rightarrow Y) \cdot A(X \rightarrow Y). \quad (34)$$

Markov Chains

We denote the probability to find the system in a certain configuration X at virtual time τ by $p(X, \tau)$. The *master equation* describes the time evolution of $p(X, \tau)$ and is given by

$$\frac{dp(X, \tau)}{d\tau} = \sum_Y p(Y)W(Y \rightarrow X) - \sum_Y p(X)W(X \rightarrow Y). \quad (35)$$

Markov Chains

A stationary state p_{st} is reached if $\frac{dp(X,\tau)}{d\tau} = 0$. The probability of the Markov chain fulfills the following properties:

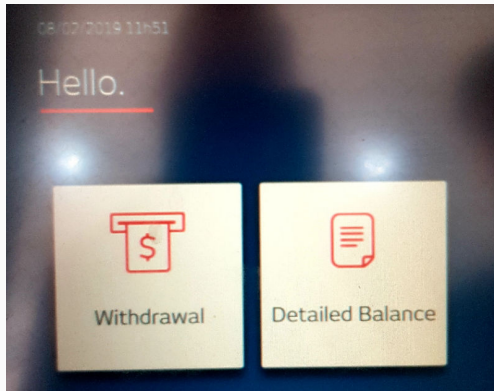
1. *Ergodicity*: any configuration must be reachable: $\forall X, Y :$
 $W(X \rightarrow Y) \geq 0,$
2. *Normalization*: $\sum_Y W(X \rightarrow Y) = 1,$
3. *Homogeneity*: $\sum_Y p_{\text{st}}(Y)W(Y \rightarrow X) = p_{\text{st}}(X).$

Markov Chains

To efficiently sample the relevant regions of the phase space, the probability of the Markov chain $W(\cdot)$ has to depend on the system properties. To achieve that, we set the stationary distribution p_{st} equal to the equilibrium distribution of the physical system p_{eq} (a real and measurable distribution):

$$\frac{dp(X, \tau)}{d\tau} = 0 \Leftrightarrow p_{\text{st}} \stackrel{!}{=} p_{\text{eq}}. \quad (36)$$

Markov Chains



Markov Chains

It then follows from the stationary state condition of the Markov chain that

$$\sum_Y p_{\text{eq}}(Y)W(Y \rightarrow X) = \sum_Y p_{\text{eq}}(X)W(X \rightarrow Y). \quad (37)$$

A sufficient condition for this to be true is

$$p_{\text{eq}}(Y)W(Y \rightarrow X) = p_{\text{eq}}(X)W(X \rightarrow Y), \quad (38)$$

which is referred to as *condition of detailed balance*.

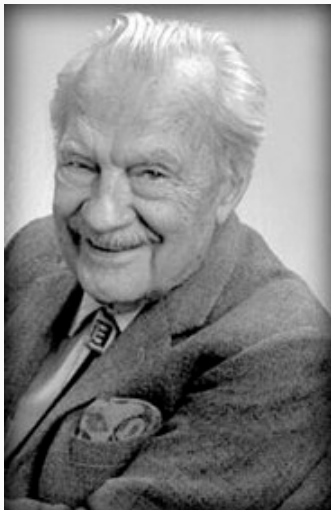
Markov Chains

As an example, in a canonical ensemble at fixed temperature T , the equilibrium distribution is given by the Boltzmann distribution

$$p_{\text{eq}}(X) = \frac{1}{Z_T} \exp \left[-\frac{E(X)}{k_B T} \right] \quad (39)$$

with the partition function $Z_T = \sum_X \exp \left[-\frac{E(X)}{k_B T} \right]$.

$M(RT)^2$ Algorithm



Nicholas C. Metropolis (1915-1999) introduced the $M(RT)^2$ sampling technique.

M(RT)² Algorithm

One possible choice of the acceptance probability fulfilling the detailed balance condition is given by

$$A(X \rightarrow Y) = \min \left[1, \frac{p_{\text{eq}}(Y)}{p_{\text{eq}}(X)} \right]. \quad (40)$$

which can be obtained by rewriting Eq. (38).

M(RT)² Algorithm

In the case of the canonical ensemble with

$p_{\text{eq}}(X) = \frac{1}{Z_T} \exp \left[-\frac{E(X)}{k_B T} \right]$, the acceptance probability becomes

$$A(X \rightarrow Y) = \min \left[1, \exp \left(-\frac{\Delta E}{k_B T} \right) \right], \quad (41)$$

where $\Delta E = E(Y) - E(X)$. The last equation implies that the Monte Carlo step is always accepted if the energy decreases, and if the energy increases, it is accepted with probability $\exp \left(-\frac{\Delta E}{k_B T} \right)$.

$M(RT)^2$ Algorithm

In summary, the steps of the $M(RT)^2$ algorithm applied to the Ising model are

$M(RT)^2$ Algorithm

- Randomly choose a lattice site i ,
- Compute $\Delta E = E(Y) - E(X) = 2J\sigma_i h_i$,
- Flip the spin if $\Delta E \leq 0$, otherwise accept it with probability $\exp\left(-\frac{\Delta E}{k_B T}\right)$,

with $h_i = \sum_{\langle i,j \rangle} \sigma_j$ and $E = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j$.

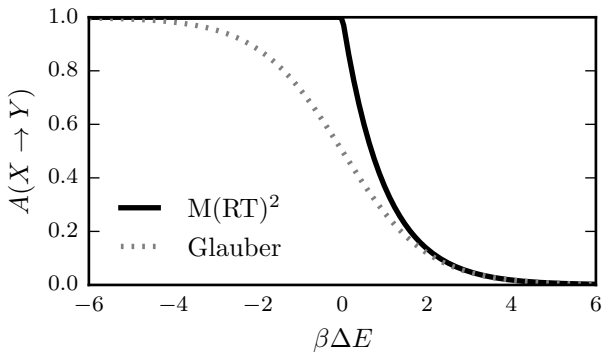
Glauber Dynamics

The Metropolis algorithm is not the only possible choice to fulfill the detailed balance condition. Another acceptance probability given by

$$A_G(X \rightarrow Y) = \frac{\exp\left(-\frac{\Delta E}{k_B T}\right)}{1 + \exp\left(-\frac{\Delta E}{k_B T}\right)} \quad (42)$$

has been suggested by Glauber.

Glauber Dynamics



A comparison of the acceptance probabilities of $M(RT)^2$ and Glauber dynamics.

Glauber Dynamics

In contrast to the $M(RT)^2$ acceptance probability, updates with $\Delta E = 0$ are not always accepted but with probability $1/2$.

To proof that Eq. (42) satisfies the condition of detailed balance, we have to show that

$$p_{\text{eq}}(Y)A_G(Y \rightarrow X) = p_{\text{eq}}(X)A_G(X \rightarrow Y), \quad (43)$$

since $T(Y \rightarrow X) = T(X \rightarrow Y)$.

Glauber Dynamics

The last equation is equivalent to

$$\frac{p_{\text{eq}}(Y)}{p_{\text{eq}}(X)} = \frac{A_G(X \rightarrow Y)}{A_G(Y \rightarrow X)} \quad (44)$$

which is fulfilled since

$$\frac{p_{\text{eq}}(Y)}{p_{\text{eq}}(X)} = \exp\left(-\frac{\Delta E}{k_B T}\right) \quad (45)$$

and

$$\frac{A_G(X \rightarrow Y)}{A_G(Y \rightarrow X)} = \frac{\exp\left(-\frac{\Delta E}{k_B T}\right)}{1 + \exp\left(-\frac{\Delta E}{k_B T}\right)} \left[\frac{\exp\left(\frac{\Delta E}{k_B T}\right)}{1 + \exp\left(\frac{\Delta E}{k_B T}\right)} \right]^{-1} = \exp\left(-\frac{\Delta E}{k_B T}\right) \quad (46)$$

Glauber Dynamics

As in the $M(RT)^2$ algorithm, only the local configuration around the lattice site is relevant for the update procedure. Furthermore, with $J = 1$, the probability to flip spin σ_i is

$$A_G(X \rightarrow Y) = \frac{\exp\left(\frac{-2\sigma_i h_i}{k_B T}\right)}{1 + \exp\left(\frac{-2\sigma_i h_i}{k_B T}\right)} \quad (47)$$

with $h_i = \sum_{\langle i,j \rangle} \sigma_j$ being the local field and $X = \{\dots, \sigma_{i-1}, \sigma_i, \sigma_{i+1}, \dots\}$ and $Y = \{\dots, \sigma_{i-1}, -\sigma_i, \sigma_{i+1}, \dots\}$ the initial and final configuration, respectively.

Glauber Dynamics

We abbreviate the probability defined by Eq. (47) as p_i . The spin flip and no flip probabilities can then be expressed as

$$p_{\text{flip}} = \begin{cases} p_i & \text{for } \sigma_i = -1 \\ 1 - p_i & \text{for } \sigma_i = +1 \end{cases} \quad \text{and} \quad p_{\text{no flip}} = \begin{cases} 1 - p_i & \text{for } \sigma_i = -1 \\ p_i & \text{for } \sigma_i = +1 \end{cases} \quad (48)$$

Glauber Dynamics

A possible implementation is

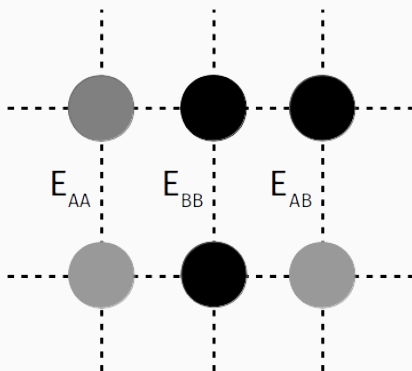
$$\sigma_i(\tau + 1) = -\sigma_i(\tau) \cdot \text{sign}(p_i - z), \quad (49)$$

with $z \in (0, 1)$ being a uniformly distributed random number,
or

$$\sigma_i(\tau+1) = \begin{cases} +1 & \text{with propability } p_i \\ -1 & \text{with propability } 1 - p_i \end{cases} \quad \text{and} \quad p_i = \frac{\exp(2\beta h_i)}{1 + \exp(2\beta h_i)}. \quad (50)$$

This method does not depend on the spin value at time t and is called *heat bath Monte Carlo*.

Binary Mixtures



An example of a binary mixture consisting of two different atoms A and B .

Binary Mixtures

Kawasaki dynamics

- Choose a $A - B$ bond,
- Compute ΔE for $A - B \rightarrow B - A$,
- Metropolis: If $\Delta E \leq 0$ flip, else flip with probability

$$p = \exp \left(\frac{-\Delta E}{k_B T} \right),$$

- Glauber: Flip with probability

$$p = \exp \left(-\frac{\Delta E}{k_B T} \right) / \left[1 + \exp \left(-\frac{\Delta E}{k_B T} \right) \right].$$

VOLUME 50, NUMBER 19

PHYSICAL REVIEW LETTERS

9 MAY 1983

Microcanonical Monte Carlo Simulation

Michael Creutz

Department of Physics, Brookhaven National Laboratory, Upton, New York 11973

(Received 24 February 1983)

A new algorithm for the simulation of statistical systems is presented. The procedure produces a random walk through configurations of a constant total energy. It is computationally simple and applicable to systems of both discrete and continuous variables.

Creutz Algorithm

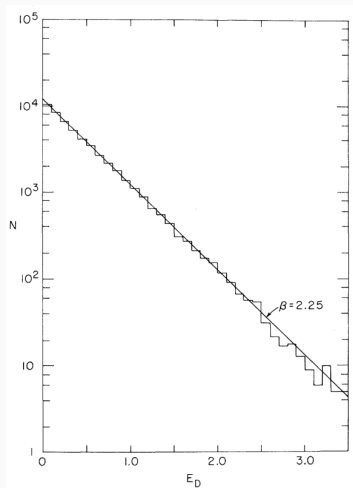
The movement in phase space is therefore not strictly constrained to a subspace of constant energy but there is a certain additional volume in which we can freely move. The condition of constant energy is softened by introducing a so-called *demon* which corresponds to a small reservoir of energy E_D that can store a certain maximum energy $E_{\max}$.

Creutz Algorithm

Creutz Algorithm

- Choose a site,
- Compute ΔE for the spin flip,
- Accept the change if $E_{\max} \geq E_D - \Delta E \geq 0$.

Creutz Algorithm



The distribution of the demon energy E_D is exponentially distributed. Based on the Boltzmann factor, it is possible to extract the inverse temperature $\beta = (k_B T)^{-1} = 2.25$.

Q2R

In the case of $E_{\max} \rightarrow 0$, the Creutz algorithm resembles a totalistic cellular automaton called Q2R [5]. The update rules on a square lattice for spins $\sigma_{ij} \in \{0, 1\}$ are given by

$$\sigma_{ij}(\tau + 1) = f(x_{ij}) \oplus \sigma_{ij}(\tau) \quad (51)$$

with

$$x_{ij} = \sigma_{i-1j} + \sigma_{i+1j} + \sigma_{ij-1} + \sigma_{ij+1} \quad \text{and} \quad f(x) = \begin{cases} 1 & \text{if } x = 2 \\ 0 & \text{if } x \neq 2 \end{cases} . \quad (52)$$

Boundary Conditions

For finite lattices, the following boundary conditions may be used:

- Open boundaries, i.e., no neighbors at the edges of the system,
- fixed boundary conditions,
- and periodic boundaries.

Temporal Correlations

According to the definition of a Markov chain, the dependence of a quantity A on virtual time τ is given by

$$\langle A(\tau) \rangle = \sum_X p(X, \tau) A(X) = \sum_X p(X, \tau_0) A(X(\tau)). \quad (53)$$

In the second step of the last equation, we used the fact that the average is taken over an ensemble of initial configurations $X(\tau_0)$ which evolve according to Eq. (35) [6].

Temporal Correlations

For some $\tau_0 < \tau$, the *non-linear correlation function*

$$\Phi_A^{\text{nl}}(\tau) = \frac{\langle A(\tau) \rangle - \langle A(\infty) \rangle}{\langle A(\tau_0) \rangle - \langle A(\infty) \rangle} \quad (54)$$

is a measure to quantify the deviation of $A(\tau)$ from $A(\infty)$ relative to the deviation of $A(\tau_0)$ from $A(\infty)$.

Temporal Correlations

The *non-linear* correlation time τ_A^{nl} describes the relaxation towards equilibrium and is defined as²

$$\tau_A^{\text{nl}} = \int_0^{\infty} \Phi_A^{\text{nl}}(\tau) \, d\tau. \quad (56)$$

²If we consider an exponential decay of $\Phi_A^{\text{nl}}(\tau)$, we find that this definition is meaningful since

$$\int_0^{\infty} \exp\left(-\tau/\tau_A^{\text{nl}}\right) d\tau = \tau_A^{\text{nl}}. \quad (55)$$

Temporal Correlations

In the vicinity of the critical temperature T_c , we observe the so-called *critical slowing down* of our dynamics, i.e., the non-linear correlation time is described by power law

$$\tau_A^{\text{nl}} \propto |T - T_c|^{-z_A^{\text{nl}}} \quad (57)$$

with z_A^{nl} being the non-linear dynamical critical exponent. This is very bad news, because the last equation implies that the time needed to reach equilibrium diverges at T_c .

Temporal Correlations

The linear correlation function of two quantities A and B in equilibrium is defined as

$$\Phi_{AB}(\tau) = \frac{\langle A(\tau_0)B(\tau) \rangle - \langle A \rangle \langle B \rangle}{\langle AB \rangle - \langle A \rangle \langle B \rangle} \quad (58)$$

with

$$\langle A(\tau_0) B(\tau) \rangle = \sum_X p(X, \tau_0) A(X(\tau_0)) B(X(\tau)).$$

As τ goes to infinity, $\Phi_{AB}(\tau)$ decreases from unity to zero.

Temporal Correlations

If $A = B$, we call Eq. (58) the *autocorrelation function*. For the spin-spin correlation in the Ising model we obtain

$$\Phi_{\sigma}(\tau) = \frac{\langle \sigma(\tau_0)\sigma(\tau) \rangle - \langle \sigma(\tau_0) \rangle^2}{\langle \sigma^2(\tau_0) \rangle - \langle \sigma(\tau_0) \rangle^2}. \quad (59)$$

Temporal Correlations

The *linear* correlation time τ_{AB} describes the relaxation in equilibrium

$$\tau_{AB} = \int_0^{\infty} \Phi_{AB}(\tau) d\tau. \quad (60)$$

Temporal Correlations

As in the case of the non-linear correlation time, in the vicinity of T_c , we observe a *critical slowing down*, i.e.,

$$\tau_{AB} \propto |T - T_c|^{-z_A} . \quad (61)$$

with z_A being the *linear* dynamical critical exponent.

Temporal Correlations

The dynamical exponents for spin correlations turn out to be

$$z_\sigma = 2.16 \text{ (2D)}, \quad (62)$$

$$z_\sigma = 2.09 \text{ (3D)}. \quad (63)$$

There is a conjectured relation between the Ising critical exponents and the critical dynamical exponents for spin σ and energy correlations E . The relations

$$z_\sigma - z_\sigma^{\text{nl}} = \beta, \quad (64)$$

$$z_E - z_E^{\text{nl}} = 1 - \alpha, \quad (65)$$

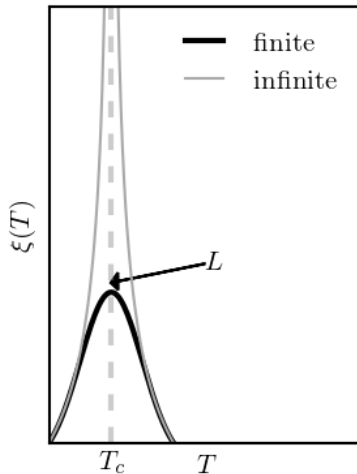
are numerically well-established, however, not yet analytically proven.

Decorrelated Configurations

Connecting this behavior with the one observed for the correlation time described by Eq. (60) yields

$$\tau_{AB} \propto |T - T_c|^{-z_{AB}} \propto L^{\frac{z_{AB}}{\nu}}. \quad (66)$$

Decorrelated Configurations



Decorrelated Configurations

To ensure not to sample correlated configurations one should

- first reach equilibrium (discard $n_0 = c\tau^{\text{nl}}(T)$ configurations),
- only sample every $n_e^{\text{th}} = c\tau(T)$ configuration,
- and at T_c use $n_0 = cL^{\frac{z_{\text{nl}}}{\nu}}$ and $n_e = cL^{\frac{z}{\nu}}$

where $c \approx 3$ is a "safety factor" to make sure to discard enough samples.

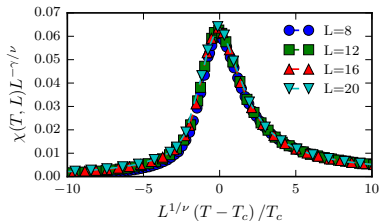
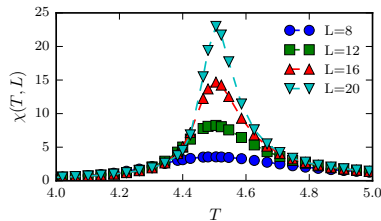
Finite Size Methods

Finite size methods

Divergent behavior at T_c

$$\begin{aligned}\chi(T) &\propto |T_c - T|^{-\gamma}, \\ C(T) &\propto |T_c - T|^{-\alpha}, \\ \xi(T) &\propto |T_c - T|^{-\nu}.\end{aligned}\tag{67}$$

Finite size methods



The system size dependence of the susceptibility and the corresponding finite size scaling.

Finite size methods

The finite size scaling relation of the susceptibility is given by

$$\chi(T, L) = L^{\frac{\gamma}{\nu}} F_{\chi} \left[(T - T_c) L^{\frac{1}{\nu}} \right], \quad (68)$$

where F_{χ} is called *susceptibility scaling function*³.

³Based on Eq. (23), we can infer that $F_{\chi} \left[(T - T_c) L^{\frac{1}{\nu}} \right] \propto \left(|T - T_c| L^{\frac{1}{\nu}} \right)^{-\gamma}$ as $L \rightarrow \infty$.

Finite size methods

In the case of the magnetization, the corresponding finite size scaling relation is

$$M_S(T, L) = L^{-\frac{\beta}{\nu}} F_{M_S} \left[(T - T_c) L^{\frac{1}{\nu}} \right]. \quad (69)$$

Binder Cumulant

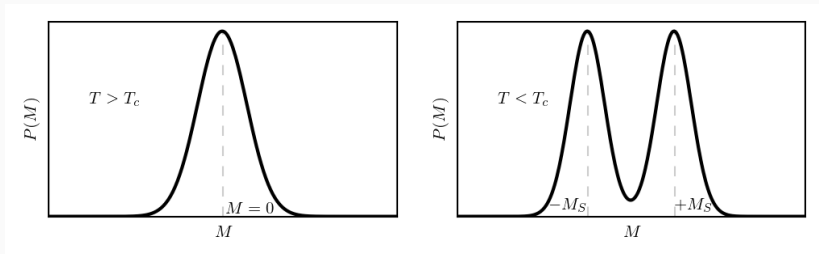
We still need a method to better determine T_c and make therefore use of the Binder cumulant

$$U_L(T) = 1 - \frac{\langle M^4 \rangle_L}{3 \langle M^2 \rangle_L^2}, \quad (70)$$

which is independent of the system size L at T_c since

$$\frac{\langle M^4 \rangle_L}{3 \langle M^2 \rangle_L^2} = \frac{L^{-\frac{4\beta}{\nu}} F_{M4} \left[(T - T_c) L^{\frac{1}{\nu}} \right]}{\left\{ L^{-\frac{2\beta}{\nu}} F_{M2} \left[(T - T_c) L^{\frac{1}{\nu}} \right] \right\}^2} = F_C \left[(T - T_c) L^{\frac{1}{\nu}} \right]. \quad (71)$$

Binder Cumulant



The distribution $P(M)$ of the magnetization M above and below the critical temperature T_c .

Binder Cumulant

For $T > T_c$, the magnetization is described by a Gaussian distribution

$$P_L(M) = \sqrt{\frac{L^d}{\pi\sigma_L}} \exp\left[-\frac{M^2 L^d}{\sigma_L}\right], \quad (72)$$

with $\sigma_L = k_B T 2\chi_L$. Since the fourth moment equals three times the second moment squared, i.e.,

$$\langle M^4 \rangle_L = 3 \langle M^2 \rangle_L^2, \quad (73)$$

it follows that $U_L(T)$ must be zero for $T > T_c$.

Binder Cumulant

Below the critical temperature ($T < T_c$), there exist one ground state with positive and one with negative magnetization and the corresponding distribution is given by

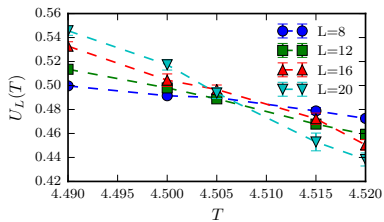
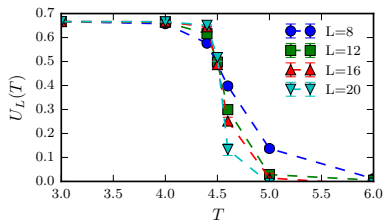
$$P_L(M) = \frac{1}{2} \sqrt{\frac{L^d}{\pi \sigma_L}} \left\{ \exp \left[-\frac{(M - M_S)^2 L^d}{\sigma_L} \right] + \exp \left[-\frac{(M + M_S)^2 L^d}{\sigma_L} \right] \right\}. \quad (74)$$

Binder Cumulant

For this distribution, it holds that $\langle M^4 \rangle_L = \langle M^2 \rangle_L^2$ and therefore $U_L(T) = \frac{2}{3}$ for $T < T_c$. In summary, we demonstrated that

$$U_L(T) = \begin{cases} \frac{2}{3} & \text{for } T < T_c \\ \text{const.} & \text{for } T = T_c \\ 0 & \text{for } T > T_c \end{cases} \quad (75)$$

Binder Cumulant



Corrections to Scaling

Far away from T_c we cannot observe a clear power law behavior anymore, and corrections to scaling have to be employed, i.e.,

$$M(T) = A_0 (T_c - T)^\beta + A_1 (T_c - T)^{\beta_1} + \dots, \quad (76)$$

$$\xi(T) = C_0 (T_c - T)^{-\nu} + C_1 (T_c - T)^{-\nu_1} + \dots, \quad (77)$$

with $\beta_1 > \beta$ and $\nu_1 < \nu$.

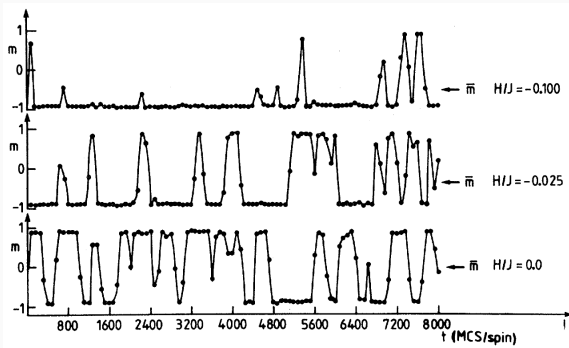
Corrections to Scaling

These corrections are very important for high quality data, where the errors are small and the deviations become visible. The scaling functions must also be generalized as

$$M(T, L) = L^{-\frac{\beta}{\nu}} F_M \left[(T - T_c) L^{\frac{1}{\nu}} \right] + L^{-x} F_M^1 \left[(T - T_c) L^{\frac{1}{\nu}} \right] + \dots \quad (78)$$

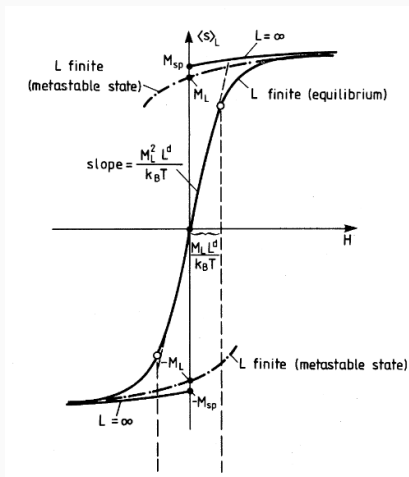
with $x = \max \left[\frac{\beta_1}{\nu}, \frac{\beta}{\nu_1}, \frac{\beta}{\nu} - 1 \right]$.

First Order Transition



The magnetization exhibits a switching behavior if the field vanishes. For non-zero magnetic fields, the magnetization is driven in the direction of the field. The figure is taken from Ref. [7].

First Order Transition



Hysteresis can be found by varying the field from negative to positive values and back. The figure is taken from Ref. [7].

First Order Transition

Binder showed that the magnetization as a function of the field H is described by $\tanh(\alpha L^d)$ if the distribution of the magnetization is given by Eq. (74) [7]. Specifically, we find for the magnetization and susceptibility

$$M(H) = \chi_L^D H + M_L \tanh(\beta H M_L L^d), \quad (79)$$

$$\chi_L(H) = \frac{\partial M}{\partial H} = \chi_L^D + \frac{\beta M_L L^d}{\cosh^2(\beta H M_L L^d)}. \quad (80)$$

First Order Transition

Similarly, to the scaling of a second order transition, we can scale the maximum of the susceptibility ($\chi_L (H = 0) \propto L^d$) and the width of the peak ($\Delta\chi_L \propto L^{-d}$). To summarize, a first order phase transition is characterized by

1. A bimodal distribution of the order parameter,
2. stochastic switching between the two states in small systems,
3. hysteresis of the order parameter when changing the field,
4. a scaling of the order parameter, or response function according to Eq. (80).

Cluster Algorithms

Potts Model

The Hamiltonian of the system is defined as

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \delta_{\sigma_i \sigma_j} - H \sum_i \sigma_i, \quad (81)$$

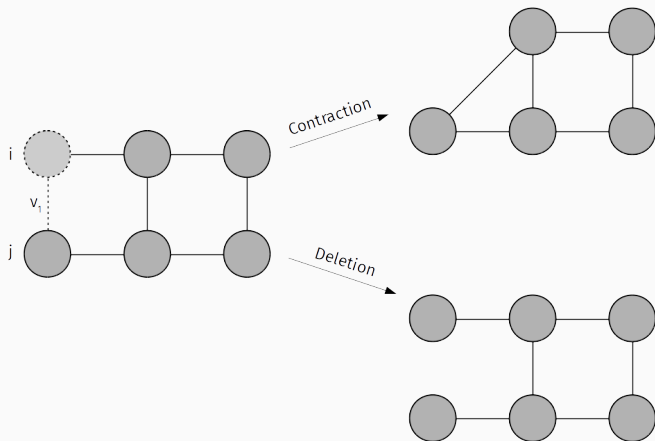
where $\sigma_i \in \{1, \dots, q\}$ and $\delta_{\sigma_i \sigma_j}$ is unity when nodes i and j are in the same state. The Potts model exhibits a *first order transition* at the critical temperature in two dimensions for $q > 4$, and for $q > 2$ for dimensions larger than two.

The Kasteleyn and Fortuin Theorem

We consider the Potts model not on a square lattice but on an arbitrary graph of nodes connected with bonds ν . Each node has q possible states and each connection leads to an energy cost of unity if two connected nodes are in a different state and of zero if they are in the same state, i.e.,

$$E = J \sum_{\nu} \epsilon_{\nu} \quad \text{with} \quad \epsilon_{\nu} = \begin{cases} 0 & \text{if endpoints are in the same state} \\ 1 & \text{otherwise} \end{cases} \quad (82)$$

The Kasteleyn and Fortuin Theorem



Contraction and deletion on a graph.

The Kasteleyn and Fortuin Theorem

The partition function is the sum over all possible configurations weighted by the Boltzmann factor and thus given by

$$Z = \sum_X e^{-\beta E(X)} \stackrel{(82)}{=} \sum_X e^{-\beta J \sum_{\nu} \epsilon_{\nu}} = \sum_X \prod_{\nu} e^{-\beta J \epsilon_{\nu}}. \quad (83)$$

The Kasteleyn and Fortuin Theorem

We now consider a graph where bond ν_1 connects two nodes i and j with states σ_i and σ_j , respectively. If we would delete bond ν_1 , the partition function is

$$Z_D = \sum_X \prod_{\nu \neq \nu_1} e^{-\beta J_{\epsilon_\nu}}. \quad (84)$$

The Kasteleyn and Fortuin Theorem

We can thus rewrite Eq. (83) as

$$\begin{aligned} Z &= \sum_X e^{-\beta J \epsilon_{v_1}} \prod_{\nu \neq v_1} e^{-\beta J \epsilon_\nu} \\ &= \sum_{X: \sigma_i = \sigma_j} \prod_{\nu \neq v_1} e^{-\beta J \epsilon_\nu} + e^{-\beta J} \sum_{X: \sigma_i \neq \sigma_j} \prod_{\nu \neq v_1} e^{-\beta J \epsilon_\nu}, \end{aligned}$$

where the first part is the partition function of the contracted graph Z_C and the second part is given by the identity

$$\sum_{X: \sigma_i \neq \sigma_j} \prod_{\nu \neq v_1} e^{-\beta J \epsilon_\nu} = \sum_X \prod_{\nu \neq v_1} e^{-\beta J \epsilon_\nu} - \sum_{X: \sigma_i = \sigma_j} \prod_{\nu \neq v_1} e^{-\beta J \epsilon_\nu} = Z_D - Z_C. \quad (85)$$

The Kasteleyn and Fortuin Theorem

Summarizing the last results, we find

$$Z = Z_C + e^{-\beta J} (Z_D - Z_C) = pZ_C + (1-p)Z_D, \quad (86)$$

where $p = 1 - e^{-\beta J}$. To be more precise, we expressed the partition function Z as the contracted and deleted partition functions at bond ν_1 . We apply the last procedure to another bond ν_2 and find

$$Z = p^2 Z_{C_{\nu_1}, C_{\nu_2}} + p(1-p) Z_{C_{\nu_1}, D_{\nu_2}} + (1-p)p Z_{D_{\nu_1}, C_{\nu_2}} + (1-p)^2 Z_{D_{\nu_1}, D_{\nu_2}}. \quad (87)$$

The Kasteleyn and Fortuin Theorem

After applying these operations to every bond, the graph is reduced to a set of separated points corresponding to clusters of nodes which are connected and in the same state out of q states. The partition function reduces to

$$Z = \sum_{\text{configurations of bond percolation}} q^{\# \text{ of clusters}} p^c (1-p)^d = \left\langle q^{\# \text{ of clusters}} \right\rangle_b, \quad (88)$$

where c and d are the numbers of contracted and deleted bonds respectively. In the limit of $q \rightarrow 1$, one obtains the partition function of bond percolation⁴.

⁴In bond percolation, an edge of a graph is occupied with probability p and vacant with probability $1 - p$.

Coniglio-Klein Clusters

The probability of a given cluster C to be in a certain state σ_0 is independent of the state itself, i.e.,

$$p(C, \sigma_0) = p^{c_C} (1 - p)^{d_C} \sum_{\substack{\text{bond percolation} \\ \text{without cluster } C}} q^{\# \text{ of clusters}} p^c (1 - p)^d. \quad (89)$$

Coniglio-Klein Clusters

This implies that flipping this particular cluster has no effect on the partition function (and therefore the energy) so that it is possible to accept the flip with probability one. This can be seen by looking at the detailed balance condition of the system

$$p(C, \sigma_1)W [(C, \sigma_1) \rightarrow (C, \sigma_2)] = p(C, \sigma_2)W [(C, \sigma_2) \rightarrow (C, \sigma_1)] \quad (90)$$

and using $p(C, \sigma_1) = p(C, \sigma_2)$.

Coniglio-Klein Clusters

We then obtain for the acceptance probabilities

$$\begin{aligned} A[(C, \sigma_2) \rightarrow (C, \sigma_1)] &= \min \left[1, \frac{p(C, \sigma_2)}{p(C, \sigma_1)} \right] = 1 \quad \text{M(RT)}^2, \\ A[(C, \sigma_2) \rightarrow (C, \sigma_1)] &= \frac{p(C, \sigma_2)}{p(C, \sigma_1) + p(C, \sigma_2)} = \frac{1}{2} \quad \text{Glauber.} \end{aligned} \tag{91}$$

Based on these insights, we introduce cluster algorithms which are much faster than single-spin flip algorithms and less prone to the problem of critical slowing down.

Swendsen-Wang Algorithm

Swendsen-Wang algorithm

- Occupy the bonds with probability $p = 1 - e^{-\beta J}$ if sites are in the same state.
- Identify the clusters with the Hoshen-Kopelman algorithm.
- Flip the clusters with probability $1/2$ for Ising or always choose a new state for $q > 2$.
- Repeat the procedure.

Wolff Algorithm

Wolff algorithm

- Choose a site randomly.
- If the neighboring sites are in the same state, add them to the cluster with probability $p = 1 - e^{-\beta J}$.
- Repeat this for any site on the boundaries of the cluster, until all the bonds of the cluster have been checked exactly once.
- Choose a new state for the cluster.
- Repeat the procedure.

Other Ising-like Models

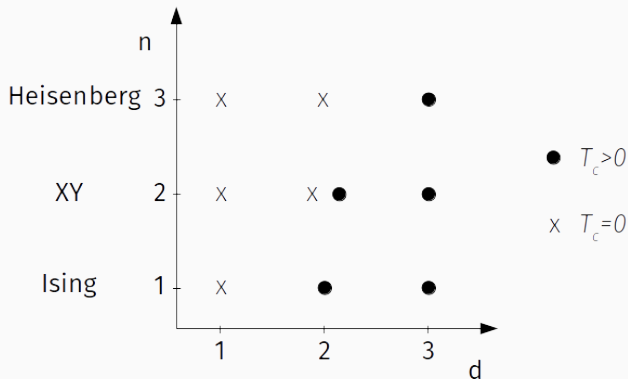
One of the possible generalizations of the Ising model is the so called n -vector model.

The Hamiltonian resembles the one of the Potts model in the sense that it favors spin alignment

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j + \vec{H} \sum_i \vec{S}_i. \quad (92)$$

with $\vec{S}_i = (S_i^1, S_i^2, \dots, S_i^n)$ and $|\vec{S}_i| = 1$.

Other Ising-like Models



The dependence of the critical temperature on the number of vector components n .

Other Ising-like Models

For Monte Carlo simulations with vector-valued spins we have to adapt our simulation methods. The classical strategy is to flip spins by modifying the spin locally through adding a small $\Delta\vec{S}$ such that $\vec{S}'_i = \vec{S}_i + \Delta\vec{S}$ and $\Delta\vec{S} \perp \vec{S}_i$. The classical Metropolis algorithm can then be used in the same fashion as in the Ising model.

Histogram Methods

Histogram Methods

For computing the thermal average defined by Eq. (14), we need to sample different configurations at different temperatures. Another possibility would be to determine an average at a certain temperature T_0 and extrapolate to another temperature T . In the case of a canonical ensemble, an extrapolation can be achieved by reweighing the histogram of energies $p_{T_0}(E)$ with the Boltzmann factor $\exp(E/T - E/T_0)$.

Histogram Methods

Such histogram methods have first been described in Ref. [8]. We now reformulate the computation of the thermal average of a quantity Q and of the partition function as a sum over all possible energies instead of over all possible configurations and find

$$Q(T_0) = \frac{1}{Z_{T_0}} \sum_E Q(E) p_{T_0}(E) \quad \text{with} \quad Z_{T_0} = \sum_E p_{T_0}(E), \quad (93)$$

where $p_{T_0}(E) = g(E) e^{-\frac{E}{k_B T_0}}$ with $g(E)$ defining the *degeneracy of states*, i.e., the number of states with energy E .

Histogram Methods

This takes into account the fact that multiple configurations can have the same energy. The goal is to compute the quantity Q at another temperature T

$$Q(T) = \frac{1}{Z_T} \sum_E Q(E) p_T(E). \quad (94)$$

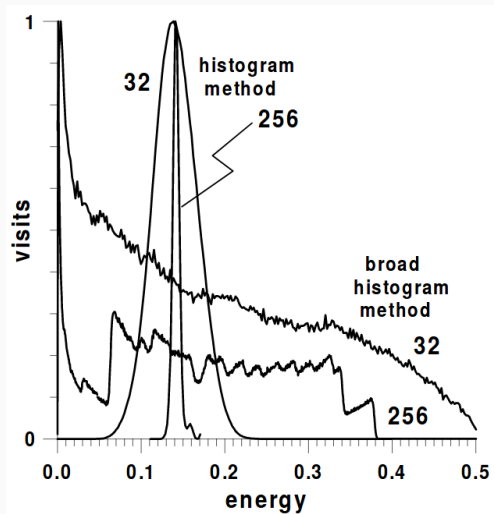
The degeneracy of states contains all the information needed. Using the definition of $g(E)$ yields

$$p_T(E) = g(E) e^{-\frac{E}{k_B T}} = p_{T_0}(E) \exp \left[-\frac{E}{k_B T} + \frac{E}{k_B T_0} \right] \quad (95)$$

and with $f_{T_0,T}(E) = \exp \left[-\frac{E}{k_B T} + \frac{E}{k_B T_0} \right]$ we finally obtain

$$Q(T) = \frac{\sum_E Q(E) p_{T_0}(E) f_{T_0,T}(E)}{\sum_E p_{T_0}(E) f_{T_0,T}(E)}. \quad (96)$$

Histogram Methods



An example of the histogram and the broad histogram method for different system sizes. The figure is taken from Ref. [9].

Broad Histogram Method

Let N_{up} and N_{down} be the numbers of processes which lead to an *increasing* and *decreasing* energy, respectively.

Furthermore, we have to keep in mind that the degeneracy of states increases exponentially with energy E , because the number of possible configurations increases with energy. To explore all energy regions equally, we find a condition equivalent to the one of detailed balance, i.e.,

$$g(E + \Delta E) N_{\text{down}}(E + \Delta E) = g(E) N_{\text{up}}(E). \quad (97)$$

Broad Histogram Method

The motion in phase space towards higher energies can then be penalized with a Metropolis-like dynamics:

- Choose a new configuration,
- if the new energy is lower, accept the move,
- if the new energy is higher then accept with probability $\frac{N_{\text{down}}(E+\Delta E)}{N_{\text{up}}(E)}$.

Broad Histogram Method

We obtain the function $g(E)$ by taking the logarithm of Eq. (97) and divide by ΔE

$$\log [g(E + \Delta E)] - \log [g(E)] = \log [N_{up}(E)] - \log [N_{down}(E + \Delta E)] . \quad (98)$$

Broad Histogram Method

In the limit of small energy differences, we can approximate the last equation by

$$\frac{\partial \log [g(E)]}{\partial E} = \frac{1}{\Delta E} \log \left[\frac{N_{\text{up}}(E)}{N_{\text{down}}(E + \Delta E)} \right] \quad (99)$$

which we can numerically integrate to obtain $g(E)$.

Broad Histogram Method

Distributions of N_{up} and N_{down} can be obtained by keeping track of these numbers for each configuration at a certain energy. In addition, we also need to store the values of the quantity $Q(E)$ we wish to compute as a thermal average according to

$$Q(T) = \frac{\sum_E Q(E) g(E) e^{-\frac{E}{k_B T}}}{\sum_E g(E) e^{-\frac{E}{k_B T}}}. \quad (100)$$

Based on a known degeneracy of states $g(E)$, we can now compute quantities at any temperature.

Flat Histogram Method

Flat histogram method

- Start with $g(E) = 1$ and set $f = e$.
- Make a Monte Carlo update with $p(E) = 1/g(E)$.
- If the attempt is successful at E : $g(E) \leftarrow f \cdot g(E)$.
- Obtain a histogram of energies $H(E)$.
- If $H(E)$ is flat enough, then $f \leftarrow \sqrt{f}$.
- Stop when $f \leq 1 + 10^{-8}$.

Umbrella Sampling

The basic idea is to multiply transition probabilities with a function that is large at the free energy barrier and to later on remove this correction in the averaging step.

$$\tilde{p}(C) = \frac{w(C) e^{-\frac{E(C)}{k_B T}}}{\sum_C w(C) e^{-\frac{E(C)}{k_B T}}} \quad \text{with} \quad \langle A \rangle = \frac{\langle A/w \rangle_w}{\langle 1/w \rangle_w}. \quad (101)$$

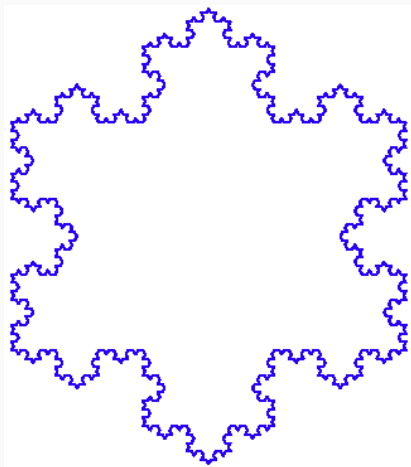
Summary Histogram Methods

Summarizing, some of the most common techniques related to the histogram methods are

- Wang-Landau method [10, 11],
- Multiple histogram method [12],
- Multicanonical Monte Carlo [13],
- Flat Histogram method [14],
- Umbrella sampling [15].

Renormalization Group

Self-similarity



The Koch snowflake as an example of a self-similar pattern.

Real Space Renormalization

H. J. Maris and L. P. Kadanoff describe in *Teaching the renormalization group* that “communicating exciting new developments in physics to undergraduate students is of great importance. There is a natural tendency for students to believe that they are a long way from the frontiers of science where discoveries are still being made.” [16]

Renormalization and Free Energy

To build some intuition for renormalization approaches, we consider a scale transformation of the characteristic length L of our system with that leads to a rescaled characteristic length $\tilde{L} = L/l$. Moreover, we consider the partition function of an Ising system as defined by the Hamiltonian given in Eq. (15). A scale transformation with $\tilde{L} = L/l$ leaves the partition function

$$Z = \sum_{\{\sigma\}} e^{-\beta\mathcal{H}} \quad (102)$$

and the corresponding free energy invariant [17].

Renormalization and Free Energy

We therefore find for the free energy densities

$$f(\epsilon, H) = l^{-d} \tilde{f}(\tilde{\epsilon}, \tilde{H}). \quad (103)$$

We set $\tilde{\epsilon} = l^{y_T} \epsilon$ and $\tilde{H} = l^{y_H} H$ to obtain

$$\tilde{f}(\tilde{\epsilon}, \tilde{H}) = \tilde{f}(l^{y_T} \epsilon, l^{y_H} H). \quad (104)$$

Renormalization and Free Energy

Since renormalization also affects the correlation length

$$\xi \propto |T - T_c|^{-\nu} = |\epsilon|^{-\nu} \quad (105)$$

we can relate the critical exponent ν to y_T . The renormalized correlation length $\tilde{\xi} = \xi/l$ scales as

$$\tilde{\xi} \propto \tilde{\epsilon}^{-\nu}. \quad (106)$$

Renormalization and Free Energy

And due to

$$l^{y_T} \epsilon = \tilde{\epsilon} \propto \tilde{\xi}^{-\frac{1}{\nu}} = \left(\frac{\xi}{l} \right)^{-\frac{1}{\nu}} \propto \epsilon l^{\frac{1}{\nu}}, \quad (107)$$

we find $y_T = 1/\nu$.

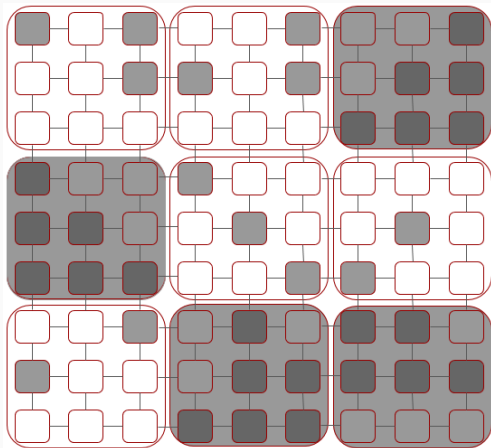
The critical point is a fixed point of the transformation since $\epsilon = 0$ at T_c and ϵ does not change independent of the value of the scaling factor.

Majority Rule

A straightforward example which can be regarded as renormalization of spin systems is the *majority rule*. Instead of considering all spins in a certain neighborhood separately, one just takes the direction of the net magnetization of these regions as new spin value, i.e.,

$$\tilde{\sigma}_i = \text{sign} \left(\sum_{\text{region}} \sigma_i \right). \quad (108)$$

Majority Rule



An illustration of the majority rule renormalization.

Decimation of the One-dimensional Ising Model

The spins only interact with their nearest neighbors and the coupling constant $K = J/(k_B T)$ is the same for all spins.

Decimation of the One-dimensional Ising Model



An example of a one-dimensional Ising chain.

Decimation of the One-dimensional Ising Model

To further analyze this system, we compute its partition function Z and obtain

$$\begin{aligned} Z &= \sum_{\{\sigma\}} e^{K \sum_i \sigma_i} = \sum_{\sigma_{2i}=\pm 1} \prod_{2i} \left[\sum_{\sigma_{2i+1}=\pm 1} e^{K(\sigma_{2i}\sigma_{2i+1} + \sigma_{2i+1}\sigma_{2i+2})} \right] \\ &= \sum_{\sigma_{2i}=\pm 1} \prod_{2i} \{2 \cosh [K (\sigma_{2i} + \sigma_{2i+2})]\} \\ &= \sum_{\sigma_{2i}=\pm 1} \prod_{2i} z(K) e^{\tilde{K} \sigma_{2i} \sigma_{2i+2}} \\ &= [z(K)]^{\frac{N}{2}} \sum_{\sigma_{2i}=\pm 1} \prod_{2i} e^{\tilde{K} \sigma_{2i} \sigma_{2i+2}}, \end{aligned} \tag{109}$$

where we used in the third step that the $\cosh(\cdot)$ function only depends on even spins.

Decimation of the One-dimensional Ising Model

According to Eq. (109), the relation

$$Z(K, N) = [z(K)]^{\frac{N}{2}} Z(\tilde{K}, N/2) \quad (110)$$

holds as a consequence of the decimation method. The function $z(K)$ is the spin-independent part of the partition function and $\tilde{K}$ is the renormalized coupling constant.

Decimation of the One-dimensional Ising Model

We compute the relation

$z(K) e^{\tilde{K} s_{2i} s_{2i+2}} = 2 \cosh [K (s_{2i} + s_{2i+2})]$ explicitly and find

$$z(K) e^{\tilde{K} s_{2i} s_{2i+2}} = \begin{cases} 2 \cosh (2K) & \text{if } s_{2i} = s_{2i+2}, \\ 2 & \text{otherwise .} \end{cases} \quad (111)$$

Decimation of the One-dimensional Ising Model

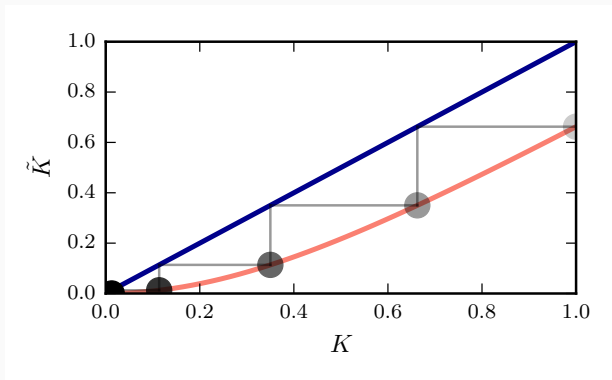
Dividing and multiplying the last two expressions yields

$$e^{2\tilde{K}} = \cosh(2K) \quad \text{and} \quad z^2(K) = 4 \cosh(2K). \quad (112)$$

And the renormalized coupling constant $\tilde{K}$ in terms of K is given by

$$\tilde{K} = \frac{1}{2} \ln [\cosh(2K)]. \quad (113)$$

Decimation of the One-dimensional Ising Model



An illustration of the fixed point iteration defined by Eq. (113).

Decimation of the One-dimensional Ising Model

Given the partition function, we now compute the free energy according to $F = -k_B T N f(K) = -k_B T \ln(Z)$ with $f(K)$ being the free energy density. Taking the logarithm of Eq. (110) yields

$$\ln [Z(K, N)] = N f(K) = \frac{1}{2} N \ln [z(K)] + \frac{1}{2} N f(\tilde{K}). \quad (114)$$

Based on the last equation, we can derive the following recursive relation for the free energy density

$$f(\tilde{K}) = 2f(K) - \ln [2 \cosh (2K)^{1/2}]. \quad (115)$$

Decimation of the One-dimensional Ising Model

There exists one *stable fixed point* at $K^* = 0$ and another *unstable one* at $K^* \rightarrow \infty$. Every fixed point ($K^* = \tilde{K}$) implies that Eq. (115) can be rewritten due to $f(\tilde{K}) = f(K^*)$.

The case of $K^* = 0$ corresponds to the high-temperature limit where the free energy approaches the value

$$F = -Nk_B T f(K^*) = -Nk_B T \ln(2). \quad (116)$$

In this case, the entropy dominates the free energy.

Decimation of the One-dimensional Ising Model

For $K^* \rightarrow \infty$, the system approaches the low temperature limit and the free energy is given by

$$F = -Nk_B T f(K^*) = -Nk_B T K = -NJ, \quad (117)$$

i.e., given by the internal energy.

Generalization

In general, multiple coupling constants are necessary as for example in the case of the two-dimensional Ising model. Thus, we have to construct a renormalized Hamiltonian based on multiple renormalized coupling constants, i.e.,

$$\tilde{H} = \sum_{\alpha=1}^M \tilde{K}_{\alpha} \tilde{O}_{\alpha} \text{ with } \tilde{O}_{\alpha} = \sum_i \prod_{k \in c_{\alpha}} \tilde{\sigma}_{i+k} \quad (118)$$

where c_{α} is the configuration subset over which we renormalize and

$$\tilde{K}_{\alpha} (K_1, \dots, K_M) \quad \text{with} \quad \alpha \in \{1, \dots, M\}. \quad (119)$$

Generalization

At T_c there exists a fixed point $K_\alpha^* = \tilde{K}_\alpha(K_1^*, \dots, K_M^*)$. A possible ansatz to solve this problem is the linearization of the transformation. Thus, we compute the Jacobian

$T_{\alpha,\beta} = \frac{\partial \tilde{K}_\alpha}{\partial K_\beta}$ and obtain

$$\tilde{K}_\alpha - K_\alpha^* = \sum_{\beta} T_{\alpha,\beta}|_{K^*} (K_\beta - K_\beta^*). \quad (120)$$

Generalization

To analyze the behavior of the system close to criticality, we consider eigenvalues $\lambda_1, \dots, \lambda_M$ and eigenvectors $\phi_1, \dots, \phi_M$ of the linearized transformation defined by Eq. (120). The eigenvectors fulfill $\tilde{\phi}_\alpha = \lambda_\alpha \phi_\alpha$ and the fixed point is unstable if $\lambda_\alpha > 1$.

Generalization

The largest eigenvalue dominates the iteration and we can identify the scaling field $\tilde{\epsilon} = l^{y_T} \epsilon$ with the eigenvector of the transformation, and the scaling factor with eigenvalue $\lambda_T = l^{y_T}$. Then, we compute the exponent ν according to

$$\nu = \frac{1}{y_T} = \frac{\ln(l)}{\ln(\lambda_T)}. \quad (121)$$

Monte Carlo Renormalization Group

Since we are dealing with generalized Hamiltonians with many interaction terms, we compute the thermal average of O_α according to

$$\langle O_\alpha \rangle = \frac{\sum_{\{\sigma\}} O_\alpha e^{\sum_\beta K_\beta O_\beta}}{\sum_{\{\sigma\}} e^{\sum_\beta K_\beta O_\beta}} = \frac{\partial F}{\partial K_\alpha} \quad (122)$$

where F is the free energy.

Monte Carlo Renormalization Group

Using the fluctuation-dissipation theorem, we can also numerically compute the response functions

$$\chi_{\alpha,\beta} = \frac{\partial \langle O_\alpha \rangle}{\partial K_\beta} = \langle O_\alpha O_\beta \rangle - \langle O_\alpha \rangle \langle O_\beta \rangle ,$$
$$\tilde{\chi}_{\alpha,\beta} = \frac{\partial \langle \tilde{O}_\alpha \rangle}{\partial K_\beta} = \langle \tilde{O}_\alpha O_\beta \rangle - \langle \tilde{O}_\alpha \rangle \langle O_\beta \rangle .$$

Monte Carlo Renormalization Group

We also find with Eq. (122) that

$$\tilde{\chi}_{\alpha,\beta}^{(n)} = \frac{\partial \langle \tilde{O}_\alpha^{(n)} \rangle}{\partial K_\beta} = \sum_\gamma \frac{\partial \tilde{K}_\gamma}{\partial K_\beta} \frac{\partial \langle \tilde{O}_\alpha^{(n)} \rangle}{\partial K_\gamma} = \sum_\gamma T_{\gamma,\beta} \chi_{\alpha,\gamma}^{(n)}. \quad (123)$$

It is thus possible to derive a value of $T_{\gamma,\beta}$ from the correlation functions by solving a set of M coupled linear equations. At point $K = K^*$, we can apply this method in an iterative manner to compute critical exponents as suggested by Eq. (121).

Monte Carlo Renormalization Group

There are many error sources in this technique, that originate from the fact that we are using a combination of several tricks to obtain our results:

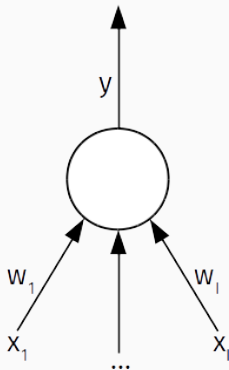
- Statistical errors,
- Truncation of the Hamiltonian to the M^{th} order,
- Finite number of scaling iterations,
- Finite size effects,
- No precise knowledge of K^* .

Boltzmann Machine

Hopfield Network

We begin our excursion to Boltzmann machines with a network consisting of neurons which are fully connected, i.e., every single neuron is connected to all other neurons. A neuron represents a node of a network and is nothing but a function of I different inputs $\{x_i\}_{i \in \{1, \dots, I\}}$ which are weighted by $\{w_i\}_{i \in \{1, \dots, I\}}$ to compute and output y .

Hopfield Network



An illustration of a single neuron with output y , inputs $\{x_i\}_{i \in \{1, \dots, I\}}$ and weights $\{w_i\}_{i \in \{1, \dots, I\}}$.

Hopfield Network

In terms of a Hopfield network, we consider discrete inputs $x_i \in \{-1, 1\}$. The activation of neuron i is given by

$$a_i = \sum_j w_{ij} x_j, \quad (124)$$

where we sum over the inputs. The weights fulfill $w_{ij} = w_{ji}$ and $w_{ii} = 0$.

Hopfield Network

Similarly to the Ising model, the associated energy is given by

$$E = -\frac{1}{2} \sum_{i,j} w_{ij} x_i x_j - \sum_i b_i x_i, \quad (125)$$

where b_i is the bias term.

Hopfield Network

The dynamics of a Hopfield network is given by

$$x_i(a_i) = \begin{cases} 1 & \text{if } a_i \geq 0, \\ -1 & \text{otherwise.} \end{cases} \quad (126)$$

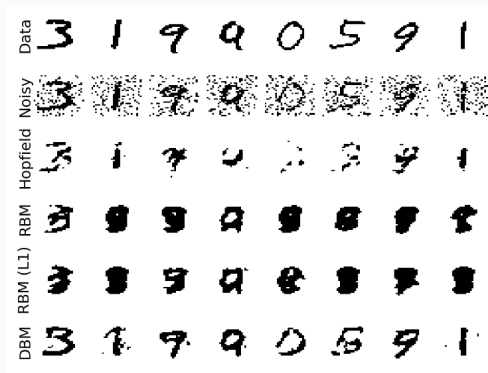
Hopfield Network

The energy difference ΔE_i after neuron i has been updated is

$$\Delta E_i = E(x_i = -1) - E(x_i = 1) = 2 \left(b_i + \sum_j w_{ij} x_j \right). \quad (127)$$

We can absorb the bias b_i in the sum by having an extra active unit at every node in the network. We thus showed that the activation defined by Eq. (124) equals one half of the energy difference ΔE_i .

Boltzmann Machine Learning



A comparison of denoising capabilities of Hopfield and Boltzmann machine models. The figure is taken from <https://arxiv.org/pdf/1803.08823.pdf>.

Boltzmann Machine Learning

For some applications, finding a local minimum based on the deterministic update rule defined by Eq. (126) might not be sufficient. Similar to the discussion of Monte Carlo methods for Ising systems, we employ an update probability

$$p_i = \frac{1}{1 + \exp(-\Delta E_i/T)} = \sigma(2a_i/T) \quad (128)$$

to set neuron i to unity independent of its state [18]. Here, $\sigma(x) = 1/[1 + \exp(-x)]$ denotes the sigmoid function.

Boltzmann Machine Learning

As defined in Eq. (127), the energy difference ΔE_i is the gap between a configuration with an active neuron i and an inactive one. The parameter T acts as temperature equivalent⁵.

A closer look at Eqs. (125) and (128) tells us that we are simulating a Hamiltonian system with Glauber dynamics. Due to the fulfilled detailed balance condition, we reach thermal equilibrium and find again for the probabilities of the system to be in state X or Y ⁶

$$\frac{p_{\text{eq}}(Y)}{p_{\text{eq}}(X)} = \exp \left(-\frac{E(Y) - E(X)}{T} \right). \quad (129)$$

⁵For $T \rightarrow 0$, we recover deterministic dynamics as described by Eq. (126).

⁶Here we set $k_B = 1$.

Boltzmann Machine Learning

We divide the Boltzmann machine units into *visible* and *hidden units* represented by the non-empty set V and the possibly empty set H , respectively. The visible units are set by the environment whereas the hidden units are additional variables which might be necessary to model certain outputs.

Boltzmann Machine Learning

Let $P'(\nu)$ be the probability distribution over the visible units ν in a freely running network. It can be obtained by marginalizing over the corresponding joint probability distribution, i.e.,

$$P'(\nu) = \sum_h P'(\nu, h), \quad (130)$$

where h represents a hidden unit.

Boltzmann Machine Learning

The goal is to come up with a method such that $P'(\nu)$ approaches the unknown environment distribution $P(\nu)$. We measure the difference between $P'(\nu)$ and $P(\nu)$ in terms of the Kullback-Leibler divergence (relative entropy)

$$G = \sum_{\nu \in V} P(\nu) \ln \left[\frac{P(\nu)}{P'(\nu)} \right]. \quad (131)$$

Boltzmann Machine Learning

To minimize G , we perform a gradient descent according to

$$\frac{\partial G}{\partial w_{ij}} = -\frac{1}{T} (p_{ij} - p'_{ij}) , \quad (132)$$

where p_{ij} is the probability that two units are active on average if the environment is determining the states of the visible units and p'_{ij} is the corresponding probability in a freely running network without a coupling to the environment.

Boltzmann Machine Learning

Both probabilities are measured at thermal equilibrium. In the literature, the probabilities p_{ij} and p'_{ij} are also often defined in terms of the thermal averages $\langle x_i x_j \rangle_{\text{data}}$ and $\langle x_i x_j \rangle_{\text{model}}$, respectively. The weights w_{ij} of the network are then updated according to

$$\Delta w_{ij} = \epsilon (p_{ij} - p'_{ij}) = \epsilon (\langle x_i x_j \rangle_{\text{data}} - \langle x_i x_j \rangle_{\text{model}}), \quad (133)$$

where ϵ is the learning rate.

Boltzmann Machine Learning

The so-called restricted Boltzmann machine is not taking into account mutual connections within the set of hidden and visible units, and turned out to be more suitable for learning tasks. In the case of restricted Boltzmann machines, the weight update is given by

$$\Delta w_{ij} = \epsilon (\langle \nu_i h_j \rangle_{\text{data}} - \langle \nu_i h_j \rangle_{\text{model}}), \quad (134)$$

where ν_i and h_j represent visible and hidden units, respectively.

Boltzmann Machine Learning

Instead of sampling the configurations for computing $\langle \nu_i h_j \rangle_{\text{data}}$ and $\langle \nu_i h_j \rangle_{\text{model}}$ at thermal equilibrium, we could also just consider a few relaxation steps. This method is called *contrastive divergence* and defined by the following update rule

$$\Delta w_{ij}^{\text{CD}} = \epsilon \left(\langle \nu_i h_j \rangle_{\text{data}}^0 - \langle \nu_i h_j \rangle_{\text{model}}^k \right), \quad (135)$$

where the superscript indices indicate the number of updates.

Parallelization

Parallelization

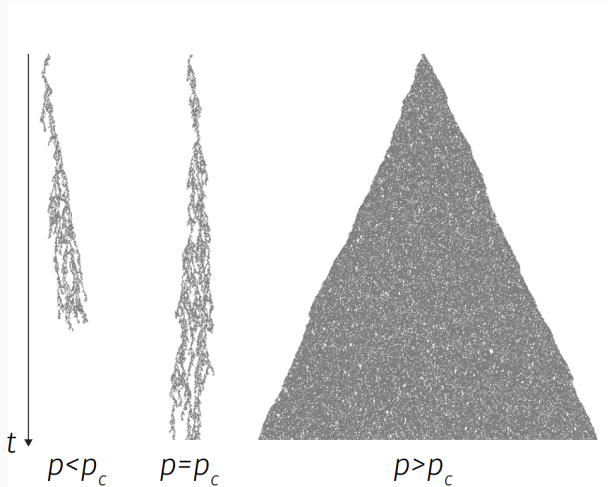
Some important concepts are summarized in the lecture notes.

Non-Equilibrium Systems

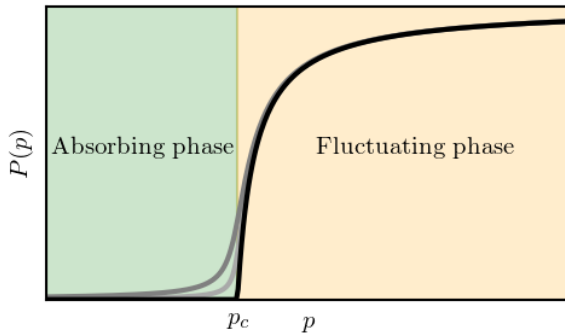
Directed Percolation

Isotropic percolation is an equilibrium process in which a site or bond is occupied with a certain probability p . However, in the case of directed percolation an additional constraint is defined, namely in each time step occupied sites or bonds occupy their neighbors with probability p only along a given direction [19]. Time is therefore a relevant parameter.

Directed Percolation



Directed Percolation



Directed Percolation

In the vicinity of p_c the percolation order parameter $P(p)$, the fraction of sites of the largest cluster (isotropic percolation) or the density of active sites (directed percolation), takes the form: $P(p) \propto (p - p_c)^\beta$.

Although, the free energy is not defined in non-equilibrium processes one can distinguish between first-order and continuous phase transitions based on the behavior of the order parameter.

Directed Percolation

Another formulation of directed percolation is given by the *contact process*. On a lattice with active ($s_i(t) = 1$) and inactive ($s_i(t) = 0$) sites, we sequentially update the system according to the following dynamics:

Let the number of active neighbors neighbors be $n_i(t) = \sum_{\langle i,j \rangle} s_j(t)$ and a new value $s_i(t + dt) \in \{0, 1\}$ is obtained according to the transition rates

$$w[0 \rightarrow 1, n] = (\lambda n)/(2d) \quad \text{and} \quad w[1 \rightarrow 0, n] = 1, \quad (136)$$

where d is the dimension of the system.

Directed Percolation

The critical infection rate λ_c and the critical exponent β on a square lattice for different dimensions d . At and above the critical dimension $d_c = 4$ the critical exponents are described by mean-field theory [17]. The values are rounded to the second decimal and taken from Refs. [20] and [21].

	$d = 1$	$d = 2$	$d = 3$	$d_c = 4$
λ_c	3.30	1.65	1.32	1.20
β	0.28	0.59	0.78	1

Directed Percolation

The contact process has the same characteristic critical exponents as directed percolation what implies that they both belong to the same universality class. According to Ref. [17] the directed percolation universality class requires:

1. a continuous phase transition from a fluctuating phase to a unique absorbing state,
2. a transition with a one-component order parameter,
3. local process dynamics,
4. no additional attributes, such as conservation laws or special symmetries.

Kinetic Monte Carlo (Gillespie Algorithm)

Non-equilibrium systems are not described by a Hamiltonian and we cannot use the same algorithms which we employed for studying the Ising model or other equilibrium models. Time is now a physical parameter and not a mere virtual property of a Markov chain.

A standard algorithm for simulating non-equilibrium dynamics is the *kinetic Monte Carlo* method which is also often referred to as *Gillespie algorithm*.

Kinetic Monte Carlo (Gillespie Algorithm)

To give an example, the kinetic Monte Carlo algorithm is applied to the contact process on a two-dimensional square lattice. Recovery and activation define $n = 2$ processes with corresponding rates $R_1 = 1$ and $R_2 = \lambda$.

Kinetic Monte Carlo (Gillespie Algorithm)

At time t , the subpopulation $N_1(t)$ consists of all active nodes which recover with rate R_1 . The total recovery rate is then given by $Q_1(t) = N_1(t)$.

On the square lattice only nearest-neighbor interactions are considered and the total rate of the second process (activation) is obtained by computing n_i and $w [0 \rightarrow 1, n_i]$ according to Eq. (136) for all nodes.

Kinetic Monte Carlo (Gillespie Algorithm)

Kinetic Monte Carlo algorithm

1. Identify all individual rates (per node) $\{R_i\}_{i \in \{1, \dots, n\}}$.
2. Determine the overall rates (all nodes) $\{Q_i\}_{i \in \{1, \dots, n\}}$, where $Q_i = R_i N_i$ and N_i defines the subpopulation which corresponds to R_i . It is possible that some subpopulations are identical but correspond to different rates.
3. Let $\eta \in [0, 1)$ be a uniformly distributed random number and $Q = \sum_i Q_i$. The process with rate R_j occurs if $\sum_{i=1}^{j-1} Q_i \leq \eta Q < \sum_{i=1}^j Q_i$.
4. Let $\epsilon \in [0, 1)$ be a uniformly distributed random number. Then the time evolves as $t \rightarrow t + \Delta t$, where $\Delta t = -Q^{-1} \log(1 - \epsilon)$ and return to step 2.

- [1] Stanley, H.
Scaling, universality, and renormalization: Three pillars of modern critical phenomena (1999).
URL http://link.springer.com/chapter/10.1007%2F978-1-4612-1512-7_39.
- [2] Stanley, H.
***Introduction to Phase Transitions and Critical Phenomena* (Clarendon, Oxford, 1971).**
- [3] Sengers, J., Sengers, J. L. & Croxton, C.
Progress in liquid physics.
Wiley, Chichester (1978).

- [4] Pelissetto, A. & Vicari, E.
Critical phenomena and renormalization-group theory.
Phys. Rep. **368**, 549–727 (2002).
- [5] Vichniac, G. Y.
Simulating physics with cellular automata.
Phys. D **10**, 96–116 (1984).
- [6] Binder, K. & Heermann, D.
Monte Carlo Simulation in Statistical Physics (Springer, 1997).
- [7] Binder, K. & Landau, D.
Finite-size scaling at first-order phase transitions.
Phys. Rev. B **30**, 1477 (1984).

- [8] Salzburg, Z., Jacobson, J., Fickett, W. & Wood, W.
Excitation of non-radial stellar oscillations by gravitational waves: a first model.
Chem. Phys. 65 (1959).
- [9] de Oliveira, P., Penna, T. & Herrmann, H.
Broad histogram method.
Braz. J. Phys. 677 (1996).
- [10] Wang, F. & Landau, D.
Efficient, multiple-range random walk algorithm to calculate the density of states.
Phys. Rev. Lett. **86**, 2050 (2001).

- [11] Zhou, C., Schulthess, T. C., Torbrügge, S. & Landau, D.
Wang-landau algorithm for continuous models and joint density of states.
Phys. Rev. Lett. **96**, 120201 (2006).
- [12] Ferrenberg, A.
Am ferrenberg and rh swendsen, phys. rev. lett. 63, 1195 (1989).
Phys. Rev. Lett. **63**, 1195 (1989).
- [13] Berg, B. A.
Introduction to multicanonical monte carlo simulations.
Fields Inst. Commun. **26**, 1–24 (2000).

- [14] Wang, F. & Landau, D.
Determining the density of states for classical statistical models: A random walk algorithm to produce a flat histogram.
Phys. Rev. E **64**, 056101 (2001).
- [15] Torrie, G. & Valleau, J.
Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling.
J. Comp. Phys. **23**, 187–199 (1977).
URL <http://www.sciencedirect.com/science/article/pii/0021999177901218>.
- [16] Maris, H. J. & Kadanoff, L. P.
Teaching the renormalization group.
Am. J. Phys. **46**, 652–657 (1978).

- [17] Henkel, M., Hinrichsen, H. & Lübeck, S.
Non-Equilibrium Phase Transitions Volume 1 – Absorbing Phase Transitions (Springer Science & Business Media, 2008).
- [18] Ackley, D. H., Hinton, G. E. & Sejnowski, T. J.
A learning algorithm for boltzmann machines.
Cogn. Sci. **9**, 147–169 (1985).
- [19] Broadbent, S. & Hammersley, J.
Percolation processes I. crystals and mazes.
Proceedings of the Cambridge Philosophical Society (1957).

- [20] Moreira, A. G. & Dickman, R.
Critical dynamics of the contact process with quenched disorder.
Phys. Rev. E (1996).
- [21] Sabag, M. M. S. & de Oliveira, M. J.
Conserved contact process in one to five dimensions.
Phys. Rev. E (2002).