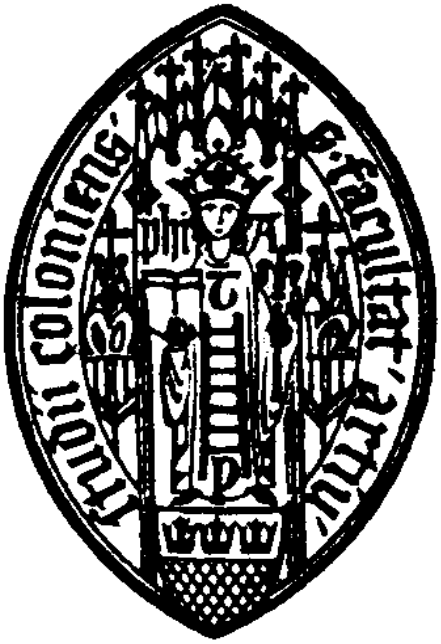
# Advanced Statistical Physics



**Lecture:** Prof. Dr. Johannes Berg
**Notes:** Gerold Busch, BSc

Wintersemester 2011/12

Institut für Theoretische Physik
Universität zu Köln

# Contents

# Preface

We proudly present the lecture notes of the master course "Advanced Statistical Physics". The notes follow the lecture given by Johannes Berg in the winter term 2011/12 at the University of Cologne.

Of course, there will be some mistakes in these notes. So read carefully and with caution. If you find something wrong or have questions, tell me (busch@ph1.uni-koeln.de)!

Cologne, February 2012
Gerold Busch

# 1 Stochastic Processes and the Approach to Equilibrium

Consider a particle moving in a gas (Brownian motion). The collisions with gas molecules can be modelled as random events, pushing the particle in random directions. By definition, there is no way of predicting the trajectory of a particle. One might think that the search for deterministic laws of motion is doomed in this case.

However, at least in principle we can repeat the experiment of suspending a particle in a gas many times over, and observe the *statistics* over many trajectories. Examples are the average position at a given moment in time, the variance, and so on. These statistics all stem from a probability density of finding the particle in a given position (and momentum, and other variables needed to describe our particle). We can ask if this probability evolves over time according to a deterministic law of motion.

One convenient way of doing this, which is easy to visualize, is to suspend many non-interacting particles at once, like a blob of ink in a glass of water. Here the variance of the particles' displacement, or the variance of the probability distribution function of a single particle, of a single particle will evolve deterministically with time (according to the law of diffusion).

This chapter will introduce the notion of a deterministic evolution of a probability density. A particular question will be how a system approaches a steady state at long times, when the statistics no longer changes (and our ink has ended up being uniformly distributed in a glass of water.) Before studying a general approach, we will look at a simple model.

## 1.1 The Ehrenfest Model

A very instructive model to investigate the approach to the steady state is the Ehrenfest model (1907). Consider two urns, termed "R" and "L", and $N$ balls, which can be in either urn. Start with all balls in one of the urns, say L. At each time step, one of the $N$ balls is picked at random (with equal probability) and moved to the other urn.

After a sufficiently long time, we observe that the number of balls in each urn has equalized. In this **steady state**, on average, each urn contains $N/2$ balls, and fluctuations around this average are of order $\sqrt{N}$. Observing the system in this state, it is impossible to tell whether initially all urns were in the left, or the right urn, or whether the system was started with an equal number of balls in either urn.

Moreover, making a movie of this process in the steady state, and playing the movie either forwards or backwards, it would be impossible to tell which direction was being shown. This is different during the approach to the steady state, where when running the movie backwards it would appear the model has a different rule which favours picking balls in the right urn and moving them to the left.

### 1.1.1 Microscopic Description

We label the balls $b = 1, 2, 3, ...N$ and assign a variable $s_b = \pm 1$ to each of them, $s_b = +1$ indicating ball $b$ sitting in the right urn, and $-1$ in the left urn. Thus, the configuration of the system is described by a $N$-dimensional vector $\mathbf{s}$. At any moment in time, a given configuration occurs with probability $p(\mathbf{s})$. The initial configuration has $p(-1, -1, \ldots) = 1$, and zero for all the other configurations. Our

aim is to find a deterministic law of motion for $p(\mathbf{s}, t)$ over time. The limit $\lim_{t\to\infty} p(\mathbf{s}, t)$ describes the steady state (if it exists).

### 1.1.2 Deterministic Description

Before this, we ask how the centre of mass of the system changes with time. If at a given time there are $i$ balls in the right-hand urn, with probability $i/N$ the move will take a ball from right to left (decreasing $i$ by one), with probability $1-i/N$ from left to right (increasing $i$ by one). The expected *net change* of $i$ is $-i/N + (1 - i/N) = 1 - 2i/N$. Denoting the expected number of balls in the right urn by $y(t)$ we have

$$\partial_t y = P(i+1|i) - P(i-1|i) = 1 - \frac{2y}{N} \,. \tag{1.1}$$

The more particles are in the left urn, the higher the probability that a particle in the left urn will be picked at the next step. This equation of motion is solved by

$$y(t) = \frac{N}{2} + \left(y(0) - \frac{N}{2}\right) \cdot e^{-\frac{2y}{N}} \tag{1.2}$$

We see that for $t \to \infty$, the second term vanishes and the steady state of $y(t) = N/2$ is reached, regardless of which initial configuration one begins in. The rate at which the expected value of $i$ approaches that of the steady state is also $N/2$. Figure 1.1 shows a simulation of the Ehrenfest model.

### 1.1.3 Macroscopic Description

Our experience with the deterministic approach suggest not monitor the position each ball individually, but to look at one macroscopic variable only, the number $I$ of balls in the right urn. This variable may on the possible values $i = 0, 1, \ldots N$. The event that a ball from the right urn is picked and

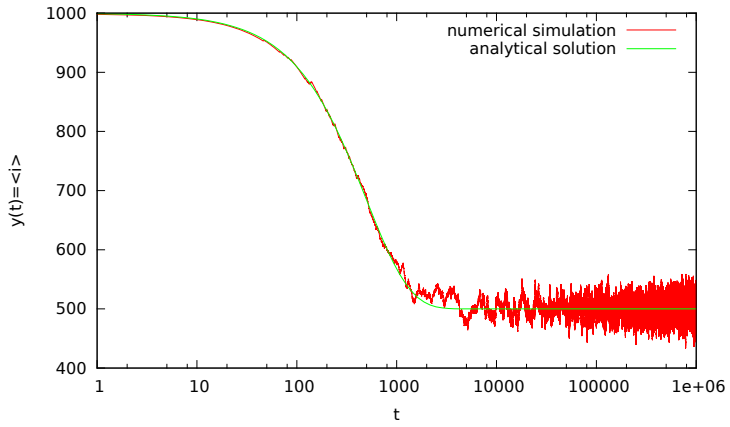



Figure 1.1: A single run of the Ehrenfest model with $N = 1000$ starting all balls in the left urn. The number of balls in the left urn is plotted against the number of moves taken.In the steady state the particle number in the left urn fluctuates around $N/2$. The analytical solution is given by equation 1.2. The apparent increase of the fluctuations at large times is an effect of the logarithmic scale on the time axis, since in larger time intervals larger deviations from the mean appear.

put into the left corresponds to $I = i$ at time $t$ changing to $I = i - 1$ at time $t + 1$. This event occurs with probability denoted

$$P(i-1, t+1|i, t) = \frac{i}{N}$$

the opposite event has

$$P(i+1, t+1|i, t) = 1 - \frac{i}{N}.$$

These are conditional probabilities of finding the system in a certain configuration, given the configuration at the previous time.

The probability $p(i, t)$ of finding variable $I$ in the configuration $i$ at time $t$ changes with time. Since at each step, $I$ either decreases or increases by one with probabilities specified above, $p(i, t+1)$ depends only on $p(i+1, t)$ and $p(i-1, t)$, giving

$$\begin{aligned} p(i, t+1) &= P(i+1, t+1|i, t)p(i+1, t) + P(i-1, t+1|i, t)p(i-1, t) && (1.3) \\ &= \sum_j W(i|j)p(j, t) && (1.4) \\ &= \mathbf{W} \cdot \mathbf{p}(t) \ . && (1.5) \end{aligned}$$

This equation defines the so-called **stochastic matrix** $(\mathbf{W})_{ij} \equiv W(i|j)$ of the system, and we defined a vector $\mathbf{p}_i(t) = p(i, t)$. In the case of the Ehrenfest model, it is

$$\mathbf{W} = \frac{1}{N} \begin{pmatrix} 0 & 1 & 0 & \cdots \\ N & 0 & 2 & \cdots \\ 0 & N-1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

We can now write $\mathbf{p}(t = 1) = \mathbf{W}\mathbf{p}(0)$, iterating gives

$$\mathbf{p}(t) = (\mathbf{W})^t \, \mathbf{p}(0) \tag{1.6}$$

for all $t$. $\mathbf{p}(t+1) = \mathbf{W}\mathbf{p}(t)$ is the sought-after deterministic equation of motion of the system, (1.6) is its solution.

For any system, the stochastic matrix $\mathbf{W}$ must have the following properties:

- $W(i|j) \geq 0 \ \forall \ i, j$ since the entries of the matrix are the probabilities of transitions from configuration $i$ to $j$.
- $\sum_i W(i|j) = 1$ i.e. with probability the system has to take on some configuration at the next step.

From these two conditions, the following two properties follow:

- All left eigenvalues of $\underline{\underline{W}}$ have the property $|\lambda^L| \leq 1$. [1]

[1] Proof:

$$\begin{aligned} |\lambda^L e_j^L| &= \left| \sum_i e_i^L W_{ij} \right| \\ \left| \sum_i e_i^L W_{ij} \right| &\leq \max_l \left| e_l^L \right| \sum_i W_{ij} \\ \Rightarrow \left| \lambda^L e_j^L \right| &\leq \max_l \left| e_l^L \right| \\ \Rightarrow \left| \lambda^L \right| &\leq 1. \end{aligned}$$

- Any stochastic matrix has a left eigenvector with $\lambda^1 = 1$ such that $e_i^{1L} = 1 \forall i$. [2]

These two properties are the key to the behaviour of the system at long times. Since in general, $\mathbf{W}$ will not be a symmetric matrix, we obtain the spectral decomposition

$$W_{ij} = \sum_k \lambda_k e_i^{kR} e_j^{kL} \tag{1.7}$$

in terms of left and right eigenvectors of $\mathbf{W}$. For now, we assume that eigenvalues $\lambda$ are not degenerate and use the orthogonality of eigenvectors $\underline{e}^{k_m L} \cdot \underline{e}^{k_n R} = \delta_{k_m k_n}$, we get

$$(W_{ij})^n = \sum_k \lambda_k^n e_i^{kR} e_j^{kL} \tag{1.8}$$

Thus, from the solution (1.6) for the probabilities $p_i(t)$, we have

$$p_i(t) = \sum_{j,k} \lambda_k^t e_i^{kR} e_j^{kL} p_j(0) \tag{1.9}$$

As $t \to \infty$, the contribution from eigenvalues different from $\pm 1$ in the sum over $k$ will go to $0$. Assuming non degeneracy and labelling, without loss of generality, the eigenvalue with value $\lambda = 1$ by $k = 1$, we obtain

$$\lim_{n\to\infty} p_i(n) = e_i^{1R} \sum_j e_j^{1L} p_j(0) = e_i^{1R} \tag{1.10}$$

as $e_i^{1L} = 1 \forall i$. We see that in the long term, all dependence on initial conditions is lost and the system goes to an steady state defined by the right eigenvector with eigenvalue $1$. Note that this approach not only generates the expected number of $i$ in the steady state, but the full distribution of $p_i$, that is all moments of the fluctuations around the expected value.

### 1.1.4 Ergodicity Breaking

Our derivation assumed a non-degenerate eigenvalue $\lambda = 1$. We briefly discuss two ways in which this condition can be violated. One simple scenario is to combine two stochastic matrices in a block-diagonal structure, such as

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

This stochastic matrix defines transitions between configurations that toggle between $1 \leftrightarrow 2, 3 \leftrightarrow 4$. If the system is started in configuration $1$, or a linear combination of configurations $1, 2$, it will never be found in configurations $3$ or $4$ (all powers of $\mathbf{W}$ are also block-diagonal).

A somewhat different scenario can occur in large systems. The stochastic matrix is exponentially large in the number of degrees of freedom (since the number of configurations grows exponentially with the system size). It is possible that the second largest eigenvalues $\lambda_2 \to 1$ as $N \to \infty$. Any *finite* system will then reach a steady state on a timescale of $1/\log(\lambda_1/\lambda_2)$, which diverges in the thermodynamic limit. Thus in a macroscopically large system, a dependence of $\mathbf{p}(t)$ on the initial state can remain for a very long time; an example is hysteresis.

[2] Proof: Assume $e_i^L = 1 \forall i$. Then

$$\sum_i e_i^L W_{ij} = \sum_i W_{ij} = 1 \times e_i^L$$

## 1.2 Time Reversal and Detailed Balance

A key simplification of the steady state occurs for systems which are said to be time-reversal invariant.

**Time Reversal:** A process is said to be reversible if for any sequence of configurations $i = i_0, i_1, i_2, i_3, ..., i_n$, the transition rates are such that

$$W(i_n|i_{n-1}) \cdot W(i_{n-1}|i_{n-2}) \cdots W(i_1|i_0) \cdot p_{\text{steady}}(i_0) = W(i_0|i_1) \cdots W(i_{n-1}|i_n) \cdot p_{\text{steady}}(i_n) \quad (1.11)$$

where $p_{\text{steady}}(i_0)$ and $p_{\text{steady}}(i_1)$ are both the steady-state distributions achieved after long times. Note that there are alternative, equivalent definitions, for instance that the probability of any circular trajectory in configuration space be independent of the direction the circle is traversed in.

Specifically, for two configurations $i, j$ one has

$$W(j|i)p_{\text{steady}}(i) = W(i|j)p_{\text{steady}}(j) \ . \quad (1.12)$$

This equation has a simple interpretation in terms of the probability flows. In the steady state, the probability of finding the system in configuration $i$ at a given, and making a transition to configuration $j$, equals the probability of the reverse process. Visualizing the probability density as a particle density, one speaks of a probability flow (in the space of configurations of the system). Equation (1.12) is called **detailed balance**, because probability flows between pairs of configurations are equal and opposite.

A steady state obeying detailed balance is called an **equilibrium state**. In a **non-equilibrium steady state** the are non-zero probability currents between pairs of configurations, which cancel each other at each configuration $i$ (like non-compressible fluid float having zero divergence everywhere but a finite circulation).

Detailed balance allows us to construct the equilibrium solution directly from the stochastic matrix. Let $p_{\text{eq}}(1) = r$ be some number. Then, the pairwise conditions of detailed balance mean that we can determine every

$$p_{\text{eq}}(k) = \frac{W(k|1)}{W(1|k)} \cdot r$$

and $r$ is determined by the normalization condition $\sum_k p_{\text{eq}}(k) = 1$. For a dynamics obeying time reversal invariance (1.11) ensures that this solution is the same regardless of which configuration we start with.

## 1.3 Stochastic Processes

We are now prepared for a broader development, leading to the equation of for a general stochastic process.

### 1.3.1 Definitions

A **stochastic variable** $X$ is defined by a set of possible values and a probability distribution over this set. A particular realisation of $X$ is denoted $X = x$. Depending on the system, $X$ may comprise many positions, momenta, spins configurations, etc. of all the components of our system.

Formally, a **stochastic process** $Y$ is defined as some function of a stochastic variable $X$

$$Y_X(t) = f(X, t) \ .$$

In many applications, the parameter $t$ the function $f(X, t)$ depends on will be time.

Statistically, a stochastic process is described completely by a probability or probability density

$$P(...; x_3, t_3; x_2, t_2; x_1, t_1) \ . \tag{1.13}$$

If $t$ denotes time, then this function specifies the probability of an entire trajectory, with the system in configuration $s_1$ at time $t_1$, in configuration $x_2$ at $t_2$, and so on. It turns out that many stochastic processes allow an important simplification, which allows to work with $P(x_{i+1}, t_{i+1}|x_i, t_i)$ (see (1.1.3)) instead of the full (1.13).

### 1.3.2 Markov process

Think of a time $t_1$ and configuration $x_1$ as "now", times $t < t_1$ as "past" and $t > t_1$ as "future". If the condition

$$P(...; x_3, t_3; x_2, t_2; x_1, t_1; x_0, t_0, x_{-1}, t_{-1}; ...) = P(...; x_3, t_3; x_2, t_2; x_1, t_1) \tag{1.14}$$

holds for all times and configuration, the corresponding stochastic process is called a **Markov Process**. Physically, this condition can be interpreted as a rapid memory decay: The present configuration $x_1$ at $t_1$ is fully sufficient to determine the statistics of the future trajectory. The Ehrenfest model specifies a Markov process, and many physical systems can be modeled at least approximated by Markov processes.

Mathematically, the Markov property is an extremely powerful restriction as

$$\begin{aligned} P(x_2, t_2; x_1, t_1|x_0, t_0) &= P(x_2, t_2|x_1, t_1; x_0, t_0)P(x_1, t_1|x_0, t_0) \\ &\overset{\text{Markov}}{=} P(x_2, t_2|x_1, t_1)P(x_1, t_1|x_0, t_0) \ . \end{aligned}$$

The joint probability of $x_2$ and $x_1$ can thus be written as the product of two conditional probabilities.

Integration over $x_1$ gives:

$$P(x_2, t_2|x_0, t_0) = \int \mathrm{d}x_1 P(x_2, t_2|x_1, t_1)P(x_1, t_1|x_0, t_0) \tag{1.15}$$

This consistency equation for two-timepoint conditional probabilities $P(x, t|x', t')$ is called the **Chapman-Kolmogorov-equation**. These conditional probabilities, specifying the probability of finding the system in state $x$ at time $t$, given that at time $t'$ the system was in state $x'$, specify the transition rates from $x'$ to $x$ (through the derivative wrt to $t$). This conditional probability is also termed the **propagator** in analogy with quantum field theory. For a Markov process, they turn out to be sufficient to determine the equation of motion for the probability function.

## 1.4 The master equation in continuous time

Consider a Markov process with

$$P(x_1, t + \tau|x_0, t) = [1 - \tau a(x_0, t)]\, \delta(x_1 - x_0) + \tau W(x_1, x_0, t) + \mathcal{O}(\tau^2) \ . \tag{1.16}$$

We consider the variable $x$ as continuous, so (1.16) describes a particle either jumping from $x_0$ to $x_1$ during the interval $t_1 - t_0$, or remaining at $x_0$.

From

$$\int \mathrm{d}x_1 P(x_1, t + \tau|x_0, t) = 1$$

(the system has to take some state at time $t + \tau$), it follows that

$$a(x_0, t) = \int \mathrm{d}x_1 W(x_1, x_0, t).$$

Thus, $a(x_0, t)$ is the probability of jumping from $x_0$ anywhere.

We plug this into the Chapman-Kolmogorov-equation and demand consistency at least in first order of $\tau$.

$$\begin{aligned} p(x, t+\tau) &= \int \mathrm{d}x' P(x, t+\tau|x', t) p(x', t) \\ &= [1 - \tau a(x,t)]\, p(x,t) + \tau \int \mathrm{d}x' W(x, x', t) p(x', t) + \mathcal{O}(\tau^2) \end{aligned}$$

Comparison with the Taylor Expansion $p(x, t+\tau) = p(x,t) + \tau \partial_t p(x,t) + \mathcal{O}(\tau^2)$ yields:

$$\partial_t p(x,t) = \int \mathrm{d}x' W(x, x', t) p(x', t) - \int \mathrm{d}x' W(x', x, t) p(x, t) \tag{1.17}$$

This is the **master equation**. Consider the first integral: $P(x', t)$ is the probability for being in configuration $x'$. $w(x, x', t)$ gives the transition rate from $x'$ to $x$. Integration over $x'$ gives the probability flow into configuration $x$. Similarly, the second integral gives probability flow moving out of $x$ and into some configuration $x'$. The rate of change of $p(x,t)$ is thus the balance between probability flows into and out of $x$. The master equation can be considered as the equation of continuity for an incompressible fluid. In general, only an integral form exists, since transitions are not restricted to states $x'$ close in some sense to $x$. For the particular case of transition rates restricting transition to small jumps, one can derive a diffusion equation for $p(x,t)$ called the Fokker-Planck equation.

Conservation of probability yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \int \mathrm{d}x\, p(x,t) = \int \mathrm{d}x \frac{\partial}{\partial t} p(x,t) = 0 \tag{1.18}$$

For a discrete configuration space, the master equation can be derived analogously, giving

$$\partial_t p_i(t) = \sum_{i'} w_{ii'} p_{i'}(t) - \sum_{i'} w_{i'i} p_i(t) = \sum_i W_{ii'} p_{i'} \tag{1.19}$$

with

$$W_{ii'} = \begin{cases} w_{ii'} & \text{for } i \neq i' \\ -\sum_{i' \neq i} w_{i'i} & \text{for } i = i' \end{cases} \tag{1.20}$$

and the same interpretation in terms of probability flows. Detailed balance (1.12) ensures that in equilibrium the probabilities $\mathbf{p}$ no longer change. The formal solution is of the discrete-states master equation is

$$\underline{p}(t) = \exp(\underline{\underline{W}} t) \underline{p}(0)\ , \tag{1.21}$$

corresponding to (1.9) for the Ehrenfest model.

## 1.5 Physical systems, free energy and the approach to equilibrium

Now we are prepared to address the question of if and how a physical system, described by a Markov process and obeying time-reversal invariance, approaches its equlibrium. Consider the following scenario:

A system is characterized by a configuration denoted $i$ (discrete configuration space, continuous systems would be discretised by making a very fine mesh imposed on the continuous coordinates). The configuration $i$ comprises the entire microscopic information of the system, e.g. positions and momenta of all molecules of a gas.

Each configuration $i$ is characterized by an energy $E_i$, e.g. total kinetic and potential energy of the molecules. Furthermore, the system is connected to a heat bath (not further specified) and exchanges energy with the heat bath.

Stochastic transitions between different configurations of the system depend only on the *change in energy* $E_j - E_i$ $(j \to i)$

$$W(i|j) = f(E_i - E_j) \ .$$

Strictly speaking, this is a property of the heat bath, not the system.

We now define a function $g$ as follows

$$\frac{w(i|j)}{w(j|i)} = \frac{f(E_i - E_j)}{f(E_j - E_i)} = g(E_i - E_j) \ ,$$

and we further make the assumption that the system obeys time reversal invariance

$$\frac{w(i|j)}{w(j|i)} = \frac{p_{\mathsf{eq}}(i)}{p_{\mathsf{eq}}(j)} = g(E_i - E_j) \tag{1.22}$$

This condition can only be fulfilled if $p_{\mathsf{eq}}$ is the exponential function:

$$p_{\mathsf{eq}}(i) = \frac{1}{Z_\beta} e^{-\beta E_i} \tag{1.23}$$

with $Z_\beta = \sum_i e^{-\beta E_i}$. If we identify $\beta = 1/k_B T$, we have derived the **Boltzmann measure** as the equilibrium measure of our system. (In the following we will always use units such that $k_B = 1$.

The big question now is: Starting in starting condition $p(i, t = 0) = p_i(t = 0) = p_i^0$. Does $p_i(t)$ evolve towards the equilibrium probability distribution $p_i^{\mathsf{eq}}$?

We define a function

$$F(t) = \sum_i p_i(t) E_i + T \sum_i p_i(t) \ln p_i(t) \ , \tag{1.24}$$

which depends on time through the probability function $\mathbf{p}(t)$. $T = 1/\beta$. $\sum_i p_i(t) E_i$ is the expected energy of the distribution, $\sum_i p_i(t) \ln p_i(t)$ is the corresponding entropy. We call $F(t)$ the **time dependent free energy**.

The equilibrium free energy for the Boltzmann distribution $p_i = 1/Z \exp(-\beta E_i)$ is given by:

$$F_{\mathsf{eq}} = \frac{1}{Z} \sum_i e^{-\beta E_i} E_i + T \frac{1}{Z} \sum_i e^{-\beta E_i} \left(-\beta E_i - \ln Z\right) = -T \ln Z \tag{1.25}$$

We will now show that $F(t)$ decreases monotonically and tends to $F_{\mathsf{eq}}$ at long times. We calculate rate of change of $F(t)$

$$\begin{aligned} \frac{\mathsf{d}}{\mathsf{d}t} F(t) = \dot{F} &= \sum_i \dot{p}_i E_i + T \sum_i \dot{p}_i \ln p_i + T \sum_i p_i \frac{\dot{p}_i}{p_i} \\ &= \sum_i \dot{p}_i \left[E_i + T \ln p_i\right] \end{aligned}$$

using $\sum_i \dot{p}_i = \partial_t \sum_i p_i = 0$. The rate of change of the probability distribution is specified by the master equation $\dot{p}_i = \sum_j w_{ij} p_j - \sum_j w_{ji} p_i$, giving for the rate of change of $F(t)$

$$\sum_{ij} \left(w_{ij} p_j - w_{ji} p_i\right) \left[E_i + T \ln p_i\right] \tag{1.26}$$

$$\overset{i \leftrightarrow j}{=} \sum_{ij} \left(w_{ji} p_i - w_{ij} p_j\right) \left[E_j + T \ln p_j\right] \tag{1.27}$$

Now take [(1.26)+(1.27)]/2 and obtain

$$= \frac{1}{2}\sum_{ij}\left(w_{ij}p_j - w_{ji}p_i\right)\left\{[E_i + T\ln p_i] - [E_j + T\ln p_j]\right\}$$

For a system invariant under time reversal, detailed balance gives for the rates $w_{ij} = w_{ji}\exp(-(E_i - E_j)/T)$, so

$$= \frac{1}{2}\sum_{ij} w_{ji}e^{-E_i/T}\left(e^{E_j/T}p_j - e^{E_i/T}p_i\right)\left\{[E_i + T\ln p_i] - [E_j + T\ln p_j]\right\}$$

The last two terms in this expression can be written in the form

$$()\{\} = (e^x - e^y)(Ty - Tx) = -T(e^x - e^y)(x - y) \leq 0 \tag{1.28}$$

which is always less than zero.

Thus the time-dependent free energy decreases with time. However, both terms in the definition of this free energy (1.24 ) are bounded from below (by the ground state energy, and the maximum of the entropy realized by a uniform distribution, respectively). $F(t)$ cannot decrease indefinitely, and must reach a fixed point. The Boltzmann measure is exactly such a fixed point because at $p_i = p_i^{\mathsf{eq}}$ we have $w_{ji}p_i - w_{ij}p_j = 0$, so $\dot{F} = 0$. However, the fact that the Boltzmann measure is a fixed point does not mean that there cannot be other fixed points, or even limit cycles. Proving for a specific system $\mathbf{p}(t)$ will converge to the Boltzmann measure from any starting configuration is challenging in most cases.

## 1.6 Properties of the Boltzmann measure

The Boltzmann measure

$$\boxed{p_i = \frac{1}{Z}e^{-\beta E_i}} \tag{1.29}$$

underlies the entire statistical physics of the equilibrium state. It specifies the probability $p_i$ that a system characterized by configurations $\{i\}$ with energies $\{E_i\}$ obeying detailed balance in equilibrium with heat bath at temperature $1/\beta = T$ is found in configuration $i$. The **partition function** $Z(\beta) = \sum_i \exp(-\beta E_i)$ is fixed by the normalization condition $\sum_i p_i = 1$. $Z(\beta)$ is the central tool of equilibrium statistical mechanics for it connects the microscopic description $\{E_i\}$ with the macroscopic expectation values$\langle E\rangle$. We briefly recapitulate some of the key properties of the Boltzmann measure and the partition function, and briefly discuss the connection with thermodynamics.

The expectation value of the energy in equilibrium can easily be calculated from the partition function $Z(\beta)$

$$\langle E\rangle_{\mathsf{eq}} = \sum_i p_i E_i = \frac{1}{Z}\sum_i e^{-\beta E_i}E_i = \frac{\partial}{\partial\beta}\ln Z(\beta) \tag{1.30}$$

Consider now the **spectrum of energies**. Denote by $\rho(E)$ the number of times an energy $E$ occurs in the set $\{E_i\}$. Then

$$Z = \sum_i e^{-\beta E_i} = \sum_E \rho(E)e^{-\beta E}\ . \tag{1.31}$$

For continuous energy levels, a corresponding density of levels can be defined. This spectrum is sufficient to calculate the partition function since $Z(\beta) = \sum_E \rho(E)e^{-\beta E}$.

We are (often) interested in physical systems consisting of a large number $N$ degrees of freedom, labeled $k = 1, 2, ..., N$. Frequently, energies $E_i$ scales linearly (extensively) with $N$, so we define the (intensive) $e_i = E_i/N = \mathcal{O}(1)$. Also, the number of configurations with energy $E_i = e_i N$ is frequently exponential in the system size, so the number of configurations with energies between $e$ and $e + \Delta e$ is $\exp(Ns(e))\Delta e$. $s(e)$ is the *canonical entropy* per degree of freedom[3].

We now focus on the partition function, and using $\rho(E) = \exp(Ns(E/N))$ to obtain

$$Z(\beta) = \sum_E \rho(E) e^{-\beta E} = \int \mathrm{d}e\, e^{Ns(e)-\beta Ne} = \int \mathrm{d}e\, \exp\left\{ N \underbrace{(s(e) - \beta e)}_{-g(e)} \right\} \tag{1.32}$$

We will solve this integral with the so called **saddle point approximation**, which is asymptotically exact for $N \to \infty$.

Consider an integral of the form

$$I = \int \mathrm{d}e\, e^{-Ng(e)} = \int \mathrm{d}e\, e^{-(Ng(e_0) + \frac{N}{2}(e-e_0)^2 g''(e_0) + ...)}$$

For large $N$ and a well-behaved function $g(e)$ , the integral is dominated by the global minimum of $g(e)$ at $e_0$, where the integrand has a sharply-peaked maximum. We expand $g(e)$ as a Taylor series around the minimum of $g(e)$ at $e_0$. Since $g'(e_0) = 0$, the first order term is zero. The necessary condition $g'(e_0) = 0$ is called the **saddle-point equation** or saddle-point condition. The zeroth-order term $\exp(-Ng(e_0))$ is independent of $e$, the first-order term is zero from the saddle point condition. The second-order term $\exp\left(\frac{N}{2}(e - e_0)^2 g''(e_0)\right)$ leads to a Gaussian integral

$$I \approx e^{-Ng(e_0)} \sqrt{\frac{2\pi}{g'' N}} \tag{1.33}$$

$$\frac{1}{N} \ln I \approx -g(e_0) - \frac{1}{2} \frac{\ln N}{N} - ... \tag{1.34}$$

where we have assumed that $g$ has only one global minimum. For contour integrals in the complex plane, we are dealing with a saddle-point rather than a minimum, which gave this method its name. Another name used in the mathematics literature is Laplace's method. In cases where there are several global minima of $g(e)$, a sum over all minima needs to be taken. Saddle-point integrals are used all over physics to derive asymptotic expression in limiting cases, including the thermodynamic limit ($N$ is large), the classical limit ($\hbar^{-1}$ is large), and others.

For large $N$, the integral in the partition function (1.32) can be evaluated by saddle-point integration

$$Z = \int \mathrm{d}e\, e^{N(s(e)-\beta Ne)} \overset{N\to\infty}{=} c e^{-N\beta f(\beta)} \tag{1.35}$$

where $-\beta f(\beta) = \max_e(s(e) - \beta e)$. Comparison with (1.25) allows to identify $f(\beta)$ is the the intensive equilibrium free energy[4].

The maximum over $e$ is specified by $\frac{\mathrm{d}}{\mathrm{d}e}(s(e) - \beta e) = 0$ giving

$$\frac{\mathrm{d}s}{\mathrm{d}e} = \beta \tag{1.36}$$

[3] This entropy actually coincides with the information entropy $-\sum_i p_i \log p_i$ if $p_i$ is the Boltzmann measure at inverse temperature $\beta(e)$ determined by $Ne = \langle \mathcal{H} \rangle_\beta$.

[4] Note that we have dropped logarithmic prefactors from the saddle-point integral (compare with (1.6)), as in the thermodynamic limit their contribution to any extensive quantity scales sublinearly with $N$.

This relationship is of limited use in practice, since the problem of finding $s(e)$ is just as hard as evaluating $Z(\beta)$. Conceptually, however, this equation recovers the definition of the inverse temperature from thermodynamics, and $-\beta f(\beta) = \max_e(s(e) - \beta e)$ is just the Legendre transformation linking energy and free energy. For large $N$, the Boltzmann factor $e^{-\beta N e}$ is a rapidly decreasing function of $e$, the density of configuration $e^{-\beta N s(e)}$ is rapidly increasing with $e$. For large $N$, only configurations with energies maximizing $(s(e) - \beta e)$ contribute to the partition function. The expected energy is then $e(\beta) = e_0$ and all other energies either have too small a Boltzmann factor or too small an entropy. Thus we expect that results obtained from the canonical ensemble defined by the Boltzmann distribution, and the microcanonical ensemble will generally agree.

# 2 Mean Field Theory

In some cases of interacting systems, the interaction can be approximated by some constant effective field acting on each degree of freedom. This effective field arises from the interaction with other degrees of freedom.

We first study a model for with this so-called **mean-field theory** turns out to be exact. It is also the simplest model of the ferromagnetic phase transition.

## 2.1 The Weiss Ferromagnet

Consider a set of degrees of freedom $s_i \in \{\pm 1\}$ with $i = 1, ..., N$ (spins) and the Hamiltonian

$$\mathcal{H}[s] = -\frac{J}{2N} \sum_{i,j} s_i s_j - h \sum_i s_i \ . \tag{2.1}$$

Each degree of freedom defines an elementary spin, which can either be up or down in some preferred direction. In this model, every spin interacts with every other spin in the same way. The interaction $J/N$ of each pair scales as $1/N$ to keep the total energy extensive. The partition function is obtained by summing over all possible spin configurations:

$$Z = \mathsf{Tr}_{\underline{s}} e^{-\beta \mathcal{H}[\underline{s}]} = \prod_{i=1}^{N} \left( \sum_{s_i = \pm 1} \right) e^{+\frac{\beta J}{2N} \sum_{i,j} s_i s_j + \beta h \sum_i s_i} \tag{2.2}$$

We will now calculate the partition function using the Gaussian integral

$$e^{\frac{x^2}{2a}} = \sqrt{\frac{a}{2\pi}} \int_{-\infty}^{\infty} \mathsf{d}m e^{-\frac{1}{2}am^2 + mx} \tag{2.3}$$

and

$$\begin{aligned} a &= N\beta J \\ x &= \beta J \sum_i s_i \end{aligned}$$

giving

$$e^{\frac{\beta J}{2N} \sum_{i,j} s_i s_j} = \sqrt{\frac{N\beta J}{2\pi}} \int_{-\infty}^{\infty} \mathsf{d}m e^{-\frac{1}{2} N\beta J m^2 + \beta J m \sum_i s_i} \ . \tag{2.4}$$

This integral representation of the exponential of a square is called the **Hubbard-Stratonovich-transformation** and allows to evaluate the partition function. Plugging equation 2.4 in equation

2.2, we obtain

$$\begin{aligned}
Z &= \prod_{i=1}^{N}\left(\sum_{s_i=\pm 1}\right)\sqrt{\frac{N\beta J}{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}m\, e^{-\frac{1}{2}N\beta Jm^2+\beta Jm\sum_i s_i+\beta h\sum_i s_i} \\
&= \sqrt{\frac{N\beta J}{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}m\, e^{-\frac{1}{2}N\beta Jm^2}\prod_i\left(\sum_{s_i} e^{\beta Jms_i+\beta hs_i}\right) \\
&= \sqrt{\frac{N\beta J}{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}m\, e^{-\frac{1}{2}N\beta Jm^2}\left[2\cosh(\beta Jm+\beta h)\right]^N \\
&= \sqrt{\frac{N\beta J}{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}m\, e^{-\frac{1}{2}N\beta Jm^2+N\ln(2\cosh(\beta Jm+\beta h))} \\
&= \sqrt{\frac{N\beta J}{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}m\, e^{-\beta Nf(\beta,m)}
\end{aligned}$$

where the so-called free energy function

$$-\beta f(\beta,h,m) \equiv -\frac{1}{2}\beta Jm^2+\ln\left(2\cosh(\beta Jm+\beta h)\right) \tag{2.5}$$

has been defined. This function depends on the integration variable $m$, apart from the parameters $\beta$ and $h$.

In the thermodynamic limit $N\to\infty$, the integrand becomes a sharply peaked function of $m$, and the integral can be performed by saddle-point integration. Neglecting sub-leading prefactors gives $Z\approx\exp(-\beta Nf(\beta))$ with

$$\beta f(\beta,h)=\min_m f(\beta,h,m)$$

The necessary condition for a minimum $\partial_m f(\beta,m)=0$ gives the equation

$$m=\tanh(\beta Jm+\beta h)\ . \tag{2.6}$$

To visualize where the minimum over $m$ lie (2.5) is plotted in figure 2.1 for a temperature below and above $J$. We see that solutions of the self-consistent equation 2.6 are:

$$m=0 \qquad \begin{array}{c}\text{global minimum for } T>T_c\\ \text{local maximum for } T<T_c\end{array} \tag{2.7}$$

$$m=\pm\overline{m}(\beta) \qquad \text{global minimum for } T<T_c \tag{2.8}$$

As shown in figure 2.2, for $T<T_c$, $\overline{m}$ shows a discontinuity as a function of $h$. Note that although $-\beta f(\beta,m)$ is perfectly analytic in all variables, the $\min[]$-operation results in the global minimum $\bar{m}(h)$ jumping discontinuously as $h$ is changed. All effects linked to this behaviour thus only occur in the limit of large $N$.

### 2.1.1 The expected magnetisation

Calculate the probability that a given spin (say $i=1$) points up

$$\begin{aligned}
\langle\delta_{s_1,1}\rangle &= \frac{\int\mathrm{d}m\exp\left\{-\frac{1}{2}\beta NJm^2+(N-1)\ln 2\cosh(\beta Jm+\beta h)+(J\beta m+h\beta)\right\}}{\int\mathrm{d}m\exp\{-\beta Nf(m,\beta)\}} \\
&\to_{N\to\infty}\frac{\exp\left\{\beta J\overline{m}+\beta h\right\}}{(\exp(J\beta\overline{m}+\beta h)+\exp(-J\beta\overline{m}-\beta h))}=\frac{1+\overline{m}}{2}\ ,
\end{aligned}$$

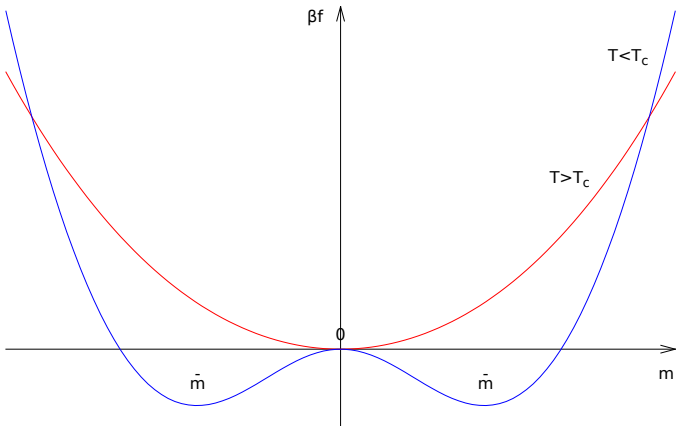



Figure 2.1: The function $\beta f(\beta, h, m)$ as a function of $m$ for $h = 0$ at temperatures $T$ below and above the critical temperature $T_c$.

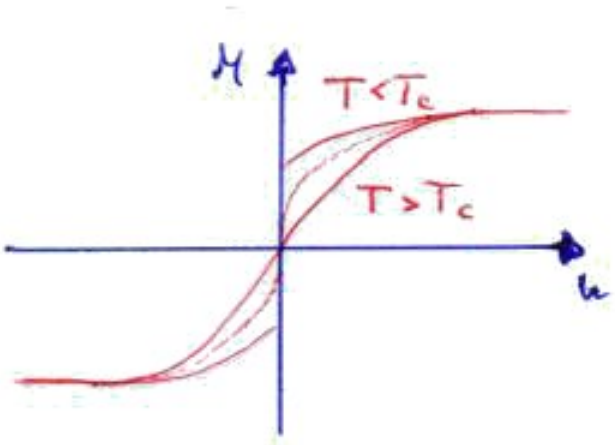



Figure 2.2: The function $\bar{m}(h)$ shows a discontinuity as a function of $h$ for $T > T_c$..

where all spins except $i = 1$ where summed over, and we used the saddle point equation $\overline{m} = \tanh(\beta J\overline{m} + \beta h)$. Similarly, the probability for a given spin to point in the down direction is $(1-\bar{m})/2$ and the expected magnetisation is $\bar{m} = \bar{m}(\beta, h)$. Hence, although the integration variable $m$ has no physical interpretation, the value of $m$ where the global minimum of the free energy function occurs is the expectation value of the magnetisation $\bar{m}(\beta,h)\langle s_1\langle = \frac{1}{N}\sum_i^N \rangle s_i\langle$ under the Boltzmann measure.

Analogously, we can calculate the probability that a pair of spins $s_1, s_2$ are up, giving $(1+\bar{m})/2 \cdot (1+\bar{m})/2$, so

$$\langle s_1 s_2 \rangle_{0^\pm} = \langle s_1 \rangle_{0^\pm} \langle s_2 \rangle_{0^\pm}$$

This means that the spins are not coupled for the Weiss ferromagnet and the connected expectation value is zero:

$$\langle s_1 s_2 \rangle_c \equiv \langle s_1 s_2 \rangle - \langle s_1 \rangle \langle s_2 \rangle = 0$$

In fact, in the thermodynamic limit the Boltzmann measure factorizes for the Weiss ferromagnet, as can be seen from

$$\begin{aligned} Z &\propto \int_{-\infty}^{\infty} \mathrm{d}m e^{-\frac{1}{2}N\beta J m^2} \Pi_i \left[ \sum_{s_i} e^{-\beta J m s_i + \beta h s_i} \right] \\ &\overset{N\to\infty}{\to} e^{-\frac{1}{2}N\beta J \overline{m}^2} \Pi_i \left[ \sum_{s_i} e^{-\beta J \overline{m} s_i + \beta h s_i} \right] \end{aligned}$$

We will explore this point further in section 2.2.

### 2.1.2 Behaviour near the critical temperature

Hoping to describe the system near a phase transition, where $m$ is small, we expand the free energy function (2.5) as a Taylor-series in $m$

$$\beta f(\beta, m) = \frac{1}{2}\beta J m^2 - \ln 2 - \frac{(\beta J m + \beta h)^2}{2} + \frac{(\beta J m + \beta h)^4}{12} + ... \tag{2.9}$$

$$= \frac{1}{2}\beta J(1-\beta J)m^2 - \beta^2 J m h + \frac{\beta^4 J^4 m^4}{12} + \mathcal{O}(h^2, m^3 h, m^2 h^2, ...) + \text{const} \tag{2.10}$$

$$\tag{2.11}$$

using $\ln(2\cosh x) = \ln 2 + \frac{x^2}{2} - \frac{x^4}{12} + \mathcal{O}(x^6)$. At low values of $m$, the quadratic term is dominating, at larger values the $m^4$-term. The phase transition occurs when the quadratic term switches sign at $1 - \beta_c J = 0$ resp. $T_c = J$. Below this temperature, the free energy function has two local minima, rather than one. Either local minimum can be a global minimum, depending on the external magnetic field $h$.

Such an expansion can also be performed for other models. Near a second-order phase transition, we expect that

$$\beta f(\beta, m) = \frac{1}{2} t m^2 + u m^4 - c m h + \text{const} \tag{2.12}$$

The coefficients $t, u, c$ particular to the the Weiss ferromagnet are

$$\begin{aligned} t &= \beta J(1-\beta J) = \beta J \beta(T - T_c) \approx \beta_c^2 J(T - T_c) = \frac{1}{J}(T - T_c) \equiv a(T - T_c) \\ u &= \frac{\beta^4 J^4}{12} \\ c &= \beta^2 J \end{aligned}$$

From the expansion (2.12), the behaviour of the system near the phase transition follows.

**Magnetization** The saddle-point of $f(\beta, m)$ for $h = 0^+$ and $t < 0$ is given by

$$tm + 4um^3 = 0 \Rightarrow m^2 = \frac{|t|}{4u} \tag{2.13}$$

The solution is $m \propto |t|^{1/2} \equiv |t|^\beta$ which is a power-law behaviour with exponent $\beta = 1/2$.

**Critical susceptibility** At $t = 0$ and for finite $h$

$$4um^3 - ch = 0 \Rightarrow m = \left(\frac{c}{4u}h\right)^{1/3} \propto h^{1/3} \equiv h^{1/\delta} \tag{2.14}$$

i.e. $h \propto m^\delta$ with $\delta = 3$.

**Zero-field susceptibility** At $t > 0$ and for small $h$

$$tm + 4um^3 - ch = 0 \Rightarrow m \approx \frac{ch}{t} \propto h \equiv h^\gamma \tag{2.15}$$

the exponent is $\gamma = 1$.

**Specific heat** For $t > 0$, $\beta f = \text{const}$. For $t < 0$, we use equation 2.13 giving

$$\begin{aligned}
\beta f &= \text{const} + \frac{1}{2}t\left(\frac{|t|}{4u}\right) + n\left(\frac{|t|}{4u}\right)^2 \\
&= \text{const} - \frac{1}{2}\frac{t^2}{4u} + \frac{t^2}{16u} \\
&= \text{const} - \frac{1}{16}\frac{t^2}{u}
\end{aligned}$$

The resulting specific heat is then

$$\begin{aligned}
c_V &= \frac{1}{N}\frac{\mathrm{d}\langle H\rangle}{\mathrm{d}T} \\
&= -\beta^2\partial_\beta^2(\beta f) = \frac{1}{8u}\ .
\end{aligned}$$

A power-law-ansatz $c_V \propto t^{-\alpha}$ yields an exponent of $\alpha = 0$.

These values for the exponents are called the **mean-field exponents**:

| | $\alpha$ | $\beta$ | $\gamma$ | $\delta$ |
|---|---|---|---|---|
| MF exponent | 0 | 1/2 | 1 | 3 |

It turns out that second-order phase transitions indeed show power-laws for various physical quantities near the phase transition. However, the exponents usually differ from the result we obtained here for the Weiss ferromagnet. Specifically, the exponents are found to depend on the dimensionality of the system[1]. To account for these power-laws in detail, we will have to wait until we can introduce the concept of renormalisation.

[1] For the Weiss ferromagnet, with every spin interacting equally with every other spin, this dimensionality is effectively infinite.

## 2.2 General mean field theory

We will now turn mean-field theory as a tool to systematically derive approximations for systems, whose partition function cannot be calculated as easily as that of the Weiss ferromagnet. In this context, mean-field theory, or rather, the mean-field approximation, can be framed as the search for a simple probability function (for which the expectation value of the energy, etc. can be calculated easily), which is in some sense "close" to the Boltzmann measure of the system under consideration.

We denote the Boltzmann measure as

$$P(\mathbf{s}) = \frac{1}{Z} e^{-\beta \mathcal{H}[\mathbf{s}]} \ , \tag{2.16}$$

where $\mathbf{s}$ is a vector denoting the configuration of the system $\mathbf{s} = (s_1, s_2, \ldots, s_N)$. A particular measure of "closeness" of two functions is the **Kullback-Leibler-divergence** between two functions $Q(\mathbf{s})$ and $P(\mathbf{s})$, defined as

$$D(Q||P) = \sum_{\mathbf{s}} Q(\mathbf{s}) \ln \frac{Q(\mathbf{s})}{P(\mathbf{s}} \tag{2.17}$$

The logarithm in this expression is positive if $Q(\mathbf{s}) > P(\mathbf{s})$, zero if $Q(\mathbf{s}) = P(\mathbf{s})$, and negative if $Q(\mathbf{s}) < P(\mathbf{s})$ for a given $\mathbf{s}$. One can show that

- $D(Q||P)$ is zero iff $Q(\mathbf{s}) = P(\mathbf{s})$ for all $\mathbf{s}$
- $D(Q||P)$ is positive iff $Q(\mathbf{s}) \neq P(\mathbf{s})$

We can now choose some family of functions $Q(\mathbf{s})$ and minimize $D(Q||P)$ across this family in order to make $Q(\mathbf{s})$ as similar to the Boltzmann measure as possible. Mean field theory consists of restricting choosing the family of functions $Q(\mathbf{s})$ which factorize in the degrees of freedom $s_1, s_2, ..., s_N$. The Weiss model of ferromagnetism (2.1) is called a mean-field model, because in this case mean field theory is exact since the Boltzmann measure factorizes in the thermodynamic limit, see section 2.1.1.

For a system with binary degrees of freedom $s_i \in \pm 1$ any measure factorizing in $i$ can be written as

$$Q(\mathbf{s}) = \prod_{i=1}^{N} \frac{1 + s_i m_i}{2} \text{ for } s_i \in \{-1, 1\} \tag{2.18}$$

This choice of parameterization is convenient because

$$\langle s_i \rangle_Q = \frac{1 + m_i}{2}(+1) + \frac{1 - m_i}{2}(-1) = m_i$$

For the Boltzmann measure (2.16) the Kullback-Leibler divergence is

$$\begin{aligned} D(Q||P) &= \sum_{\mathbf{s}} Q(\mathbf{s}) \ln \frac{Q(\mathbf{s})}{P(\mathbf{s})} \\ &= \sum_{\mathbf{s}} Q(\mathbf{s}) \ln Q(\mathbf{s}) - \sum_{\mathbf{s}} Q(\mathbf{s}) \ln P(\mathbf{s}) \\ &= \langle \ln Q \rangle_Q + \beta \langle \mathcal{H}[\mathbf{s}] \rangle_Q + \ln Z \end{aligned}$$

The first term gives the information entropy of the distribution $Q(\mathbf{s})$, $S_Q = -\sum_{\mathbf{s}} Q(\mathbf{s}) \ln Q(\mathbf{s})$, the second term the expected energy $E_Q = \sum_{\mathbf{s}} Q(\mathbf{s}) \mathcal{H}[\mathbf{s}]$ under the distribution $Q$. Taking these together, the part of the Kullback-Leibler distance which depends on $Q(\mathbf{s})$ can be considered a free energy of $Q(\mathbf{s})$.

To see how that these terms can be evaluated easily since $Q(\mathbf{s})$ factorizes, we consider a particular Hamiltonian $\mathcal{H}[\mathbf{s}] = -\sum_{i,j} J_{ij} s_i s_j - \sum_i h_i s_i$ for some coupling matrix $\mathbf{J}$ and local fields $\mathbf{h}$. We obtain

$$\langle \ln Q \rangle_Q = \sum_i \frac{1+m_i}{2} \ln \frac{1+m_i}{2} + \frac{1-m_i}{2} \ln \frac{1-m_i}{2} \tag{2.19}$$

$$\langle \mathcal{H} \rangle_Q = \sum_{\mathbf{s}} Q(\mathbf{s}) \mathcal{H}[\mathbf{s}] \tag{2.20}$$

$$= \sum_{\{s_i\}} \prod_i \left( \frac{1+m_i s_i}{2} \right) \left( -\sum_{ij} J_{ij} s_i s_j - \sum_i h_i s_i \right) \tag{2.21}$$

$$= -\frac{1}{2} \sum_{ij} J_{ij} m_i m_j - \sum_i h_i m_i \tag{2.22}$$

where we used

$$\sum_{s_i, s_j} \frac{1+m_i s_i}{2} \frac{1+m_j s_j}{2} J_{ij} s_i s_j = 2 \cdot 2 \frac{m_i}{2} \frac{m_j}{2}$$

as only even powers of $s_i$ contribute to the sum of $s_i$.

The minimum of $D(Q||P)$ within the space of factorizing function $QQ(\mathbf{s})$ by minimizing the free energy with respect to $\mathbf{m}$

$$\min_{\mathbf{m}} D(Q_m||P) = \min_{\mathbf{m}} \left( \langle \ln Q \rangle_Q + \beta \langle \mathcal{H} \rangle_Q \right) .$$

Setting $\partial_{m_i} D(Q||P) = 0$ gives with

$$\partial_{m_i} \left( \frac{1 \pm m_i}{2} \ln \frac{1 \pm m_i}{2} \right) = \pm \frac{1}{2} \ln \frac{1 \pm m_i}{2} \pm \frac{1}{2}$$

the following condition

$$\frac{1}{2} \ln \frac{1+m_i}{2} - \frac{1}{2} \ln \frac{1-m_i}{2} - \beta \sum_j J_{ij} m_j - \beta h_i = 0 \forall i \ . \tag{2.23}$$

Using arctanh$x = (1/2) \log \frac{1+x}{1-x}$ this can be written as

$$m_i = \tanh \left( \beta \sum_j J_{ij} m_j + \beta h_i \right) \tag{2.24}$$

whose solution gives the magnetisations $\mathbf{m}$ for the mean-field approximation to $\mathcal{H}[\mathbf{s}]$. Given these magnetisations, the mean-field approximation to the entropy and energy (2.19) can be evaluated.

For a translationally invariant model, such as the Weiss ferromagnet, $m_i = m \forall i$, and (2.24) reduces to the saddle-point equation of the Weiss ferromagnet (2.6).

# 3 Statistical Field Theory and Landau-Ginzburg Models

## 3.1 Statistical Fields

In principle, it is possible to extract all thermodynamical properties of a system from the partition function

$$Z(T) = \mathsf{Tr}\ \left[e^{-\beta \mathcal{H}_{\mathsf{mic}}}\right] \tag{3.1}$$

obtained by summing over all microscopic degrees of freedom. In many cases, the partition function cannot be evaluated exactly.

For many purposes, progress can be made by changing the focus from microscopic to **mesoscopic** scales much larger than the lattice spacing but much smaller than the system size. The resulting effective **statistical field theories**, are more amenable to analytic treatment.

We introduce coarse grained variables and define a magnetization field $\mathbf{m}(\mathbf{x})$ which represents the average of the elemental spins in the vicinity of a point $\mathbf{x}$. The scale $\ell$ of the coarse graining is chosen such that

- $\mathbf{m}(x_1), \mathbf{m}(x_2), ...$ contain enough microscopic variables such that the law of large numbers applies
- and $\mathbf{m}(\mathbf{x})$ can still be described by a smoothly varying function.

The transformation from the original degrees of freedom to the field $\mathbf{m}(\mathbf{x})$ is a a non-invertible mapping, since microscopic details are washed out due to averaging. It is possible to obtain the probabilities for field configurations $\mathbf{m}(\mathbf{x})$ by adding all the Boltzmann weights $e^{-\beta \mathcal{H}_{\mathsf{mic}}}$ of spin configurations corresponding to a particular field $\mathbf{m}(\mathbf{x})$. The partition function is then preserved

$$Z = \mathsf{Tr}\ \left[e^{-\beta \mathcal{H}_{\mathsf{mic}}}\right] \equiv \int \mathcal{D}\mathbf{m}(\mathbf{x})\, e^{-S[\mathbf{m}(\mathbf{x})]} \tag{3.2}$$

where $\int \mathcal{D}\mathbf{m}(\mathbf{x})$ denotes the integral over all field configurations $\mathbf{m}(\mathbf{x})$.

Obtaining the precise form of $e^{-S[\mathbf{m}(\mathbf{x})]}$ can be even harder than solving the full problem. Instead, we hope to guess $S[\mathbf{m}(\mathbf{x})]$ on the basis of very general principles

**Locality and uniformity** For a system consisting of disconnected parts, the overall probability is obtained as a product of independent probabilities. $S[\mathbf{m}]$ then decomposes into a sum going over to an integral in the continuum representation:

$$S[\mathbf{m}(\mathbf{x})] = \int \mathsf{d}^d\mathbf{x}\, f(\mathbf{m}(\mathbf{x}), \mathbf{x})$$

For a material that is uniform, all places are equivalent and there should be no explicit $\mathbf{x}$ dependence. However, if there are couplings between different parts of the system, we have to consider gradients of the field, leading to

$$S[\mathbf{m}(\mathbf{x})] = \int \mathsf{d}^d\mathbf{x} f(\mathbf{m}(\mathbf{x}), \nabla\mathbf{m}, \nabla^2\mathbf{m}, ...) \tag{3.3}$$

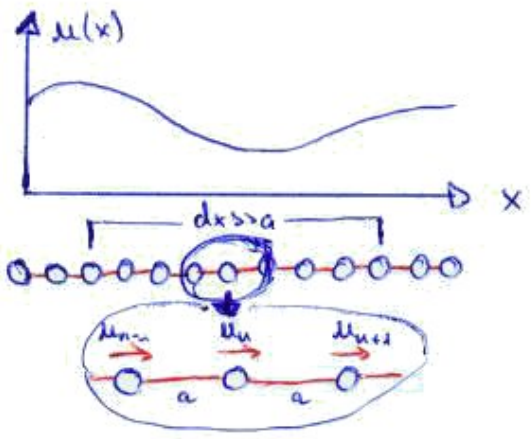



Figure 3.1: Displacements $\{u_n\}$ of a one-dimensional chain of particles and the coarse-grained field $u(x)$ of the continuous string in analogy to the coarse-grained magnetization $m(x)$.

**Analyticity** We expect that, as in the case of the free energy function, $S[\mathbf{m}]$ is analytic. Then, we can write $S[\mathbf{m}]$ as a series expansion in powers of $\mathbf{m}$ and its gradients.

**Symmetries** As any symmetries will survive the averaging process, only terms compatible with the symmetries of $\mathcal{H}_{\text{mic}}$ are considered.

**Stability** Since the coarse grained Boltzmann weights are supposed to describe a real problem, we will not tolerate unphysical solutions as divergence for infinite values of $\mathbf{m}$. E.g. this leads to the condition that the coefficient of the highest order power of $\mathbf{m}$ has to be positive.

## 3.2 Landau-Ginzburg theory for the ferromagnetic phase transition

As an example, we will now consider the component $m(\mathbf{x})$ in the direction of a magnetic field $h$.

The simplest functional $S[m(\mathbf{x})]$ in agreement with the above requirements is

$$S[m(\mathbf{x})] = \int \mathrm{d}^d\mathbf{x} \left[ -hm(\mathbf{x}) + \frac{t}{2}m^2(\mathbf{x}) + um^4(\mathbf{x}) + \frac{k}{2}(\nabla m)^2 \right] \tag{3.4}$$

- $m^3$,$\nabla m$-terms etc. are missing for reasons of symmetry ($S[m] = S[-m]$ for $h = 0$)
- We could add a term $(\nabla m)^2 m^2$ but this does not lead to new results
- $(\nabla m)^2$ penalizes fluctuations of $m(\mathbf{x})$ in $\mathbf{x}$ (adjacent regions of different $m(\mathbf{x})$)
- The parameters $h, t, k, u$ are effective parameters whose relationship to the microscopic $J, T$ is not known (to us). Calculating these relationships would be has hard as calculating the full partition function of the model
- Nevertheless, we make the assumption that $h, k, u$ are approximately constant near the phase transition, $t$ depends linearly on $T$.

The functional (3.4) is called the **Landau-Ginzburg** functional. One can view this functional as the Hamiltonian of some field $m(\mathbf{x})$. It is important to note that the Landau-Ginzburg functional

contains also entropic contributions from microscopic degrees of freedom which were averaged out (in the same way the quartic term in the small-$m$ expansion of the Weiss ferromagnet free energy function (2.9) arose from expanding the $\log\cosh$ originating from the sum over spins).

Indeed, we begin our analysis of the Landau-Ginzburg model with the ansatz of a constant magnetization $m(\mathbf{x}) = \overline{m}$. Integration over $\mathbf{x}$ yields the volume $V$ and

$$S[\overline{m}] = V(-h\overline{m} + \frac{t}{2}\overline{m}^2 + u\overline{m}^4) \tag{3.5}$$

giving the small-$m$ expansion of the Weiss ferromagnet free energy functional.

## 3.3 Stability analysis and upper critical dimension

In the following section, we will investigate the role of (small) fluctuations $\phi(\mathbf{x})$ around $m(\mathbf{x}) = \overline{m}$. Thus, we write $m$ as $m(\mathbf{x}) = \overline{m} + \phi(\mathbf{x})$. This gives us:

$$\begin{aligned} m^2(x) &= \overline{m}^2 + 2\overline{m}\phi(\mathbf{x}) + \phi^2(\mathbf{x}) \\ m^4(x) &= \overline{m}^4 + 4\overline{m}^3\phi(\mathbf{x}) + 6\overline{m}^2\phi^2(\mathbf{x}) + \mathcal{O}(\phi^3) \\ (\nabla m)^2 &= (\nabla\phi)^2 \end{aligned}$$

Inserting these formulae gives the Landau-Ginzburg functional with small fluctuations:

$$S[m] = V(-h\overline{m} + \frac{t}{2}\overline{m}^2 + u\overline{m}^4) + \int \mathsf{d}^d\mathbf{x}\left[\frac{k}{2}(\nabla\phi)^2 + \left(\frac{t}{2} + 6u\overline{m}^2\right)\phi^2(\mathbf{x})\right] + \mathcal{O}(\phi^3) \tag{3.6}$$

The terms coupling linearly to $\phi$

$$\left(t + 4u\overline{m}^2\right)\overline{m}\phi$$

vanish since at the saddle point $\overline{m}^2 = |t|/4u$ for $t < 0$ and $\overline{m} = 0$ for $t > 0$.

The first part of this expression gives us the contribution of the saddle point in $\overline{m}$, the second the contribution to $S[m]$ from small fluctuations around the saddle point.

Defining

$$\xi^{-2}\frac{k}{2} = \frac{t}{2} + 6u\overline{m} \tag{3.7}$$

the partition function can be written as

$$Z = \int \mathcal{D}m\, e^{-S[m]} \propto \int \mathcal{D}\phi \exp\left\{-\frac{k}{2}\int \mathsf{d}^d\mathbf{x}\left((\nabla\phi)^2 + \xi^{-2}\phi^2(\mathbf{x})\right)\right\} \tag{3.8}$$

Since the functional (3.6) is translation invariant, Fourier transformations are key to its analysis. In particular, derivatives of fields are then transformed into simple multiples of the Fourier modes; $\nabla\phi(\mathbf{x}) \to i\underline{k}\phi_k$ and $\nabla^2\phi(\mathbf{x}) \to -k^2\phi_k^2$.

We use the following conventions for discrete Fourier expansion/transform:

$$\phi(\mathbf{x}) = \sum_{\mathbf{q}} \phi_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{x}} \tag{3.9}$$

$$\phi_{\mathbf{q}} = \frac{1}{L^d}\int \mathsf{d}^d\mathbf{x}\,\phi(\mathbf{x})e^{-i\mathbf{q}\cdot\mathbf{x}} \tag{3.10}$$

For $L \to \infty$ we get (for simplicity in 1D)

$$\sum_q \phi_q(...) = \sum_n \phi_{2\pi n/L}(...)\int_{2\pi(n-1/2)/L}^{2\pi(n+1/2)/L} \frac{\mathsf{d}^d q}{(2\pi/L)^d} = \frac{1}{(2\pi)^d}\int \mathsf{d}^d q\,\phi(q)(...)\ ,$$

where we defined $\phi(q) = Lq$. This leads us to the continuous form:

$$\phi(\mathbf{x}) = \frac{1}{(2\pi)^d} \int \mathrm{d}^d\mathbf{q}\, \phi(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{x}} \tag{3.11}$$

$$\phi(\mathbf{q}) = \int \mathrm{d}^d\mathbf{x}\, \phi(\mathbf{x}) e^{-i\mathbf{q}\cdot\mathbf{x}} \tag{3.12}$$

Using the Fourier transform we obtain:

$$\begin{aligned}\int \mathrm{d}^d\mathbf{x}\, \phi^2(\mathbf{x}) &= \sum_{\mathbf{q},\mathbf{q}'} \phi_\mathbf{q}\phi_{\mathbf{q}'} \underbrace{\int \mathrm{d}^d\mathbf{x}\, e^{i(\mathbf{q}+\mathbf{q}')\cdot\mathbf{x}}}_{=\delta(\mathbf{q}-\mathbf{q}')L^d} \\ &= L^d \sum_\mathbf{q} \phi_\mathbf{q}\phi_{-\mathbf{q}} = L^d \sum_\mathbf{q} |\phi_\mathbf{q}|^2\end{aligned}$$

i.e. the integral transforms into a sum of Fourier modes. Analogously, the integral over the gradient transforms as

$$\int \mathrm{d}^d\mathbf{x} (\nabla\phi)^2 = L^d \sum_\mathbf{q} q^2 |\phi_\mathbf{q}|^2$$

Thus, we can write $S[m]$ as

$$S[\phi] = \frac{k}{2} L^d \sum_\mathbf{q} (q^2 + \xi^{-2}) |\phi_\mathbf{q}|^2. \tag{3.13}$$

We see that there are no couplings between the $\phi_q$ with different $q$ anymore (apart from couplings between modes with equal and opposite momenta, as $|\phi_\mathbf{q}|^2 = \phi_\mathbf{q}\phi^*_\mathbf{q} = \phi_\mathbf{q}\phi_{-\mathbf{q}}$. This expression (3.13) we have obtained can be seen as the Hamiltonian of a kinetic theory in $q$-space. The partition function is

$$Z = \int \mathcal{D}\phi\, e^{\frac{k}{2} L^d \sum_\mathbf{q} (q^2+\xi^{-2}) |\phi_\mathbf{q}|^2} \tag{3.14}$$

The integration over fields can be carried out noting that each pair of momenta appears twice in (3.13), once as $\phi_\mathbf{q}\phi_{-\mathbf{q}}$ and once as $\phi_{-\mathbf{q}}\phi_\mathbf{q}$, giving

$$\int \mathcal{D}\phi e^{-S[\phi]} = \prod_\mathbf{q}{}' \left( \int \mathrm{d}(\phi_q, \phi_q^*) e^{-kL^d(q^2+\xi^{-2})|\phi_q|^2} \right)$$

$\prod'$ means that every element $q, -q$ appears once, and $\phi_q$ determines $\phi_{-q} = \phi_q^*$. This leads us to a Gaussian integration:

$$\int \mathrm{d}(z, z^*) e^{-z^* w z} = \int \mathrm{d}(x, y) e^{-(x^2+y^2)w} = \sqrt{\frac{\pi}{w}}\sqrt{\frac{\pi}{w}} = \frac{\pi}{w}$$

Choosing $w = \frac{k}{2} L^d (q^2 + \xi^{-2})$, we obtain

$$Z \propto \prod_\mathbf{q}{}' \tfrac{1}{kL^d(q^2+\xi^{-2})} = \exp\left\{ \tfrac{1}{2} \textstyle\sum_\mathbf{q} \ln(k(q^2+\xi^{-2}))^{-1} \right\} \tag{3.15}$$

$$= \exp\left\{ -\tfrac{1}{2} \tfrac{L^d}{(2\pi)^d} \textstyle\int \mathrm{d}^d q \ln(k(q^2+\xi^{-2})) \right\} \tag{3.16}$$

$$\equiv e^{-\beta L^d f_{\mathsf{fluc}}(\beta)} \tag{3.17}$$

with $-\beta f_{\mathsf{fluc}}(\beta) = -\frac{1}{2} \frac{1}{(2\pi)^d} \int \mathrm{d}^d q \ln(k(q^2+\xi^{-2}))$.

What is the contribution of these fluctuations to observables near the phase transition? Consider the contribution of fluctuations to the heat capacity.

$$\partial_t \partial_\beta (\beta f_{\text{fluc}}(\beta)) \propto \frac{1}{\beta^2} \partial_t^2 (\beta f_{\text{fluc}})$$

The most singular term near the phase transition in this derivation is proportional to

$$\frac{1}{(2\pi)^d} \int \mathrm{d}^d q \frac{1}{(q^2 + \xi^{-2})^2}$$

This integral changes its behaviour drastically at $d = 4$.

- for $d > 4$ the integral is divergent at short distances or $q \to \infty$. Introducing a cutoff in $q$-space corresponding to the inverse lattice spacing $a^{-1}$

$$\int \mathrm{d}^d \mathbf{q} \frac{1}{(q^2 + \xi^{-2})^2} \propto \int \mathrm{d}q \frac{q^{d-1}}{(q^2 + \xi^{-2})^2} \approx \left(\frac{1}{a}\right)^d \left(\frac{1}{\left(\frac{1}{a}\right)^2 + \xi^{-2}}\right)^2 \approx a^{4-d}$$

- for $d < 4$ the integral is convergent for both $q \to 0$ and $q \to \infty$ (provided $\xi^{-2}$ is finite)

$$\int \mathrm{d}^d \mathbf{q} \frac{1}{(q^2 + \xi^{-2})^2} \propto \xi^{-d} \xi^4 \int \mathrm{d}^d \mathbf{q}' \frac{1}{(q'^2 + 1)^2} \approx \xi^{4-d}$$

  This is finite for finite $\xi$ but diverges ($d < 4$) if $\xi \to \infty$. Since $\xi \propto 1/\sqrt{t}$ this divergence occurs at the critical point!

$d = d_c = 4$ is called the **upper critical dimension**. Below $d_c$ mean-field predictions are (singularly) wrong and fluctuations around $m(x) = \overline{m}$ dominate the partition function.

The parameter $\xi$ defined in (3.7), whose divergence signals the phase transition, has an important physical interpretation – it is the length over which fluctuations $\phi(\mathbf{x})$ are correlated. To show this, we consider the connected correlation function between field variables $\phi$ a distance $x$ apart

$$G(\mathbf{x}) = \langle \phi(\mathbf{x}) \phi(0) \rangle - \langle \phi(\mathbf{x}) \rangle \langle \phi(0) \rangle$$

Again, Fourier transform:

$$\begin{aligned} G_{\mathbf{q}} &= \frac{1}{L^d} \int \mathrm{d}^d \mathbf{x}\, G(\mathbf{x}) e^{-iqx} \\ &= \frac{1}{L^d} \int \mathrm{d}^d \mathbf{x} \langle \phi(\mathbf{x}) \phi(0) \rangle e^{-iqx} \\ &= \frac{1}{L^d} \int \mathrm{d}^d \mathbf{x} \sum_{\mathbf{q}_1, \mathbf{q}_2} \langle \phi(\mathbf{q}_1) \phi(\mathbf{q}_2) \rangle e^{iq_1 0 + iq_2 x - iqx} \\ &= \sum_{\mathbf{q}_1} \langle \phi(\mathbf{q}_1) \phi(\mathbf{q}) \rangle \end{aligned}$$

For $S[\phi_q] = \frac{k}{2} L^d \sum_{\mathbf{q}} (q^2 + \xi^{-2}) |\phi_q|^2$:

$$\langle \phi_q \phi_{q'} \rangle = \begin{cases} \frac{1}{kL^d(q^2 + \xi^{-2})} & \text{for } q' = -q \\ 0 & \text{for } q' \neq -q \end{cases} \tag{3.18}$$

In $d = 1$, $G(x) = \sum_q G_q e^{iqx}$ can easily be calculated, and we restrict our discussion to this case. We get similar results in $d > 1$ apart from prefactors

$$\begin{aligned} G(x) &= \sum_q G_q \, e^{iqx} = \sum_{q,q'} \langle \phi_q \phi_{q'} \rangle \, e^{iqx} \\ &= \frac{1}{kL} \sum_q \frac{e^{iqx}}{q^2 + \xi^{-2}} \to_{L\to\infty} \frac{1}{2\pi k} \int_{-\infty}^{\infty} \frac{\mathrm{d}q \, e^{iqx}}{(q + i\xi^{-1})(q - i\xi^{-1})} \end{aligned}$$

We solve this integral by contour integration (poles at $\pm i\xi^{-1}$). According to the residual theorem, this integral gives us for $x > 0$

$$I_1 = 2\pi i \frac{e^{-x\xi^{-1}}}{2i\xi^{-1}}$$

and for $x < 0$

$$I_2 = -2\pi i \frac{e^{-x\xi^{-1}}}{-2i\xi^{-1}}$$

Putting this together gives us:

$$G(x) = \frac{\xi}{2k} e^{-|x|/\xi} \tag{3.19}$$

We see that $\xi$ is the correlation length of the fluctuations. This length diverges as $t \to 0$.

# 4 Lattice Models

## 4.1 Some examples

We assume a microscopic Hamiltonian (which may or may not be a good description of a specific system). The degrees of freedom ("spins") are located on lattice points. Some models can be solved exactly, i.e. $Z(\beta, h)$ can be calculated in the thermodynamic limit.

Some famous examples are

**The Ising model** I. At each site $i$ of some lattice, there is a spin (degree of freedom) which takes the two values of $+1$ and -1. The Hamiltonian modeling a ferromagnet is given by

$$\mathcal{H} = -\sum_{<i,j>} J_{ij} s_i s_j - \sum_i h_i s_i \ , \tag{4.1}$$

where the sum is over all pairs of spins whose lattice sites are adjacent to each other. In particular, $\mathcal{H} = -J\sum_{<i,j>} s_i s_j - h\sum_i s_i$ with a coupling constant $J > 1$ and a magnetic field $h$ is a good description for a ferromagnet.

**The O(n) model:** Each lattice site is occupied by a vector $\mathbf{s}_i = (s_i^1, s_i^2, ..., s_i^d)^T$ which is a unit vector (i.e. $\sum_m (s_i^m)^2 = 1$).

Then the Hamiltonian is given by

$$\mathcal{H} = -J \sum_{<i,j>} \mathbf{s}_i \cdot \mathbf{s}_j - \mathbf{h} \cdot \sum_i \cdot \mathbf{s}_i \tag{4.2}$$

It is called O(n)-model because the Hamiltonian is invariant under rotation.

**The Potts model:** Each site of the lattice is occupied by a $q$-valued spin $s_i \in \{1, 2, 3, ..., q\}$. The interactions between the spins are described by a Hamiltonian

$$\mathcal{H} = -J \sum_{<i,j>} \delta_{s_i s_j} - h \sum_i \delta_{s_i 1} \tag{4.3}$$

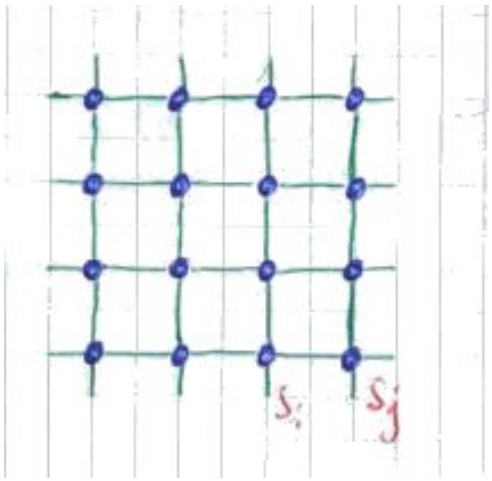


Figure 4.1: 2d lattice

where neighbouring spins give a non-zero contribution only if they are in the same configuration.

**Spin s-models:** The spins at every site of the lattice take the $2s-1$ values $s_i \in -s, -s+1, ..., +s$. A general nearest-neighbour Hamiltonian is

$$\mathcal{H} = -\sum_{<i,j>} J_1 s_i s_j + J_2 (s_i s_j)^2 + ... + J_{2s} (s_i s_j)^{2s} - h \sum_i s_i \tag{4.4}$$

For $s = 1/2$ we get back the Ising model.

Models which are mathematically equivalent to the Ising model or others also appear in contexts different from ferromagnetism. An example is a lattice gas where $s_i = -1$ means that particles are absent and $s_i = 1$ that particles are present).

## 4.2 Monte-Carlo simulation

In physics, the term **Monte Carlo** refers to the use of randomness to approximate deterministic quantities. The basic idea can be explained with the following example: Suppose a lake whose average depth is to be determined. The exact solution is

$$\frac{\int_{\text{lake}} \mathrm{d}x\, \mathrm{d}y\, \text{depth}(x,y)}{\int_{\text{lake}} \mathrm{d}x \mathrm{d}y}$$

Choosing $M$ points *randomly and uniformly* we can approximate the exact result by $M^{-1} \sum_{n=1}^{M} \text{depth}(x_n, y_n)$.

Compare now the formula for the average depth $\int \mathrm{d}x \mathrm{d}y d(x,y) / \int \mathrm{d}x \mathrm{d}y$ with

$$\langle \mathcal{H} \rangle = \frac{1}{Z} \text{Tr}\ \mathcal{H} e^{-\beta \mathcal{H}} = \frac{\text{Tr}\ \mathcal{H} e^{-\beta \mathcal{H}}}{\text{Tr}\ e^{-\beta \mathcal{H}}}$$

In the "lake-formula", the trace is replaced by the integral and the statistical weight of each point is equal to one. This leads to the idea of using Monte Carlo methods in statistical physics. We aim to calculate $\langle f(\mathbf{s}) \rangle = Z^{-1} \sum_{\mathbf{s}} f(\mathbf{s}) \exp(-\beta \mathcal{H}(\mathbf{s}))$. In order to do this, we draw $M$ configurations independently from the Boltzmann measure and approximate:

$$\langle f(\mathbf{s}) \rangle \approx \frac{1}{M} \sum_{\mu=1}^{M} f(\mathbf{s}^\mu)$$

### 4.2.1 Markov chain Monte Carlo (MCMC)

The first step is thus to generate a (large) set of states $\mathbf{s}$ independently drawn from the Boltzmann distribution. Suppose a Markov process on the configurations $\mathbf{s}$, i.e. a set of transition rates from configuration $\mathbf{s}$ to $\mathbf{s}'$, $W_{\mathbf{s}'\mathbf{s}}$. Our goal is to choose these rates such that the equilibrium distribution $p^{eq}(\mathbf{s})$ is the Boltzmann distribution.

An outline of a typical MCMC algorithm is:

1. pick an initial configuration $\mathbf{s}$ (in any way)
2. choose a new configuration $\mathbf{s}'$ with probability $W_{\mathbf{s}'\mathbf{s}}$
3. repeat "many times" until equilibrium has been reached.

The rates $W_{\mathbf{s}'\mathbf{s}}$ specify an artificial dynamics of the system with no relation to the actual dynamics of the system we model. The only requirement on the rates $W_{\mathbf{s}'\mathbf{s}}$ is that the steady state (invariant measure) is the Boltzmann measure.

Different such rate matrices can be chosen according to speed of convergence and computational efficiency.

As the steady state of MCMC should be the Boltzmann measure, the rates $W_{\mathbf{s}'\mathbf{s}}$ have to obey detailed balance.

$$\frac{W_{\mathbf{s}'\mathbf{s}}}{W_{\mathbf{s}\mathbf{s}'}} = \frac{e^{-\beta\mathcal{H}[\mathbf{s}']}}{e^{-\beta\mathcal{H}[\mathbf{s}]}} = e^{-\beta(\mathcal{H}[\mathbf{s}']-\mathcal{H}[\mathbf{s}])}$$

In detailed balance $W_{\mathbf{s}'\mathbf{s}}e^{-\beta\mathcal{H}[\mathbf{s}]}/Z = W_{\mathbf{s}\mathbf{s}'}e^{-\beta\mathcal{H}[\mathbf{s}']}/Z$ holds, so probability fluxes between $\mathbf{s}$ and $\mathbf{s}'$ balance if the probability of finding the system in a configuration $\mathbf{s}$ is $e^{-\beta\mathcal{H}[\mathbf{s}]}/Z$

One set of rates which obey detailed balance is **Metropolis dynamics** defined as follows:

1. Given the system is in a configuration $\mathbf{s}$, pick a new configuration $\mathbf{s}'$ from some neighbour set of $\mathbf{s}$. In practice, $\mathbf{s}'$ is picked such that $\mathbf{s}'$ and $\mathbf{s}$ differ by one spin flip.

2. "Accept" the transition from $\mathbf{s}$ to $\mathbf{s}'$ with probability $p_{\text{acc}}$

$$\begin{aligned} p_{\text{acc}} &= 1 && \text{if } \mathcal{H}[\mathbf{s}'] < \mathcal{H}[\mathbf{s}] \\ p_{\text{acc}} &= e^{-\beta(\mathcal{H}[\mathbf{s}']-\mathcal{H}[\mathbf{s}])} && \text{if } \mathcal{H}[\mathbf{s}'] \geq \mathcal{H}[\mathbf{s}] \end{aligned}$$

   else "reject" and remain at $\mathbf{s}$.

These rates obey detailed balance

$$\frac{W_{\mathbf{s}'\mathbf{s}}}{W_{\mathbf{s}\mathbf{s}'}} = \frac{\begin{cases} 1 & \Delta\mathcal{H} < 0 \\ e^{-\beta\Delta\mathcal{H}} & \Delta\mathcal{H} \geq 0 \end{cases}}{\begin{cases} e^{+\beta\Delta\mathcal{H}} & \Delta\mathcal{H} < 0 \\ 1 & \Delta\mathcal{H} \geq 0 \end{cases}} = e^{-\beta\Delta\mathcal{H}}$$

Another way is the **Glauber dynamics**:

1. Pick a spin at random, compute $\Delta\mathcal{H}$ resulting from a potential flip of this spin.

2. Flip the spin with probability

$$p_{\text{flip}} = \frac{1}{1+e^{\beta\Delta\mathcal{H}}}.$$

## 4.3 The lower critical dimension

In LG-theory, the effects of fluctuations are more pronounced in low dimensions. In low dimensions, fluctuations can become so strong that the ordered phase no longer exists.

**Discrete systems** Consider a system of Ising spins (or other degrees of freedom with a discrete symmetry). $d = 1$ with $L$ spins:

$$\uparrow\uparrow\uparrow\uparrow\downarrow\downarrow\downarrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow$$

We are interested in the energy difference of introducing a domain somewhere. The energy cost is $2J$ and the entropy cost is $\ln L$. $\Rightarrow -\Delta(\beta F) = +\ln L - 2\beta J$.

$$\Rightarrow \Delta F = -k_B T \ln L + 2J$$

$-k_BT \ln L \to \infty$ as $L \to \infty$ if $T > 0$. The free energy of a system with an additional domain wall is thus lower than a system without it. The so called **Peierls-argument** shows that the ordered phase is unstable against the introduction of domains provided $T > 0$ (no finite $T$ phase transition).

Consider now a domain of down-spins with domain length $l$ in $d = 2$. The energy difference is $\approx 2Jl$. As at each point you can go in three directions, the entropy difference is $\approx \ln$[number of random walks of length l], therefore $\approx \ln[\mu]^l = l \ln \mu$ with $\mu \approx 3$.

$$\Delta F = -k_BTl \ln \mu + Jl = l(-k_BT \ln \mu + J)$$

This goes to 0 at a critical temperature of $T_c = J/k_B \ln \mu$. For $T < T_c$, $\Delta F$ is increased by the introduction of domains, for $T > T_c$, $\Delta F$ is decreased in that case. We thus expect, that unlike in 1d, 2d-systems can sustain a low-temperature phase with long-range order.

**Continuous systems** The situation in continuous systems (degrees of freedom with continuous symmetry) is quite different. Here, domain walls acquire a spatial extent $l_{\text{domain}}$. The energy contribution of a domain wall is $\langle \Delta \mathcal{H} \rangle / N \propto \langle (\Delta\theta)^2 \rangle \approx (1/l_{\text{domain}})^2$ where $\Delta\theta$ is the angle between neighboring spins. Taking the domain wall size to be comparable to the size of the domain itself (minimum domain size), we find

$$\Delta \mathcal{H} = l_{\text{domain}}^d l_{\text{domain}}^{-2} = l_{\text{domain}}^{d-2}$$

This is different from the scaling $l_{\text{domain}}^{d-1}$ in discrete systems. This means in continuous systems, domains are easier to introduce because they cost less energy.

Already in $d = 2$, the ordered state is destroyed by fluctuations due to the introduction of domains. (Mermin-Wagner-theorem). The lower critical dimension for discrete systems is $d_{lc} = 1$ and for continuous systems $d_{lc} = 2$. For $d \leq d_{lc}$ there is no ordered state at a finite $T$.

## 4.4 The Ising Model in $d = 1$

Following these general arguments, we will attempt to solve some models exactly, starting with the Ising model in one dimension.

Writing $\beta J = k$ and $\beta \tilde{h} = h$, we have the Hamiltonian

$$-\beta \mathcal{H}[\mathbf{s}] = k \sum_{<i,j>} s_i s_j + h \sum_i s_i \tag{4.5}$$

With periodic boundary conditions ($s_{N+1} \equiv s_1$) we obtain the following partition function

$$\begin{aligned} Z &= \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \ldots \sum_{s_N=\pm 1} \prod_{i=1}^{N} \exp\{k s_i s_{i+1} + \frac{h}{2}(s_i + s_{i+1})\} \\ &= Tr\left[\langle s_1|\mathbf{T}|s_2\rangle \langle s_2|\mathbf{T}|s_3\rangle \ldots \langle s_N|\mathbf{T}|s_1\rangle\right] \\ &= \sum_{s_1} \langle s_1|\mathbf{T}^N|s_1\rangle = \text{Tr } \mathbf{T}^N \end{aligned}$$

where we introduced the "transfer matrix" $\mathbf{T}$ with elements

$$\langle \sigma|T|\sigma'\rangle = \exp\left(k\sigma\sigma' + \frac{h}{2}(\sigma + \sigma')\right), \text{i.e. } \mathbf{T} = \begin{pmatrix} e^{k+h} & e^{-k} \\ e^{-k} & e^{k-h} \end{pmatrix} \tag{4.6}$$

The trace of $\mathbf{T}^N$ can be calculated by diagonalizing $\mathbf{T}$

$$Z = \text{Tr } \mathbf{T}^N = \text{Tr } \begin{pmatrix} \lambda_1^N & 0 \\ 0 & \lambda_2^N \end{pmatrix} = \lambda_1^N + \lambda_2^N$$

with the eigenvalues of denoted $\mathbf{T}$ $\lambda_1, \lambda_2$.

Calculating the partition function thus reduces to computing eigenvalues of $\mathbf{T}$.

$$\begin{aligned}
\det(\mathbf{T} - \lambda \underline{\underline{1}}) &= 0 \\
\left(e^{k+h} - \lambda\right)\left(e^{k-h} - \lambda\right) - e^{-2k} &= 0 \\
\lambda^2 - 2\lambda e^k \cosh h + \left(e^{2k} - e^{-2k}\right) &= 0
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \lambda_{+,-} &= e^k \cosh(h) \pm \sqrt{e^{2k}\cosh^2(h) - e^{2k} + e^{-2k}} \\
&= e^k \cosh(h) \pm \sqrt{e^{2k}\sinh^2(h) + e^{-2k}}
\end{aligned}$$

For $N \to \infty$, if $\lambda_+ > \lambda_-$ then $\lambda_+^N \gg \lambda_-^N$:

$$-\beta f = \frac{1}{N} \ln Z = \frac{1}{N} \ln(\lambda_+^N + \lambda_-^N) \to \ln \lambda_+$$

For $T \to 0$ ($\beta \to \infty$)

$$\lambda_{+,-} \to_{T\to 0} e^k(\cosh h \pm \sinh h) = e^{k \pm h}$$

This means the free energy is analytic in $\beta, h, J$ if $T > 0$.

The magnetization can be computed from $\partial_h(-\beta h)$. An alternative is based on the transfer matrix. We focus on the spontaneous magnetisation ($h = 0$).

$$\begin{aligned}
\langle s_j \rangle &= \frac{1}{Z} \sum_{s_1 s_2 \ldots s_N} s_j \prod_{i=1}^{N} e^{k s_i s_{i+1}} \\
&= \frac{1}{Z} \mathsf{Tr}\ \left(\langle s_1|T|s_2\rangle \langle s_2|T|s_3\rangle \ldots \langle s_{j-1}|T|s_j\rangle\, s_j \,\langle s_j|T|s_{j+1}\rangle \ldots \langle s_N|T|s_1\rangle\right) \\
&= \frac{1}{Z} \mathsf{Tr}\ \left(\mathbf{T}^{j-1} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mathbf{T}^{N-j+1}\right) \\
&= \frac{1}{Z} \mathsf{Tr}\ \left(\mathbf{T}^{N} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right)
\end{aligned}$$

The eigenvectors of $\mathbf{T}$ are $(1,-1)^T/\sqrt{2}$ and $(1,1)^T/\sqrt{2}$. Transforming to an eigenvector-system yields

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

and we get

$$\langle s_j \rangle = \frac{1}{Z} \mathsf{Tr}\ \left(\begin{pmatrix} \lambda_1^N & 0 \\ 0 & \lambda_2^N \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) = \frac{1}{Z} \mathsf{Tr}\ \begin{pmatrix} 0 & \lambda_1^N \\ \lambda_2^N & 0 \end{pmatrix} = 0$$

This result follows already from symmetry at $h = 0$. The same method, however, can be used to calculate the correlation function $\langle s_i s_j \rangle$ (equals the connected correlation function $\langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle$

because $\langle s_i \rangle = \langle s_j \rangle = 0$).

$$\begin{aligned}
\langle s_i s_j \rangle &= \frac{1}{Z} \sum_{s_1, s_2, \ldots, s_N} s_i s_j \prod_{l=1}^{N} e^{k s_l s_{l+1}} \\
&= \frac{1}{Z} \mathsf{Tr} \left( \mathbf{T}^{i-1} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mathbf{T}^{j-i} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mathbf{T}^{N-j+1} \right) \\
&= \frac{1}{Z} \mathsf{Tr} \left( \mathbf{T}^{j-i} \underline{\underline{\sigma}}_z \, \mathbf{T}^{N-(j-i)} \underline{\underline{\sigma}}_z \right) \\
&= \frac{1}{Z} \mathsf{Tr} \left( \left( \begin{pmatrix} \lambda_+^{j-i} & 0 \\ 0 & \lambda_-^{j-i} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_+^{N-(j-i)} & 0 \\ 0 & \lambda_-^{N-(j-i)} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right) \right) \\
&= \frac{1}{Z} \mathsf{Tr} \left( \left( \begin{pmatrix} 0 & \lambda_+^{j-i} \\ \lambda_-^{j-i} & 0 \end{pmatrix} \begin{pmatrix} 0 & \lambda_+^{N-(j-i)} \\ \lambda_-^{N-(j-i)} & 0 \end{pmatrix} \right) \right) \\
&= \frac{1}{Z} \mathsf{Tr} \begin{pmatrix} \lambda_+^{j-i} \lambda_-^{N-(j-i)} & 0 \\ 0 & \lambda_-^{j-i} \lambda_+^{N-(j-i)} \end{pmatrix} = \frac{\lambda_+^r \lambda_-^{N-r} + \lambda_-^r \lambda_+^{N-r}}{\lambda_+^N + \lambda_-^N}
\end{aligned}$$

with $r = j - i$. Suppose $r \ll N$ and $\lambda_- < \lambda_+$

$$\langle s_i s_j \rangle \approx \frac{\lambda_-}{\lambda_+}^r = e^{-r/\xi}$$

where $\xi$ is the correlation length:

$$\xi^{-1} \overset{h=0}{=} -\ln \left( \frac{\lambda_+}{\lambda_-} \right) = -\ln \left( \frac{e^k + e^{-k}}{e^k - e^{-k}} \right) = -\ln \tanh k \tag{4.7}$$

Physical insight: Phase transition, signalled by the diverging correlation length is at $T = 0$. At finite temperature no ordered state is possible because it is destroyed by fluctuations.

## 4.5 Approximation at high and at low temperature

### 4.5.1 The low temperature expansion

At low temperature, the Boltzmann measure is dominated by the low-energy configurations.

| | ground configuration | 1st excited configuration | 2nd excited configuration |
|---|---|---|---|
| energy | $\mathcal{H}_0$ | $\mathcal{H}_1$ | $\mathcal{H}_2$ |
| degeneracy | $m_0$ | $m_1$ | $m_2$ |
| contribution to $Z$ | $m_0 e^{-\beta \mathcal{H}_0}$ | $m_1 e^{-\beta \mathcal{H}_1}$ | $m_2 e^{-\beta \mathcal{H}_2}$ |

At low temperature, $\beta(\mathcal{H}_1 - \mathcal{H}_0) \gg 1$:

$$\frac{e^{-\beta \mathcal{H}_1}}{e^{-\beta \mathcal{H}_0}} = e^{-\beta(\mathcal{H}_1 - \mathcal{H}_0)} \ll 1$$

Then, the partition function

$$Z = m_0 e^{-\beta \mathcal{H}_0} + m_1 e^{-\beta \mathcal{H}_1} + m_2 e^{-\beta \mathcal{H}_2} + \ldots$$

involves successively smaller terms, provided $m_0, m_1, m_2, \ldots$ does not grow too quickly.

As an example, let us consider the Ising model in $d$ dimensions.

$$Z = 2 e^{Ndk} \left[ 1 + N e^{-4dk} + dN e^{-4k(2d-1)} + \ldots \right] \tag{4.8}$$

Pushing the series to higher orders allows to extract aspects of critical behaviour.

| | | | |
|---|---|---|---|
| ground configuration | all spins point up (or down) | $\beta\mathcal{H}_0 = -kdN$ | $m_0 = 2$ |
| 1st excited configuration | one spin up, all others down (or v.v.) | $\beta(\mathcal{H}_0 - \mathcal{H}_1) = 2k \cdot 2d$ | $m_1 = 2N$ |
| 2nd excited configuration | two adjacent spins up, all others down (or v.v.) | $\beta(\mathcal{H}_2 - \mathcal{H}_0) = 2k \cdot 2(2d-1)$ | $m_2 = 2Nd$ |

### 4.5.2 High-temperature approximation (general)

For $\beta\mathcal{H}[\mathbf{s}] \ll 1 \forall \mathbf{s}$ (high temperatures), the Taylor expansion $\exp(-\beta\mathcal{H}(\mathbf{s})) \approx 1 - \beta\mathcal{H}[\mathbf{s}] + \beta^2/2\mathcal{H}^2[\mathbf{s}] + ...$ is dominated by the first few terms. We exploit this expansion to approximate $Z$:

$$\begin{aligned}
\ln Z = \ln \mathsf{Tr}\ \left(e^{-\beta\mathcal{H}}\right) &= \ln \mathsf{Tr}\ \left(1 - \beta\mathcal{H}[\mathbf{s}] + \beta^2/2\mathcal{H}^2[\mathbf{s}] + ...\right) \\
&= \ln\left[(\mathsf{Tr}\ (1))\left(1 - \beta\frac{\mathsf{Tr}\ \mathcal{H}}{\mathsf{Tr}\ 1} + \frac{\beta^2}{2}\frac{\mathsf{Tr}\ \mathcal{H}^2}{\mathsf{Tr}\ 1} + ...\right)\right] \\
&= \ln\left(\mathsf{Tr}\ (1)\right) + \ln\left(1 - \beta\frac{\mathsf{Tr}\ \mathcal{H}}{\mathsf{Tr}\ 1} + \frac{\beta^2}{2}\frac{\mathsf{Tr}\ \mathcal{H}^2}{\mathsf{Tr}\ 1} + ...\right) \\
&= \ln 2^N + \ln\left(1 - \beta\left\langle\mathcal{H}\right\rangle_0 + \frac{\beta^2}{2}\left\langle\mathcal{H}^2\right\rangle_0 + ...\right)
\end{aligned}$$

For small $\beta$ we can make another approximation $\ln(1-x) = -x - x^2/2 + ...$ and get

$$\begin{aligned}
&= \ln Z_0 - \beta\left\langle H\right\rangle_0 - \frac{1}{2}\beta^2\left\langle\mathcal{H}\right\rangle_0^2 + \frac{\beta^2}{2}\left\langle\mathcal{H}^2\right\rangle_0 \\
&= \ln Z_0 - \beta\left\langle\mathcal{H}\right\rangle_0 + \frac{1}{2}\beta^2\left(\left\langle\mathcal{H}^2\right\rangle_0 - \left\langle\mathcal{H}\right\rangle_0^2\right) + ...\ ,
\end{aligned}$$

with $Z_0 = \log(2)$.

## 4.6 The Ising Model and graphs on the lattice

The following identity valid for binary spins is useful to derive another high-T expansion (and much more)

$$\begin{aligned}
e^{k\sigma\sigma'} &= \begin{cases} e^k & \text{if } \sigma\sigma' = 1 \\ e^{-k} & \text{if } \sigma\sigma' = -1 \end{cases} \\
&= \frac{1}{2}\left(e^k + e^{-k}\right) + \frac{1}{2}\left(e^k - e^{-k}\right)\sigma\sigma' \\
&= \cosh k + \sigma\sigma' \sinh k \\
&= \cosh k(1 + \sigma\sigma' t)
\end{aligned}$$

with $t = \tanh k$.

The Boltzmann weight for Ising spins can be written as

$$\exp\left\{k\sum_{<i,j>}\sigma_i\sigma_j\right\} = \prod_{<i,j>} e^{k\sigma_i\sigma_j} = \prod_{<i,j>}\cosh k(1 + \sigma\sigma' t) = \cosh^n k \prod_{<i,j>}(1 + \sigma\sigma' t)$$

what is a polynomial of t. Expanding in powers of $t$ gives a new type of high-T expansion.

More importantly, any odd power-term of $\sigma_i$ in $\prod_{<i,j>}(1 + \sigma\sigma' t)$ gives zero when the sum over the possible values $\sigma = \pm 1$ is performed.

This is the basis for a representation of $Z$ in terms of graphs:

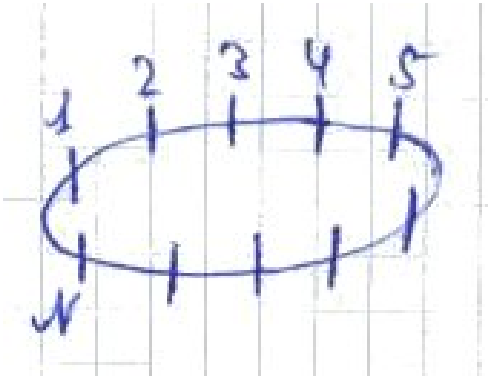



Figure 4.2: 1d ising model with closed boundaries

$$\prod_{<i,j>} (1 + \sigma_i \sigma_j t) = \text{sum over terms where each bond contributes 1 or } \sigma\sigma' t$$

In each term of this sum, each bond (i,j) either contributes a term $t$ (then we call it *occupied*), or a term $1$ (*unoccupied*). The occupied bonds form a **graph** on the lattice.

Summing over $\sum_{\{\sigma_i\}}$ only terms represented by *closed* graphs contribute to $Z$.

$$Z = 2^N \cosh^{\text{number of bonds}} k \sum_{\text{all closed graphs}} t^{\text{number of occupied bonds in graph}}$$

**Example: 1d-Ising model** (periodic boundary conditions), there is only one closed graph on this lattice (and the trivial case: no bonds)

$$Z = 2^N \cosh^N k (1 + \tanh^N k) = 2^N (\cosh^N k + \sinh^N k) \tag{4.9}$$

Compare this with the result obtained above: $Z = \lambda_1^N + \lambda_2^N \overset{h=0}{=} (2\cosh k)^N + (2\sinh k)^N$.

Correlations can also be calculated

$$\langle \sigma_m \sigma_n \rangle = \frac{1}{Z} \sum_{\sigma_i} \sigma_m \sigma_n \prod_{<i,j>} (1 + \sigma_m \sigma_n t)$$

only graphs with an even number of occupied bonds at all nodes except $m$ and $n$ contribute. $m$ and $n$ must have an odd number of occupied bonds.

$$\langle \sigma_m \sigma_n \rangle = \frac{t^{|n-m|} + t^{N-|n-m|}}{1 + t^N}$$

## 4.7 The Ising model in 2 dimensions

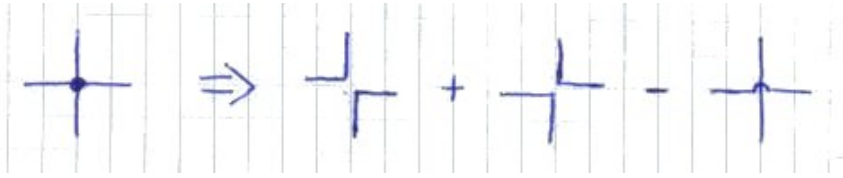
This section closely follows the exposition in Kardar, which is based on an approach pioneered by M. Kac.

### 4.7.1 Graphs and walks

A graph consists of bonds which are either absent or present. Walks where each bond can be traversed more than once are easier to deal with. After each step you have exactly three possibilities (in 2d) to go on, independently of where the walk has gone before. Claim:

$$\begin{aligned} Z &= 2^N \cosh^{\text{number of bonds}} k \sum_{\text{all closed graphs}} t^{\text{number of occupied bonds in graph}} \\ &= 2^N \cosh^{\text{number of bonds}} k \sum_{\text{all loops of walks without u-turns}} t^{\text{number of walks}} (-1)^{\text{number of crossings}} \end{aligned}$$

Consider a single site with 4 occupied bonds (0,2 are trivial). We have three walks, one of these has a crossing. This leads to a sum of 1 walk.



For a single bond we have the walk between these to points. Additionally, there is an infinite number of walks that contain the bond, but for every of these walk, another one can be generated by cutting the bond and exchanging the lines which gives us a sign of $-1$.

### 4.7.2 Multiple Loops

$$\begin{aligned} \sum \text{all (weighted) walks} &= 1 + 1 \text{ loop walks} + 2 \text{ loop walks} + 3 \text{ loop walks} + ... \\ &= 1 + \Xi + \frac{1}{2}(\Xi)^2 + \frac{1}{3!}(\Xi)^3 + ... \\ &= \exp\{\Xi\} \end{aligned}$$

$$\Rightarrow \ln Z/N = \ln(2\cosh^2 k) + \frac{1}{N}\left(\sum \text{all weighted walks with one loop}\right)$$

Notation: denote the sum over all (weighted) walks of a single loop with length $l$, starting at $(x', y')$ and ending at $(x, y)$ by

$$\langle x, y|W^*(l)|x', y'\rangle$$

* stands for weighted walks, i.e. $(-1)^{n_c}$.

$$\frac{1}{N}\left(\sum 1\right) = \frac{1}{N}\sum_{x,y}\sum_{l=1}^{\infty}\frac{1}{l}t^l \langle x, y|W^*(l)|x, y\rangle \tag{4.10}$$

is the key contribution to the partition function (4.10)

### 4.7.3 Walks on the lattice and transfer matrices

We show how walks on the lattice can be enumerated using transfer matrices. For now we study unweighted walks, $\langle x, y|W|x', y'\rangle$. The appropriate phase factor depending on the number of crossings will be introduced in the next section.

Define the adjacency matrix of a lattice

$$T_{ij} = \begin{cases} 1 & \text{if nodes i,j are neighbours} \\ 0 & \text{otherwise} \end{cases}$$

Walks on the lattice of length $l$ can be generated by raising the adjacency matrix to the $l$-th power.

**Example: 1d**

$$\mathbf{T} = \begin{pmatrix} 0 & 1 & 0 & \cdots \\ 1 & 0 & 1 & \cdots \\ 0 & 1 & 0 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

$$\langle x|W(l=1)|x'\rangle = T_{xx'}$$

$$\langle x|W(l=2)|x'\rangle = \sum_{x''} \langle x|W(l=1)|x''\rangle \langle x''|W(l=1)|x'\rangle = \sum_{x''} T_{xx'}T_{x''x'} = \left(\mathbf{T}^2\right)_{xx'}$$

Notation $|x\rangle$ was chosen in analogy to QM, so use $\underline{1} = \sum_x |x\rangle\langle x|$, in real space, with $|x\rangle$ corresponding to $(0, ..., 0, 1, 0, ...0)^t$. Since $\mathbf{T}$ is translation-invariant, Bloch-states are eigenstates of $\mathbf{T}$.

$$|q\rangle = (..., e^{-2iq}, e^{-iq}, 1, e^{iq}, e^{2iq}, ...)^t/\sqrt{N}$$

We get $\langle x|q\rangle = e^{iqx}/\sqrt{N}$, $\langle q|x\rangle = e^{-iqx}/\sqrt{N}$,

$$\langle x|x'\rangle = \sum_q \langle x|q\rangle \langle q|x'\rangle = \frac{1}{N}\sum_q e^{iq(x-x')} = \delta_{x,x'}$$

Proof that $|q\rangle$ is an eigenstate of $\mathbf{T}$:

$$\begin{aligned} \langle x|T|q\rangle &= \sum_{x'} \langle x|T|x'\rangle \langle x'|q\rangle \\ &= \sum_{x'} \left(\delta_{x',x+1} + \delta_{x',x-1}\right) e^{iqx'}/\sqrt{N} \\ &= \left(e^{iq(x+1)} + e^{iq(x-1)}\right)/\sqrt{N} = \left(e^{iq} + e^{-iq}\right) e^{iqx}/\sqrt{N} \\ &= (2\cos q)\langle x|q\rangle \equiv T(q)\langle x|q\rangle \end{aligned}$$

Now we can compute

$$\begin{aligned} \langle 0|W(l)|0\rangle = \langle 0|T^l|0\rangle &= \sum_q \langle 0|T^l|q\rangle \langle q|0\rangle \\ &= \sum_q T^l(q)\langle 0|q\rangle \langle q|0\rangle = \frac{1}{N}\sum_q T^l(q) \end{aligned}$$

This step uses that the Bloch states are also eigenstates of the product $\mathbf{T}^l$.

**Example: 2d**

$$\langle x,y|T|x',y'\rangle = \delta_{x,x'}(\delta_{y,y'+1}+\delta_{y,y'-1})+\delta_{y,y'}(\delta_{x,x'+1}+\delta_{x,x'-1})$$

Again, $|q_x,q_y\rangle$ form eigenstates of T:

$$\begin{aligned}\langle x,y|T|q_x,q_y\rangle &= \sum_{x',y'}\langle x,y|T|x',y'\rangle\langle x',y'|q_x,q_y\rangle \\ &= \frac{1}{\sqrt{N}}\left[e^{iq_x x}\left(e^{iq_y(y+1)}+e^{iq_y(y-1)}\right)+e^{iq_y y}\left(e^{iq_x(x+1)}+e^{iq_x(x-1)}\right)\right] \\ &= e^{i(q_x x+q_y y)}(\cos q_y+\cos q_x)\frac{1}{\sqrt{N}}e^{i(q_x x+q_y y)} \equiv T(q_x,q_y)\langle x,y|q_x,q_y\rangle\end{aligned}$$

Number of loops starting and ending at $0,0$ of length $l$:

$$\begin{aligned}\langle 0,0|W(l)|0,0\rangle &= \sum_{q_x,q_y}\langle 0,0|T^l|q_x,q_y\rangle\langle q_x,q_y|0,0\rangle \\ &= \sum_{q_x,q_y}T^l(q_x,q_y)\langle 0,0|q_x,q_y\rangle\langle q_x,q_y|0,0\rangle \\ &= \frac{1}{N}\sum_{q_x,q_y}T^l(q_x,q_y)\end{aligned}$$

We are still missing $(-1)^{N_c}$, i.e. we have not (yet) excluded U-turns.

### 4.7.4 The phase term $(-1)^{n_c}$

We need to compute the number of walks $\langle 0|W^*|0\rangle$ weighting each loop with $(-1)^{n_c}$. If $(-1)^{n_c}$ can be calculated using local information from each step, we have a chance!

**Theorem** (Whitney) The number $n_c$ of self-crossings of a planar loop is related to the total angle $\Theta$ its tangent turns through by

$$(n_c) \mod 2 = \left(1+\frac{\Theta}{2\pi}\right) \mod 2$$

Thus, we can compute $(-1)^{n_c}$ as

$$(-1)^{n_c} = \exp\left\{i\pi\left(1+\frac{\Theta}{2\pi}\right)\right\} = -e^{\frac{i}{2}\sum_{i=1}^{l}\theta_i}$$

with $\theta_i$ being the angle the walk turns through at step $i$.

$$\langle 0|W^*(l)|0\rangle \;=\; -\frac{1}{2}\sum \text{directed loops of } l \text{ steps, with no u-turns, from 0 to 0} \times e^{\frac{i}{2}\sum_{i=1}^{l}\theta_i}$$

To introduce changes of direction into the transfer matrix, we define $\langle x,y,\mu|W^*(l)|x',y',\mu'\rangle$ which gives us the sum over all directed paths from $x',y'$ departing along direction $\mu'$ ending at $x,y$ and leaving $x,y$ in direction $\mu$ weighted with $e^{\frac{i}{2}\sum_{i=1}^{l}\theta_i}$.

Each lattice point $x,y$ and direction $\mu$ specifies a particular state vector. Now also the weighted walks allow a transfer-matrix approach

$$\langle x,y,\mu|W^*(l)|x',y',\mu'\rangle = \sum\langle x,y,\mu|W^*(l-1)|x'',y'',\mu''\rangle\langle x'',y'',\mu''|W^*(l=1)|x',y',\mu'\rangle$$

Let us define with $|\mu\rangle \in \{|ket{\rightarrow}, |\uparrow\rangle, |\leftarrow\rangle, |\downarrow\rangle\}$ and calculate

$$\begin{aligned}\langle x,y,\mu|W^*(l=1)|x',y',\mu'\rangle &\equiv \langle x,y,\mu|T^*|x',y',\mu'\rangle \\ &= T^* = \begin{pmatrix} \langle x,y|x'+1,y'\rangle & \langle x,y|x',y'+1\rangle\, e^{-i\pi/4} & 0 & \langle x \\ \langle x,y|x'+1,y'\rangle\, e^{i\pi/4} & \langle x,y|x',y'+1\rangle & \langle x,y|x'-1,y'\rangle\, e^{-i\pi/4} & \\ 0 & \langle x,y|x',y'+1\rangle\, e^{i\pi/4} & \langle x,y|x'-1,y'\rangle & \langle x \\ \langle x,y|x'+1,y'\rangle\, e^{-i\pi/4} & 0 & \langle x,y|x'-1,y'\rangle\, e^{i\pi/4} & \end{pmatrix}\end{aligned}$$

where $\langle x,y|x',y'\rangle = \delta x x' \delta_{yy'}$. $\langle x,y,\mu|T^*|x',y',\mu'\rangle$ is a $4N \times 4N$ matrix, it is "block-diagonal" in Fourier basis (diagonal up to the $\mu\mu'$-blocks).

$$\sum_{x',y'} \langle x,y,\mu|T^*|x',y',\mu'\rangle \langle x',y'|q_x,q_y\rangle = \langle \mu|T^*(q_x,q_y)|\mu'\rangle \langle x,y|q_x,q_y\rangle$$

with

$$T^*(q_x,q_y) = \begin{pmatrix} e^{iq_x} & e^{-i(q_x+\pi/4)} & 0 & e^{iq_x+\pi/4)} \\ e^{-i(q_x-\pi/4)} & e^{-iq_x} & e^{i(q_x-\pi/4)} & 0 \\ 0 & e^{-i(q_x-\pi/4)} & e^{iq_x} & e^{i(q_x-\pi/4)} \\ e^{-i(q_x+\pi/4)} & 0 & e^{i(q_x+\pi/4)} & e^{iq_x} \end{pmatrix}$$

The quantity we are after

$$\begin{aligned}\langle 0,0|W^*(l)|0,0\rangle &= -\frac{1}{2}\sum_{\mu=1}^{4} \langle 0,0,\mu|T^{*l}|0,0,\mu\rangle \\ &= -\frac{1}{2}\sum_{\mu,q_x,q_y} \langle 0,0,\mu|T^{*l}|q_x,q_y,\mu\rangle \langle q_x,q_y,\mu|0,0,\mu\rangle \\ &= -\frac{1}{2}\frac{1}{N}\sum_{\mu,q_x,q_y} \left[T^{*l}(q_x,q_y)\right]_{\mu\mu} \\ &= -\frac{1}{2}\frac{1}{N}\sum_{q_x,q_y} tr\mathbf{T}^{*l}\end{aligned}$$

$$\begin{aligned}\frac{1}{2}\sum_{l=1}\frac{t^l}{l}\langle 0,0|W^*(l)|0,0\rangle &= \frac{1}{2N}\sum_{l,\mathbf{q}}\frac{1}{l} Tr\left(\mathbf{T}^*(\mathbf{q})t\right)^l \\ &= \frac{1}{2N}\sum_{\mathbf{q}} Tr\log\left(\underline{\underline{1}} - tT^*(q_x,q_y)\right)\end{aligned}$$

because $\ln(1-x) = -\sum l = 1^{\infty}\frac{1}{l}x^l$. Thus

$$\frac{1}{N}\ln Z = \ln(2\cosh^2 k) + \frac{1}{2N}\sum_{\mathbf{q}} \log\det\left(\underline{\underline{1}} - tT^*(q_x,q_y)\right)$$

using $tr(\log M) = \sum_\alpha \ln\lambda_\alpha = \log\prod_\alpha \lambda_\alpha = \log\det M$.

Working out the determinant and replacing the sum by an integral gives

$$\frac{1}{N}\ln Z = \ln(2\cosh^2 k) + \frac{1}{2}\int \frac{\mathsf{d}^2 q}{(2\pi)^2}\log\left[(1+t^2)^2 - 2t(1-t^2)(\cos q_x + \cos q_y)\right]$$

### 4.7.5 Critical behaviour

We expect singularities in the derivations of $\ln Z$ to arise from long lengthscales i.e. small $q_x, q_y$.

Let us therefore consider $q_x, q_y \to 0$. The integrand becomes

$$\begin{aligned} \log\left[(1+t^2)^2 - 2t(1-t^2)(\cos q_x + \cos q_y)\right] &\to \log\left[(1+t)^2 - 2t(1-t^2)\right] \\ &= \ln\left[1 + 2t^2 + t^4 - 4t + 4t^3\right] \\ &= \ln(1-t^2-2t)^2 \geq 0 \end{aligned}$$

The minimum of $(1-t^2-2t)^2$ is reached at the critical temperature (upper root of) $t = t_c = -1 \pm \sqrt{2}$ which leads us to a critical $\beta$ of

$$\beta_c = \frac{\log(1+\sqrt{2})}{2J}$$

Expansion $t = t_c + \Delta t$ and $\cos q_x + \cos q_y = 1 - \frac{1}{2}(q_x^2 + q_y^2) + \ldots$ gives

$$\frac{\ln Z}{N} \approx -\frac{1}{\pi}\left(\frac{\Delta t}{t_c}\right)^2 \log\left(\frac{\Delta t}{t_c}\right)$$

The second derivative (heat capacity) diverges logarithmically $C \propto \ln \Delta t \Rightarrow \alpha = 0$.

# 5 Renormalization

Suppose a system whose behaviour is fully determined by some (microscopic) Hamiltonian. Can the behaviour of this system on large scales be described in terms of a new (effective, large-scale) Hamiltonian? What are the couplings of this new Hamiltonian?

- Renormalization Group (RG) first arose in quantum field theory
- Kadanoff developed the idea to apply RG in statistical physics, Wilson performed first calculations in the 70s (Nobel prize to Wilson in 1982).
- tremendous impact on statistical physics and condensed matter physics; allows a deep understanding of critical phenomena.

## 5.1 The idea

Consider an ensemble of configurations specified by $\frac{1}{Z}e^{-\mathcal{H}[\mathbf{s}]}$, where $\mathcal{H}$ contains all (microscopic) couplings, fields, ... between spins.

All these configurations (1d for simplicity) have statistical weight $e^{-\mathcal{H}[\mathbf{s}]}$ relative to each other.

$$\uparrow\downarrow\uparrow\downarrow\uparrow\uparrow\downarrow\uparrow\downarrow\uparrow\uparrow\uparrow$$
$$\downarrow\uparrow\uparrow\uparrow\downarrow\downarrow\uparrow\downarrow\downarrow\downarrow\uparrow\uparrow$$
$$\downarrow\uparrow\uparrow\downarrow\uparrow\downarrow\downarrow\downarrow\uparrow\uparrow\downarrow\uparrow$$
$$\uparrow\uparrow\downarrow\uparrow\uparrow\downarrow\downarrow\uparrow\downarrow\uparrow\uparrow\downarrow$$

We can label these $2^N$ configurations $\mathbf{s_1}, \mathbf{s_2}, \mathbf{s_3}, \ldots$ with relative weights in equilibrium $e^{-\mathcal{H}[\mathbf{s}_1]}, e^{-\mathcal{H}[\mathbf{s}_2]}, e^{-\mathcal{H}[\mathbf{s}_3]}, \ldots$.

Consider disregarding a subset of the spins, e.g. every second spin:

$$\uparrow\cdot\uparrow\cdot\uparrow\cdot\downarrow\cdot\downarrow\cdot\uparrow\cdot$$
$$\downarrow\cdot\uparrow\cdot\downarrow\cdot\uparrow\cdot\downarrow\cdot\uparrow\cdot$$
$$\downarrow\cdot\uparrow\cdot\uparrow\cdot\downarrow\cdot\uparrow\cdot\downarrow\cdot$$
$$\uparrow\cdot\downarrow\cdot\uparrow\cdot\downarrow\cdot\downarrow\cdot\uparrow\cdot$$

This list contains only $2^{N/2}$ configurations. The relative weights of one of these configurations follows from combining the weights of all the configurations in the first list which correspond to the given configuration in the second (these differ only in the spins which were disregarded in going from the first list to the second).

However, the second list, and the corresponding statistical weights, can also be regarded as a statistical ensemble in its own right, that is, a new ensemble of spins $s'$, whose statistical weights are described by a new Hamiltonian $\mathcal{H}'[\mathbf{s}']$, with new couplings, fields etc.

The big question is: Is there a physical system corresponding to the new configuration that is a Hamiltonian $\mathcal{H}'[\mathbf{s}']$ such that the new relative weights are $e^{-\mathcal{H}'[\mathbf{s}']}$? $\mathcal{H}'$ describes the original system on a length scale twice the original lattice spacing.

We can, at least in principle, construct $\mathcal{H}'[\mathrm{s}']$ exploiting that "disregarding" degrees of freedom is equivalent to summing over them in the Boltzmann measure:

$$e^{-\mathcal{H}'[\mathrm{s}']} := \sum_{\text{subset of spins}} e^{-\mathcal{H}[\mathrm{s}]} \Rightarrow Z' = Z$$

This procedure can be repeated, yielding successive descriptions of the original system on increasing length scales. It is convenient to rescale the system at each step, resulting in a new system with half the number of degrees of freedom and new couplings between them. The behaviour of the new system on length scales $l'$ is the same as the behaviour of the old system on length scales $l$.

The behaviour of couplings $J$ under iteration of these steps gives information about the large-scale behaviour of the original system. If $J, J', J'', ... \to \infty$, we are in the low-T phase. If $J, J', J'' \to 0$ in a high-T phase.

## 5.2 Three examples

### 5.2.1 1d Ising model

This is a pedagogical example since we already have an exact solution. However, here RG-analysis by decimation can be applied directly.

The original Hamiltonian is $\mathcal{H}[s] = -J\sum_i s_i s_{i+1}$ with periodic boundary conditions. The Boltzmann factor can be written as $e^{-\mathcal{H}[s]} = \prod_i e^{Js_i s_{i+1}} = e^{Js_1 s_2} e^{Js_2 s_3} ...$

To construct $e^{-\mathcal{H}'[s']}$, take $s'_1 = s_1, s'_2 = s_3, ...$ and sum over $s_2, s_4, ....$

Writing $e^{Js_1 s_2} = \cosh J(1 + s_1 s_2 \tanh J)$ for $s_1 s_2 = \pm 1$ with $x = \tanh J$ results in

$$\begin{aligned} e^{-\mathcal{H}[s]} &= \cosh J(1 + s_1 s_2 x) \cosh J(1 + s_2 s_3 x) \cosh J(1 + s_3 s_4 x) ... && (5.1) \\ \Rightarrow e^{-\mathcal{H}'[s']} &= \sum_{s_2, s_4, ... = \pm 1} e^{-\mathcal{H}[s]} = 2\cosh^2 J(1 + s_1 s_3 x^2) 2\cosh^2 J(1 + s_3 s_5 x^2) ... && (5.2) \\ &= \prod_i \frac{2\cosh^2 J}{\cosh J'} \cosh J'(1 + s'_i s'_{i+1} x') = \frac{2\cosh^2 J}{\cosh J'} e^{-J' \sum_i s'_i s'_{i+1}} && (5.3) \end{aligned}$$

for the sum over $s_2$ stems only from even powers of $s_2$ here. We see that the couplings in $\mathcal{H}'[s']$ are given by

$$x' = x^2 \tag{5.4}$$

or $\tanh J' = \tanh^2 J$

$$\Rightarrow J' = \mathsf{artanh}(\tanh^2 J) \tag{5.5}$$

The new Hamiltonian is

$$\mathcal{H}'[s'] = -N' \ln \left( \frac{2\cosh^2 J}{\cosh J'} \right) - J' \sum_i s'_i s'_{i+1} \tag{5.6}$$

The coupling behaves as

$$x = \tanh J \equiv \tanh \beta J_{\mathsf{phys}} = \begin{cases} \to 0 & T \to \infty \Leftrightarrow \beta J_{\mathsf{phys}} \to 0 \\ \to 1 & T \to 0 \Leftrightarrow \beta J_{\mathsf{phys}} \to \infty \end{cases} \tag{5.7}$$

Thus, after many such decimation steps we arrive at

$$x^{(\infty)} = \begin{cases} 0 & 0 \le x < 1 \\ 1 & x = 1 \end{cases} \tag{5.8}$$

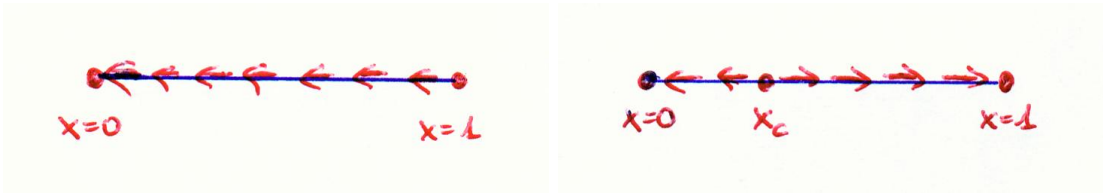



Figure 5.1: RG-flow of the Ising model in 1d (left) and 2d (right). You can easily see that the 1d model has only one stable fixed point. The 2d model has to stable and one unstable fixed point. This implies the existence of a phase transition.

If we start at exactly $x = 1$, we will stay there, otherwise, we will go to $x = 0$. This **RG-flow** of the coupling constant implies that at increasing length scales, the system looks increasingly disordered. There are two fixed points which map couplings $J$ onto themselves. $x = 0$ is a **stable fixed point**, $x = 1$ is an **unstable fixed point**.

At any temperature $T > 0$, the system is on long scales described by $x^{(\infty)} = 0$. Neighbouring spins (in the original model those can be far apart) are uncorrelated. Thus, we can conclude from RG alone that the 1d Ising model has no finite temperature phase transition.

The correlation length can also be calculated by RG arguments. The correlation length is unchanged under decimation alone since the statistics of remaining spins is unchanged. Under rescaling of the system the correlation length $\xi$ is halved so $\xi(J') = \xi(J)/2$. Rewritten in terms of $x$:

$$\xi(x^2) = \frac{1}{2}\xi(x)$$

This functional equation is solved by the function

$$\xi \propto \frac{1}{\ln x} = \frac{1}{\ln \tanh J} \tag{5.9}$$

recovering the result of the exact calculation.

### 5.2.2 2d Ising model

The Ising model in two dimensions is of course much harder and also less suited for decimation.

The key difference to the 1d case is that the RG flow keeps generating new types of couplings, RG flow equations (like $x' = x^2$) do not close.

In 1d, we have at low temperature $J' \approx J \langle s_2 \rangle_{s_1,s_3=1} \approx J$, in 2d, we have $J \approx 2J \langle s_2 \rangle_{s_1,s_3=1} \approx 2J$.

Hence, the $J = \infty, x = 1, T = 0$-fixed point is locally stable in 2d. That means there must be a third, unstable fixed point. At a fixed point, the correlation length is either $0$ or $\infty$ (the only solutions of $\xi(J) = \xi(J)/2$).

This new, unstable fixed point describes the critical behaviour of the 2d Ising model. Temperatures above $T_c$ flow into the $T = \infty$ fixed point, temperatures below $T_c$ into the $T = 0$ fixed point.

### 5.2.3 The Gaussian model

Consider the Landau-Ginzburg model, keeping only the quadratic terms ($m$-component in direction of the field):

$$Z = \int \mathcal{D}m(\mathbf{x}) \exp\left\{-\int \mathrm{d}^d\mathbf{x} \left(\frac{1}{2}tm^2(\mathbf{x}) + \frac{k}{2}(\nabla m)^2 - hm(\mathbf{x})\right)\right\} \tag{5.10}$$

Fourier transform:

$$\begin{aligned} m_{\mathbf{q}} &= \int \mathrm{d}^d\mathbf{x}\, m(\mathbf{x}) e^{-i\mathbf{q}\cdot\mathbf{x}} \\ m(\mathbf{x}) &= \tfrac{1}{L^d}\textstyle\sum_{\mathbf{q}} m_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{x}} \end{aligned}$$

The thermodynamic limit $m(q) = L^{-d} m_q$ gives us:

$$\begin{aligned} m(\mathbf{q}) &= \int \mathrm{d}^d\mathbf{x}\, m(\mathbf{x}) e^{-i\mathbf{q}\cdot\mathbf{x}} \\ m(\mathbf{x}) &= \int \tfrac{\mathrm{d}^d\mathbf{q}}{(2\pi)^d}\, m(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{x}} \end{aligned}$$

Terms in the action:

$$\int \mathrm{d}^d\mathbf{x}\, m^2(\mathbf{x}) = \int \mathrm{d}^d\mathbf{x} \int \frac{\mathrm{d}^d\mathbf{q}\mathrm{d}^d\mathbf{q}'}{(2\pi)^{2d}} e^{i(\mathbf{q}+\mathbf{q}')\cdot\mathbf{x}} m(\mathbf{q}) m(\mathbf{q}' = \int \frac{\mathrm{d}^d\mathbf{q}\mathrm{d}^d\mathbf{q}'}{(2\pi)^d} \delta(\mathbf{q}+\mathbf{q}') m(\mathbf{q}) m(\mathbf{q}') = \int \frac{\mathrm{d}^d\mathbf{q}}{(2\pi)^d} |m(\mathbf{q})|^2$$

$$\int \mathrm{d}^d\mathbf{x} (\nabla m)^2 = \int \frac{\mathrm{d}^d\mathbf{q}\mathrm{d}^d\mathbf{q}'}{(2\pi)^d} \delta(\mathbf{q}+\mathbf{q}')(i\mathbf{q})(i\mathbf{q}') m(\mathbf{q}) m(\mathbf{q}') = -\int \frac{\mathrm{d}^d\mathbf{q}}{(2\pi)^d} \mathbf{q}^2 |m(\mathbf{q})|^2$$

$$\int \mathrm{d}^d\mathbf{x} m(\mathbf{x}) = m(\mathbf{q}=0)$$

Plugging this into equation 5.10 yields the Fourier transform of the partition function:

$$\begin{aligned} Z &= \textstyle\prod_{\mathbf{q}} \left(V^{-1/2} \int \mathrm{d}m_{\mathbf{q}}\right) \exp\left\{-V \sum_{\mathbf{q}} \left(\frac{t+kq^2}{2}\right) |m_{\mathbf{q}}|^2 + V h m_{\mathbf{q}=\mathbf{0}}\right\} && (5.11) \\ &\overset{\text{cont.lim.}}{\to} \textstyle\int_0^\Lambda \mathcal{D}m(\mathbf{q}) \exp\left\{-\int \frac{\mathrm{d}^d\mathbf{q}}{(2\pi)^d} \left(\frac{t+kq^2}{2}\right) |m(\mathbf{q})|^2 + h m(\mathbf{0})\right\} && (5.12) \end{aligned}$$

where we went into the continuum limit and defined $\Lambda = 1/a$ as an ultra-violet cutoff, originating from the microscopic lattice spacing (Brillouin zone). The different q-modes do not couple to each other, i.e. there are no $m(q) \times m(q')$ terms in the action (except $\mathbf{q} = -vectq'$).

Now we carry out one RG step in three separate developments

1. "coarse-graining": **integrate out** short-wavelength degrees of freedom (e.g. every second spin in the 1d Ising model)
2. "rescaling": **rescale** the system to restore the microscopic cutoff (corresponds to $\uparrow\cdot\uparrow\cdot\uparrow \to \uparrow\uparrow\uparrow$, restoring the lattice spacing)
3. "renormalization": **renormalize** the fields, keeping a particular term in the Hamiltonian constant

**step 1** Short-length interaction corresponds to high frequencies. So, we integrate out all frequencies between $\Lambda/b$ and $\Lambda$. In $d$ dimensions, this is a momentum-shell. We write:

$$\begin{aligned} Z = \int \mathcal{D}m^<(\mathbf{q}) \exp\left\{-\int_0^{\Lambda/b} \frac{\mathrm{d}^d\mathbf{q}}{(2\pi)^d} \left(\frac{t+kq^2}{2}\right) |m^<(\mathbf{q})|^2 + h m^<(\mathbf{0})\right\} \\ \times \int \mathcal{D}m^>(\mathbf{q}) \exp\left\{-\int_{\Lambda/b}^{\Lambda} \frac{\mathrm{d}^d\mathbf{q}}{(2\pi)^d} \left(\frac{t+kq^2}{2}\right) |m^>(\mathbf{q})|^2\right\} \end{aligned}$$

The integral over "fast" or "short-range" degrees of freedom $m^>$ is

$$= \prod_{\Lambda/b<q<\Lambda} \sqrt{\frac{2\pi V}{t+kq^2}} \times \exp\left\{-\frac{1}{2}\int_{\Lambda/b}^{\Lambda} \frac{\mathrm{d}^d\mathbf{q}}{(2\pi)^d} \ln(t+kq^2)\right\} =: e^{-V\beta f_b(t,k)} \qquad (5.13)$$

Thus, after the first step of renormalization, the partition function is:

$$Z = \int \mathcal{D}m(\mathbf{q}) \exp\left\{ -\int_0^{\Lambda} \frac{\mathrm{d}^d\mathbf{q}}{(2\pi)^d} \left(\frac{t+kq^2}{2}\right) |m(\mathbf{q})|^2 + hm(\mathbf{0}) \right\} \tag{5.14}$$

$$= \int \mathcal{D}m^{<}(\mathbf{q}) \exp\left\{ -\int_0^{\Lambda/b} \frac{\mathrm{d}^d\mathbf{q}}{(2\pi)^d} \left(\frac{t+kq^2}{2}\right) |m^{<}(\mathbf{q})|^2 + hm^{<}(\mathbf{0}) \right\} \times e^{-V\beta f_b(t,k)} \tag{5.15}$$

**step 2** Rescaling the system $\mathbf{x}' = \mathbf{x}/b$ in real space is equivalent to a change of variable $\mathbf{q}' = b\mathbf{q}$ in momentum space. Additionally, we introduce the new fields $\tilde{m}(q') = m^{<}(q)$:

$$Z \propto \int \mathcal{D}\tilde{m}(\mathbf{q}') \exp\left\{ -\int \frac{\mathrm{d}^d\mathbf{q}'}{(2\pi)^d} b^{-d} \left(\frac{t+kb^{-2}q'^2}{2}\right) |\tilde{m}(\mathbf{q}')|^2 + h\tilde{m}(\mathbf{q}'=\mathbf{0}) \right\} \times e^{-V\beta f_b} \tag{5.16}$$

**step 3** Renormalize the fields $m' = \tilde{m}/z$:

$$Z \propto \int \mathcal{D}m'(\mathbf{q}') \exp\left\{ -\int_0^{\Lambda} \frac{\mathrm{d}^d q'}{(2\pi)^d} b^{-d} z^2 \left(\frac{t+kb^{-2}q'^2}{2}\right) |m'(\mathbf{q})|^2 + zhm'(\mathbf{q}'=\mathbf{0}) \right\} \times e^{-V\beta f_b} \tag{5.17}$$

This partition function can be considered the partition function of a new system with new couplings.

$$\begin{aligned} t' &= b^{-d}z^2 t \\ k' &= b^{-d-2}z^2 k \\ h' &= zh \end{aligned}$$

With this definition of $t', k', h'$, the partition function of the new system takes on the same form as the partition function of the old system. In technical language, the Gaussian model is **renormalizable**.

It is convenient to choose $z$ such that $k' = k$, so spatial fluctuations do not change their variance under steps 1-3: $z^2 = b^{d+2} \Rightarrow z = b^{1+d/2}$. Then, the couplings are:

$$t' = b^2 t \tag{5.18}$$

$$k' = k \tag{5.19}$$

$$h' = b^{1+d/2} h \tag{5.20}$$

The system has a fixed point at $t^* = 0, h^* = 0$. Since $b > 1$, the fixed point is unstable: small deviations in $t, h$ from $t^*, h^*$ grow under successive RG-steps.

Remarkably, the behaviour of the model near the critical point is entirely determined by the flow of the coupling constants. We begin with the correlation length, obeying

$$\xi(t,h) = b\xi(t',h') = b\xi(b^2 t, b^{1+d/2}h)$$

Choose $b$ such that $b^2 t = 1$. Then $b = t^{-1/2}$ and

$$\xi(t,h) = t^{-1/2}\xi(1, t^{-1/2-d/4}h) \overset{h=0}{\propto} t^{-1/2} \equiv t^{-\nu} \tag{5.21}$$

with the critical exponent $\nu = 1/2$.

- Let us have a look at the magnetization

$$\begin{aligned} m(t,h) &= \frac{1}{V}\frac{\partial \ln Z(t,h)}{\partial h} \\ &= \frac{1}{b^d V'}\frac{\partial h'}{\partial h}\frac{\partial \ln Z'}{\partial h'} \\ &= \frac{1}{b^d V'}\frac{1}{b^{-1(+d/2)}}\frac{\partial \ln Z'}{\partial h'} = b^{1+d/2-d}\left(\frac{1}{V'}\frac{\partial \ln Z'}{\partial h'}\right) \\ &= b^{1-d/2}m(t',h') = b^{1-d/2}m(b^2t, b^{1+d/2}h) \\ \overset{b^2t=1}{=} & \; t^{-\frac{1-d/2}{2}}m(1, t^{-1/2-d/4}h) \\ \overset{h=0}{\propto} & \; t^{-1/2+d/4} = t^{\beta} \end{aligned}$$

  The critical exponent is $\beta = -1/2 + d/4 \overset{d=3}{=} (-2+3)/4 = 1/4$.

- For the susceptibility, we yield

$$\begin{aligned} \chi(t,h) &= \frac{\partial m}{\partial h} = \frac{1}{V}\frac{\partial^2 \ln Z}{\partial h^2} \\ &= b^{-d}(b^{1+d/2})^2\left(\frac{1}{V'}\frac{\partial^2 \ln Z'}{\partial h'^2}\right) \\ &= b^2\chi(t',h') \propto t^{-1} = t^{-\gamma} \end{aligned}$$

  The critical exponent is $\gamma = 1$.

- The heat capacity is given by

$$\begin{aligned} c(t,h) &= \frac{1}{V}\frac{\partial^2 \ln Z}{\partial t^2} = b^{-d}(b^2)^2\left(\frac{1}{V'}\frac{\partial^2 \ln Z'}{\partial t'^2}\right) \\ &= b^{4-d}c(t',h') \propto t^{-\frac{4-d}{2}} = t^{-\alpha} \end{aligned}$$

  with the critical exponent $\alpha = \frac{4-d}{2} \overset{d=3}{=} 1/2$.

- Finally, we want to determine the field dependence of the magnetization at $t = 0$. Recapitulate

$$m(t,h) = b^{1-d/2}m(b^2t, b^{1+d/2}h)$$

  Choosing $b$ now such that $b^{1+d/2}h = 1$, we get $b = h^{\frac{1}{1+d/2}}$ and

$$\begin{aligned} m(t,h) &= h^{-\frac{1-d/2}{1+d/2}}m\left(h^{-\frac{2}{1+d/2}}t, 1\right) \\ &\overset{t=0}{\propto} h^{-\frac{1-d/2}{1+d/2}} = h^{1/\delta} \end{aligned}$$

  with a critical exponent $\delta = \frac{d/2+1}{d/2-1} \overset{d=3}{=} 5$.

Table 5.1 sums up these exponents and compares them with experiment and other methods.

Compared to the Landau-Ginzburg functional (3.4) the Gaussian model of course lacks a key term, the quartic $m^4(\mathbf{x})$. The behaviour of the couplings of the LG-hamiltonian can be worked out

| | experiment | MF | Gaussian | $\epsilon^1$ | $\epsilon^5$ |
|---|---|---|---|---|---|
| $\alpha$ | 0 - 0.14 | 0 | $1/2$ | $1/3$ | 0.109 |
| $\beta$ | 0.32 - 0.39 | $1/2$ | $1/4$ | $1/3$ | 0.327 |
| $\gamma$ | 1.3 - 1.4 | 1 | 1 | $1/4$ | 1.238 |
| $\delta$ | 4 - 5 | 3 | 5 | 4 | 4.786 |
| $\nu$ | 0.6 - 0.7 | - | $1/2$ | 1 | 0.631 |

Table 5.1: The table sums up the values of the critical exponents for different methods. $\epsilon$ corresponds to the expansion parameter of $\epsilon$-expansion, a method of pertubative renormalization group.

in detail, see for instance Kardar. Here we restrict ourselves to considering some power of $m$ as a small perturbation to Gaussian model

$$u_n \int \mathrm{d}^d x \, m^n(x) \tag{5.22}$$

We ask how this term behaves under rescaling $\mathbf{x}' = \mathbf{x}/b$ and renormalization $m' = zm$

$$u_n b^d z^{-n} \int \mathrm{d}^d x' \, m'^n(x') \tag{5.23}$$

Choosing $z$ again so that $b^{d-2}z^{-2} = 1$ (to keep the coupling constant of $(\nabla m)^2$ unchanged), leads to

$$u_n' = b^d \left(b^{d/2-1}\right)^{-n} u_n \tag{5.24}$$

$$u_4' = b^d \left(b^{4-2d}\right) u_4 = b^{4-d} u_4 \tag{5.25}$$

For $d > 4$, the coefficient of the $m^4$-term decreases under renormalization, for $d < 4$ it increases. In the latter case, the treatment of the $m^4$-term as a pertubation is inherently unstable. One speaks of the $m^4$-term as a **relevant** (**irrelevant**) term in the Hamiltonian for $d < 4$ ($d > 4$).

## 5.3 General Framework of GR

In the last chapter, we saw that all critical properties of a system can be described by the recursion relations $t' = t'(t, k, h)$ etc. We will outline a more formal procedure and emphasize this.

Assume a set of parameters $\{k\}$ that describe the same system.

- Of course we will choose $b > 1$.
- The renormalized parameters $\{k'\}$ are functions of the original ones $\{k\}$. I.e. They define a mapping $\{k'\} = R_b(\{k\})$ in parameter space.
- The mapping $\{k^*\} = R_b(\{k^*\})$defines a fixed point.

We can now investigate the stability of a fixed point by linearizing the recursion relation in his vicinity:

$$k_a' - k_a^* = \sum_b T_{ab}(k_b - k_b^*) \tag{5.26}$$

with $T_{ab} = \left.\frac{\partial k_a'}{\partial k_b}\right|_{k=k^*}$ (Taylor expansion).

Now, we diagonalize the Matrix $T_{ab}$ to get the eigenvalues:

$$\sum_a \phi_a^i T_{ab} = \lambda^i \phi_b^i = b^{y_i} \phi_b^i$$

Scaling variables $u_i \equiv \sum_a \phi_a^i (k_a - k_a^*)$ and transformation them yields:

$$u_i' = \sum_a \phi_a^i (k_a' - k_a^*) = \sum_a \phi_a^i \sum_b T_{ab}(k_b - k_b^*) = \sum_c b^{y_i} \phi_c^i (k_c - k_c^*) = b^{y_i} u_i \tag{5.27}$$

The eigenvalues are used to classify the parameters:

- $y_i > 0$, $u_i$ relevant
- $y_i < 0$, $u_i$ irrelevant
- $y_i = 0$, $u_i$ marginal

**Universality** is the prediction of RG that the thermodynamical properties of a system near a fixed point depend only on a small number of features and are not sensitive to the microscopic properties of this system. All systems that flow onto the same fixed points are called to be in the same universality class.

We will now define the reduced free energy (per spin): $f = -\frac{1}{N} \ln Z$. Then

$$Z = e^{-N f(\{k\})}$$

The partition function $Z$ of a system is invariant under RG

$$Z = \mathsf{Tr}_{\{s\}} e^{-\mathcal{H}[s]} = \mathsf{Tr}_{\{s'\}} e^{-\mathcal{H}'[s']} = Z'$$

Thus,

$$Z' = \mathsf{const} \times e^{-N' f'(\{k\})}$$

with $N' = b^{-d} N$ with

$$f(\{k\}) = g(\{k\}) + b^{-d} f(\{k'\})$$

$g(\{k\})$ represent the "analytic first modes" and

$$f(\{k\}) = b^{-d} f(\{k'\})$$

defines the singular behavior around fixed point.

For the universality class of the Ising model $u_t, u_h$ are relevant ($y_t, y_h > 0$) and

$$f(u_t, u_h) = b^{-d} f(\underbrace{u_t'}_{= b^{y_t} u_t}, u_h') = ... = b^{-nd} f(b^{n y_t} u_t, b^{-n y_h} u_h) \tag{5.28}$$

We choose the parameters such that $b^{n y_t} u_t = 1 = u_{t0}$. Self-consistency requires that $u_t, u_h$ do not grow out of the linear approximation.

$$\Rightarrow f(u_t, u_h) = |u_t|^{\frac{d}{y_t}} f(1, |u_t|^{-\frac{y_h}{y_t}} u_h)$$

and $f$ is a function of one variable only.

Assume that we are close to the fixed point and linearize $u_t, u_h$ according to previous analysis:

$$u_t = \frac{t}{t_0} + \mathcal{O}(t^2, h^2) \tag{5.29}$$

$$u_h = +\mathcal{O}(th) \tag{5.30}$$

$$\Rightarrow f(t,h) = \left|\frac{t}{t_0}\right|^{d/y_t} \Phi\left(\frac{h/h_0}{|t/t_0|^{y_h/y_t}}\right) \tag{5.31}$$

which is a funtion in one variable $x = h/t^{\Delta}$ with $\Delta = y_h/y_t$.

From this equation, we can calculate the *critical exponents*: We know that the correlation length will change under renormalization as

$$\xi(k) = b\xi(k') = ... = b^n \xi(k^{(n)}) \propto |u_t|^{-1/y_t} \propto |t|^{-1/y_t} \Rightarrow \nu = \frac{1}{y_t} \tag{5.32}$$

**specific heat**

$$\frac{\partial^2 f}{\partial t^2}|_{h=0} \propto |t|^{\frac{d}{y_t}-2} \Rightarrow \alpha = 2 - \frac{d}{y_t} \tag{5.33}$$

**magnetization**

$$\frac{\partial f}{\partial h}|_{h=0} \propto |t|^{\frac{d-y_h}{y_t}} \Rightarrow \beta = \frac{d-y_h}{y_t} \tag{5.34}$$

**susceptibility**

$$\frac{\partial^2 f}{\partial h^2}|_{h=0} \propto |t|^{\frac{d-2y_n}{y_t}} \Rightarrow \gamma = \frac{2y_n - d}{y_t} \tag{5.35}$$

**magnetization at $T_c$ as a function of $h$**

$$M = \frac{\partial f}{\partial h} = \left|\frac{t}{t_0}\right|^{\frac{d-y_h}{y_t}} \Phi'\left(\frac{h/h_0}{|t/t_0|^{y_h/y_t}}\right)$$

Look at $\Phi'(x)$. $M$ is finite for $t \to 0$. Thus, $x^{\frac{d}{y_h}-1} = \left(|t|^{-y_h/y_t}\right)^{\frac{d}{y_h}-1} = |t|^{-\frac{d-y_h}{y_t}}$

$$\Rightarrow M \propto h^{\frac{d}{y_h}-1} \Rightarrow \delta = \frac{y_h}{d-y_h} \tag{5.36}$$

You can proof the following scaling relations:

$$\begin{aligned} \alpha + 2\beta + \gamma &= 2 \\ \alpha + \beta(1+\delta) &= 2 \end{aligned}$$

that are valid for all system with a single relevant thermal eigenvalue and a single relevant symmetry breaking field.

# 6 Disordered Systems

Much of physics concerns homogeneous systems:

- homogeneous and isotropic space in classical mechanics and electrodynamics
- regular lattices (crystals) in solid state physics
- ferromagnets (all interactions between neighboring spins are of the same kind) in statistical mechanics

Common tools and concepts to treat homogeneous systems are:

- inverse space methods: Fourier transform
- statistical mechanics: transfer matrix
- mean field: each degree of freedom is in the same "environment" as all the others

Frequently, "homogeneity" and "order" are valid idealizations. A real crystal may have defects, impurities etc. but it may still be possible to use the perfect lattice as starting point for a pertubation analysis. However, there are systems with strong disorder where this approach is not possible:

- folding of a heteropolymer
- disordered magnetic system: "spin glasses".
  - magnetic impurities are placed in random positions on lattice sites
  - different distances between magnetic impurities lead to both ferromagnetic and antiferromagnetic bonds
  - ground state, excited states, thermodynamics depends on the disordered interactions
- amorphous solids: glasses, e.g. $SiO_2$ do not form a lattice many thermal and optical properties depend on the (disordered) long-range structure
- granular matter (sand): packings depend on the (disordered) particle shapes
- neural networks (brain): behaviour depends on (disordered) interaction between neurons
- economics
- theoretical computer science

It turns out that these systems share some aspects of their behaviour. This has given rise to a "statistical mechanics of disordered systems". The common themes of disordered systems are

- "frozen" low-temperature regime: each variable is constrained by configurations of all other variables $\Rightarrow$ What is the order parameter?

- frustration leading to many metastable states: for a ferromagnet, $s_i = 1 \forall i$ (or $-1$) satisfies all interactions in a plaquette. In a system where all couplings are positive, only one negative, at least one bond must be unsatisfied (frustrated). This situation arises wherever $\prod_{\text{plaquette}} J_{ij}$ is negative. In low-energy configuration, some bond will be unsatisfied. There is often a large number of ways of placing the unsatisfied bonds which leads to an unstable state. A rough energy landscape can lead to complex low-T behaviour.
  - slow equilibration as the system spends long time in metastable states
  - non-trivial equilibrium state with contributions from many different valleys

## 6.1 A mean-field model of spin glasses

In spin glasses not all bounds have the same couplings. This leads to interesting properties. Mezard et al. give an easy example: Consider three characters (A,B,C). In a tragedy there is a fight between two groups and the characters have to choose sides. If all characters like each other, the decision is easy: all will choose the same side. If A and B are friends and both do not like C, A and B will form one group and C the other. But if A,B,C all hate each other, frustation follows because enemies have to fight on the same side. If we assign a number $J_{AB}$ to each couple, $J_{AB} = +1$ denoting that A and B are friends, $J_{AB} = -1$ that they are enemies, the condition for frustration is

$$J_{AB} \cdot J_{BC} \cdot J_{CA} = -1$$

In our spin glass model, we choose Ising spin $s_i = \pm 1$ instead and have a Hamiltonian

$$\mathcal{H} = -\frac{1}{\sqrt{N}} \sum_{(i,j)} J_{ij} s_i s_j - \sum_i h_i s_i \tag{6.1}$$

where the first sum is over all pairs $i < j$ and $J_{ij}$ are drawn *randomly*, independently from a Gaussian distribution with mean 0 and variance 1.

This spin-glass model was proposed by D. Sherrington and S. Kirkpatrick in 1975 and solved by Giorgio Parisi in 1979. The proof of the correctness of this solution was done by Tallagrand in 2005!

For the Weiss ferromagnet, we use the magnetization $m$ as order parameter and have a symmetry breaking through a magnetic field $h$: If there is a freezing transition, what would be the order parameter? The symmetry breaking for ferromagnet

$$\lim_{h^+ \to 0^+} \lim_{N \to \infty} \frac{1}{N} \sum_i \langle s_i \rangle_{h^+} = m_+$$

For spin glasses, the approach is quite similar:

$$\lim_{\epsilon \to 0} \lim_{N \to \infty} \frac{1}{N} \sum_i \langle s_i \rangle^2_{\epsilon \underline{\sigma}} = q_{\text{EA}}$$

with $h_i = \epsilon \sigma_i$ and $\sigma_i \in \pm 1$.

As in the ferromagnet, we expect $\langle s_i \rangle$ to be zero in the absence of a magnetic field and at high temperatures and to be different from zero for low temperatures. Nevertheless, as the couplings $J_{ij}$ are all different, $\langle s_i \rangle$ will be sometimes positive, sometimes negative. We therefore decide to square it. The parameter is named after its inventors S. Edwards and P. Andersen.

There are several possibilities to investigate the SK-model:

- numerical simulation (exercise)
- set up self-consistent equations for $m_i$, given $J_{ij}$.
- model the disorder by random variables and average over the $J_{ij}$.

### 6.1.1 Modeling the disorder by random variables

So far, we considered a specific choice of $J_{ij}$. It turns out, different realizations of $J_{ij}$ have the same thermodynamics.

If an intensive quantity is self-averaging, $\frac{1}{N}\langle\mathcal{H}\rangle_{\beta,J_{ij}} = \langle\langle\langle\mathcal{H}\rangle_{\beta,J_{ij}}\rangle\rangle_{J_{ij}}$.

Look at the different averages:

$$\begin{aligned}\langle\langle\langle\mathcal{H}\rangle_{\beta,J_{ij}}\rangle\rangle_{J_{ij}} &= \langle\langle\frac{\mathsf{Tr}_{\ s}e^{-\beta\mathcal{H}[s,J]}\mathcal{H}}{\mathsf{Tr}_{\ s}e^{-\beta\mathcal{H}[s,J]}}\rangle\rangle_{J_{ij}} \\ &= \langle\langle-\partial_\beta\ln Z(\beta,J)\rangle\rangle_{J_{ij}} \\ &= -\partial_\beta\langle\langle\ln Z(\beta,J)\rangle\rangle_{J_{ij}} \\ &= -\partial_\beta\langle\langle\ln\mathsf{Tr}_{\ s}e^{-\beta\mathcal{H}}\rangle\rangle\end{aligned}$$

There is a distinction between $s_i$ and $J_{ij}$ variables: $J_{ij}$ in numerator and denominator are the same, $J_{ij}$ average after $\ln Z$ is taken. $s_i$ are averaged in both numerator and denominator independently, $s_i$ average is taken inside the $\ln$-function.

### 6.1.2 The "annealed" average

We will first reverse the order of averaging and compute $\ln\langle\langle Z\rangle\rangle_J$ (instead of $\langle\langle\ln Z\rangle\rangle_J$). As $J_{ij}$ is a Gaussian random variable, we can use the following disorder average, as well as the know phase space average and Boltzmann factor:

$$\text{Disorder average} \quad \prod_{i<j}\int\frac{\mathsf{d}J_{ij}}{\sqrt{2\pi}}e^{-\frac{1}{2}\sum_{i<j}J_{ij}^2} \tag{6.2}$$

$$\text{Phase space average} \quad \prod_{i=1}^{N}\sum_{s_i=\pm1} \tag{6.3}$$

$$\text{Boltzmann factor} \quad e^{-\frac{\beta}{\sqrt{N}}\sum_{i<j}J_{ij}s_is_j} \tag{6.4}$$

Thus, we get

$$\langle\langle Z\rangle\rangle_J = \langle\langle\mathsf{Tr}\ e^{-\beta\mathcal{H}}\rangle\rangle \tag{6.5}$$

$$= \prod_{i=1}^{N}\sum_{s_i=\pm1}\prod_{i<j}\left[\int\frac{\mathsf{d}J_{ij}}{\sqrt{2\pi}}e^{-\frac{1}{2}J_{ij}^2-\frac{\beta}{\sqrt{N}}J_{ij}s_is_j}\right] \tag{6.6}$$

$$= \prod_{i=1}^{N}\sum_{s_i=\pm1}\prod_{i<j}\left[e^{\frac{1}{2}\frac{\beta^2}{N}s_i^2s_j^2}\right] \tag{6.7}$$

$$= \prod_{i=1}^{N}\sum_{s_i=\pm1}e^{\frac{1}{2}\frac{\beta^2}{N}\frac{N(N-1)}{2}\times1} \tag{6.8}$$

$$= 2^Ne^{\frac{1}{4}\beta^2(N-1)} \tag{6.9}$$

As $\ln\langle Z\rangle \geq \langle\ln Z\rangle$ due to the concavity of the logarithm, the annealed average gives a upper *bound* on the quenched average.

### 6.1.3 The "quenched" average

We will now discuss the "quenched average", where we assume that the couplings $J_{ij}$ are constant on time scales over which the $s_i$ fluctuate.

We expand the logarithm

$$\ln Z = \lim_{n\to 0} \frac{Z^n - 1}{n} = \lim_{n\to 0} \partial_n Z_n \tag{6.10}$$

The idea is to calculate $\langle\langle Z^n \rangle\rangle$ for integer $n$ and *hope* to take $n \to 0$ successfully at the end. This is called "Replica-Trick".

$$\begin{aligned}
\langle\langle Z^n \rangle\rangle_J &= \langle\langle \underbrace{Z}_{a=1} \times \underbrace{Z}_{a=2} \times Z \times ... \rangle\rangle & (6.11)\\
&= \langle\langle \mathsf{Tr}_{\ s_1} e^{-\beta\mathcal{H}(J,s_1)} \times \mathsf{Tr}_{\ s_2} e^{-\beta\mathcal{H}(J,s_2)} \times ... \rangle\rangle_J & (6.12)\\
& & (6.13)
\end{aligned}$$

Again, we have:

$$\begin{aligned}
\text{Disorder average} \quad & \prod_{i<j} \int \frac{\mathrm{d}J_{ij}}{\sqrt{2\pi}} e^{-\frac{1}{2}\sum_{i<j} J_{ij}^2} & (6.14)\\
\text{Phase space average} \quad & \prod_{a=1,i=1}^{n,N} \sum_{s_i=\pm 1} & (6.15)\\
\text{Boltzmann factor} \quad & e^{-\frac{\beta}{\sqrt{N}} \sum_{i<j,a} J_{ij} s_i^a s_j^a} & (6.16)
\end{aligned}$$

Then,

$$\begin{aligned}
\langle\langle Z^n \rangle\rangle_J &= \prod_{a=1,i=1} \sum_{s_i^a=\pm 1} \prod_{i<j} \left[ \int \frac{\mathrm{d}J_{ij}}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} J_{ij}^2 - \frac{\beta}{\sqrt{N}} J_{ij} \sum_a s_i^a s_j^a \right\} \right] & (6.17)\\
&= \prod_{a=1} \sum_{s_i^a} \exp\left\{ \frac{1}{2} \frac{\beta^2}{N} \sum_{i<j} \sum_{a,b} s_i^a s_i^b s_j^a s_j^b \right\} & (6.18)
\end{aligned}$$

using $(\sum_a s_i^a s_j^b)^2 = \sum_{a,b} s_i^a s_i^b s_j^a s_j^b$

After taking the disorder average, we arrive at an effective model without disorder but with couplings between the replica.

The $i < j$ summation is awkward

$$\begin{aligned}
\sum_{i<j} s_i^a s_i^b s_j^a s_j^b &= \frac{1}{2} \sum_{i,j} s_i^a s_i^b s_j^a s_j^b - \frac{1}{2} \sum_i s_i^a s_i^b s_i^a s_i^b \\
&= \frac{1}{2} \left( \sum_i s_i^a s_i^b \right)^2 - \frac{1}{2} \sum_i 1 \\
\Rightarrow \exp\left\{ \frac{\beta^2}{2N} \sum_{i<j} \sum_{a,b} s_i^a s_i^b s_j^a s_j^b \right\} &= \exp\left\{ \frac{\beta^2}{4N} \sum_{a,b} \left( \sum_i s_i^a s_i^b \right)^2 - \frac{\beta^2}{4N} \sum_{i,a,b} 1 \right\}
\end{aligned}$$

We will now perform a HS-trafo

$$\int \frac{\mathrm{d}q}{\sqrt{2\pi}} e^{-\frac{1}{2}q^2 + xq} = e^{\frac{1}{2}x^2} \tag{6.19}$$

yielding

$$\exp\left\{ \frac{\beta^2}{4N} \sum_{a,b} \left( \sum_i s_i^a s_i^b \right)^2 \right\} = \prod_{a<b} \int \frac{\mathrm{d}q_{ab}}{\sqrt{2\pi/N\beta^2}} \exp\left\{ -\frac{N\beta^2}{2} \sum_{a<b} q_{ab}^2 + \beta^2 \sum_{a<b} q_{ab} \sum_i s_i^a s_i^b + \frac{\beta^2 n}{4} N \right\}$$

Collecting all terms, we have finally

$$\langle\langle Z^n\rangle\rangle = \prod_{a<b}\int \frac{\mathsf{d}q_{ab}}{\sqrt{2\pi/N\beta^2}} e^{-\frac{N\beta^2}{2}\sum_{a<b} q_{ab}^2 + \frac{\beta^2 n}{4}N} \underbrace{\prod_{a,i}\sum_{s_i^a} e^{\beta^2\sum_{a<b} q_{ab}\sum_i s_i^a s_i^b}}_{\left[\prod_a \sum_{s_a} e^{\beta^2\sum_{a<b} q_{ab}s_a s_b}\right]^N} \tag{6.20}$$

$$= \prod_{a<b}\int \frac{\mathsf{d}q_{ab}}{\sqrt{2\pi/N\beta}} \exp\{-\beta N n f_{\mathsf{eff}}(\{q_{ab}\})\} \tag{6.21}$$

with an effective free energy function

$$-n\beta f_{\mathsf{eff}}(\{q_{ab}\}) = -\frac{\beta^2}{2}\sum_{a<b} q_{ab}^2 + \beta^2 n/4 + \ln(\prod_a \sum_{s_a=\pm 1} e^{\beta^2\sum_{a<b} q_{ab}s_a s_b}) \tag{6.22}$$

comparable to the Weiss ferromagnet.

To take the limit of $n \to 0$, we need to simplify our approach to the saddle point. We make the assumption that a saddle point is "replica symmetric" meaning

$$q_{ab} = q \quad \forall a < b$$

This leads to $\sum_{a<b} q_{ab}^2 = \frac{n(n-1)}{2}q^2 \to -\frac{n}{2}q^2$ for $n \to 0$ and

$$\begin{aligned}
\prod_a \sum_{s^a=\pm 1} e^{\beta^2\sum_{a<b} q_{ab}s_a s_b} &= \prod_a \sum_{s^a=\pm 1} \exp\left\{\frac{1}{2}\beta^2 q \underbrace{\sum_{a,b} s^a s^b}_{(\sum_a s^a)^2} - \frac{1}{2}\beta^2 q\sum_a (s^a)^2\right\} \\
&\overset{\mathsf{HST}}{=} \int \frac{\mathsf{d}z}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \prod_a \sum_{s^a} e^{n\beta\sqrt{q}z\sum_a s^a} e^{-\frac{1}{2}\beta^2 q n} \\
&= \int \mathcal{D}z \left[\sum_s e^{\beta\sqrt{q}zs - \frac{1}{2}\beta^2 q}\right]^n \\
&= \int \mathcal{D}z \left[2\cosh(\beta\sqrt{q}z)e^{-\frac{1}{2}\beta^2 q}\right]^n \\
&= \int \mathcal{D}z \left(1 + n\ln(2\cosh(\beta\sqrt{q}z))\right) - e^{\frac{n}{2}\beta^2 q} + \mathcal{O}(n^2) \\
&= \exp\left\{\frac{n}{2}\beta^2 q + n\int \mathcal{D}z \ln(2\cosh(\beta\sqrt{q}z))\right\} + \mathcal{O}(n^2)
\end{aligned}$$

Thus,

$$-n\beta f_{\mathsf{eff}}(q) = n\frac{\beta^2}{4}(1-q)^2 + n\int \mathcal{D}z \ln(2\cosh(\beta\sqrt{q}z))$$

Evaluating $\frac{\mathsf{d}f_{\mathsf{eff}}}{\mathsf{d}q} = 0$:

$$-\frac{\beta^2}{2}(1-q) + \int \mathcal{D}z \tanh(\beta\sqrt{q}z)\frac{\beta z}{2\sqrt{q}} = 0$$

Integration by parts yields:

$$q = \int \mathcal{D}z \tanh^2(\beta\sqrt{q}z)$$

compare this with the Weiss ferromagnet $m = \tanh(\beta J m ...)$!

We find that

$$q_{ab}^{\text{saddle}} = \langle s^a s^b \rangle = \frac{1}{N} \sum_i \langle s_i^a s_i^b \rangle$$

This solution looks great but one problem is that we find an entropy $< 0$ at low temperature. For that, something has to be wrong with our solution. Maybe, we caught an unstable saddle point. Perhaps the whole replica symmetry ansatz is unstable and free energy $f(\{q_{ab}\})$ is not a minimum at $q_{ab}$ for $n \to 0$.

Parisi came up with the idea of replica symmetry breaking by introducing block matrices of size $m$.... With every step he gets better results, infinite steps solve the problem correctly.

# 7 Diffusion and growth processes

We used the conditional probability

$$p(x, t + \Delta t|z, t) = [1 - a(z, t)\Delta t]\, \delta(x - z) + \Delta t w(x, z) + \mathcal{O}(\Delta t^2)$$

to derive the master equation.

For a continuous state-space, this property leads to jumps (discontinuities) in the trajectory. Consider $a(z, t) = a = \text{const}$. Divide a time interval $T$ in many small intervals $\Delta t$. The probability not to jump away in a time interval $\Delta t$ is $(1 - a\Delta t)$, the probability not to jump away in the intervall $T$ is then $(1 - a\Delta t)^{T/\Delta t} \to e^{-aT}$. This means, the probability to make a jump at some time is $= 1$.

## 7.1 Fokker-Planck equation

Consider a continuous process

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} p(x, t + \Delta t|z, t) = 0$$

for $|x - z| > \epsilon > 0$ i.e. the probability for the final position $x$ to be finitely different from $z$ goes to zero faster than $\Delta t$ goes to zero. The first moment can be written as

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|<\epsilon} \mathrm{d}x\, (x - z)\, p(x, t + \Delta t|z, t) = A(z, t) + \mathcal{O}(\epsilon) \tag{7.1}$$

which is called a "drift". The second moment is given by:

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|<\epsilon} \mathrm{d}x\, (x - z)^2\, p(x, t + \Delta t|z, t) = B(z, t) + \mathcal{O}(\epsilon) \tag{7.2}$$

which is called "diffusion". We could go on with higher orders but in diffusion processes, higher moments vanish, e.g.

$$\begin{aligned}
\frac{1}{\Delta t} \left| \int_{|x-z|<\epsilon} \mathrm{d}x (x - z)^3 p(x, t + \Delta t|z, t) \right| &\leq \frac{1}{\Delta t} \int_{|x-z|<\epsilon} \mathrm{d}x |x - z| (x - z)^2 p(x, t + \Delta t|z, t) \\
&< \frac{\epsilon}{\Delta t} \int_{|x-z|<\epsilon} \mathrm{d}x (x - z)^2 p(x, t + \Delta t|z, t) \\
&\overset{\Delta t \to 0}{\to} \epsilon B(z, t) = \mathcal{O}(\epsilon)
\end{aligned}$$

Hence, knowing drift and diffusion are sufficient to characterize the system.

To derive the equation of motion of $p(x, t)$ consider the expectation value

$$\int_{-\infty}^{\infty} \mathrm{d}x\, p(x, t|y, t') f(x)$$

of some function $f(x)$ which is twice continuously differentiable.

$$\begin{aligned}\partial_t \int \mathrm{d}x\, f(x)\, p(x,t|y,t') &= \lim_{\Delta t\to 0} \frac{1}{\Delta t}\left\{\int \mathrm{d}x f(x)\left[p(x,t+\Delta t|y,t') - p(x,t|y,t')\right]\right\} \\ &= \lim_{\Delta t\to 0} \frac{1}{\Delta t}\left\{\int \mathrm{d}x \int \mathrm{d}z f(x) p(x,t+\Delta t|z,t) p(z,t|y,t')\right. \\ &\qquad \left. - \int \mathrm{d}z f(z) p(z,t|y,t')\right\}\end{aligned}$$

Let us expand $f(x)$ in a Taylor series around $z$

$$f(x) = f(z) + \frac{\mathrm{d}f}{\mathrm{d}z}(x-z) + \frac{1}{2}\frac{\mathrm{d}^2 f}{\mathrm{d}z^2}(x-z)^2 + |x-z|^2 R(x,z)$$

where we know that $R(x,z) \to 0$ as $|x-z| \to 0$. Inserting this yields:

$$\begin{aligned}&= \lim_{\Delta t\to 0} \tfrac{1}{\Delta t}\left\{\int \mathrm{d}x \int \mathrm{d}z \left[\tfrac{\mathrm{d}f}{\mathrm{d}z}(x-z) + \tfrac{1}{2}\tfrac{\mathrm{d}^2 f}{\mathrm{d}z^2}(x-z)^2 + |x-z|^2 R(x,z) + f(z)\right]\right. \\ &\qquad \left. \times\, p(x,t+\Delta t|z,t)\, p(z,t|y,t') - \int \mathrm{d}z\, f(z)\, p(z,t|y,t')\right\} + \mathcal{O}(\epsilon)\end{aligned}$$

Terms 1 and 2 give:

$$\int \mathrm{d}z \left(A(z)\frac{\mathrm{d}f}{\mathrm{d}z} + \frac{1}{2}B(z)\frac{\mathrm{d}^2 f}{\mathrm{d}z^2}\right) p(z,t|y,t') + \mathcal{O}(\epsilon)$$

Term 3 goes to zero for $\epsilon \to 0$:

$$\begin{aligned}\left|\frac{1}{\Delta t}\int \mathrm{d}x |x-z|^2 R(x,z) p(x,t+\Delta t|z,t)\right| &\le \tfrac{1}{\Delta t}\int \mathrm{d}x |x-z|^2 p(x,t+\Delta t|z,t) \max_{|x-z|<\epsilon} R(x,z) \\ &\overset{\Delta t\to 0}{\to} \left[B(z,t) + \mathcal{O}(\epsilon)\right] \times \max_{|x-z|<\epsilon} R(x,z) \overset{\epsilon\to 0}{\to} 0\end{aligned}$$

Terms 4+5 cancel:

$$\lim_{\Delta t\to 0} \int \mathrm{d}x\, p(x,t+\Delta t|z,t) = \mathcal{O}(\epsilon)$$

Thus, we get for $\epsilon \to 0$:

$$\begin{aligned}\partial_t \int \mathrm{d}z\, f(z)\, p(z,t|y,t') &= \int \mathrm{d}z \left[A(z)\frac{\mathrm{d}f}{\mathrm{d}z} + \frac{1}{2}B(z)\frac{\mathrm{d}^2 f}{\mathrm{d}z^2}\right] p(z,t|y,t') \\ &\overset{\text{P.I.}}{=} -\int \mathrm{d}z f(z)\frac{\partial}{\partial z}\left(A(z,t)\, p(z,t|y,t')\right) + \frac{1}{2}\int \mathrm{d}z f(z)\frac{\partial^2}{\partial z^2}\left(B(z,t)\, p(z,t|y,t')\right)\end{aligned}$$

As $f(z)$ is arbitrary, we get the *Chapman-Kolmogorov differential equation*

$$\boxed{\partial_t p(x,t|y,t') = -\frac{\partial}{\partial x}\left(A(x)p(x,t|y,t')\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(B(x)p(x,t|y,t')\right)} \tag{7.3}$$

Since $p(x,t) = \int \mathrm{d}x_0 p(x,t|x_0,t_0)p_0(x_0,t_0)$ and equation 7.3 is linear in $p(x,t|y,t')$, we get the *Fokker-Planck equation*:

$$\boxed{\partial_t p(x,t) = -\frac{\partial}{\partial x}\left(A(x)p(x,t)\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(B(x)p(x,t)\right)} \tag{7.4}$$

Some properties are:

- probability is preserved:

$$\partial_t \int_{-\infty}^{\infty} \mathrm{d}x\, p(x,t) = \int \mathrm{d}x \frac{\partial}{\partial x}\left[-A(x)p(x,t) + \frac{1}{2}\frac{\partial}{\partial x}\left(B(x)p(x,t)\right)\right] = [...]_{-\infty}^{\infty} = 0$$

  as $p(x,t) = 0$ at the boundaries.

- We can define a probability current from $\partial_t p = \partial_x j(x,t)$:

$$j(x,t) \equiv A(x)p(x,t) - \frac{1}{2}\partial_x(B(x)p(x,t))$$

**Example:** Consider $A(x,t) = 0$ and $B(x,t) = 1$ ("Wiener process"). This leads to

$$\partial_t p(x,t|x_0,t_0) = \frac{1}{2}\partial_x^2 p(x,t|x_0,t_0) \tag{7.5}$$

With the initial condition $p(x,t_0) = \delta(x - x_0)$ we can use the Fokker-Planck equation

$$\partial_t p(x,t) = \frac{1}{2}\partial_x^2 p(x,t) \tag{7.6}$$

Fourier transform yields

$$\begin{aligned}
\phi(k,t) &= \int \mathrm{d}x e^{ikx} p(x,t) \\
\partial_t \phi(k,t) &= \int \mathrm{d}x e^{ikx}\left(\frac{1}{2}\partial_x^2\right) p(x,t) = -\frac{1}{2}k^2\phi(k,t)
\end{aligned}$$

Integration gives

$$\phi(k,t) = e^{-\frac{1}{2}k^2(t-t_0)}\phi(k,t_0)$$

and $\phi(k,t_0) = e^{ikx_0}$ from initial conditions.

Inverse Fourier transform leads us to the result:

$$p(x,t) = \frac{1}{2\pi}\int \mathrm{d}k e^{-ikx} e^{-\frac{1}{2}k^2(t-t_0)} e^{ikx_0} \tag{7.7}$$

$$= \frac{1}{\sqrt{2\pi(t-t_0)}} e^{-\frac{(x-x_0)^2}{2(t-t_0)}} \tag{7.8}$$

This is a Gauss shaped distribution. Initially sharp peaked, it spreads in time, as you can see in figure 7.1.

Trajectories generated by this propagator $p(x,t|x_0,t_0)$ are continuous for

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|x|>\epsilon}\mathrm{d}x\,\frac{1}{\sqrt{2\pi\Delta t}}e^{-\frac{1}{2}\frac{x^2}{\Delta t}} \overset{x'=x/\sqrt{\Delta t}}{=} \lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|x'|>\epsilon/\sqrt{\Delta t}}\frac{\mathrm{d}x'}{\sqrt{2\pi}}e^{-\frac{1}{2}x'^2} = 0$$

as $\epsilon/\sqrt{\Delta t} \to \infty$.

But with probability one, trajectories are not differentiable:

$$\begin{aligned}
\Pr\left(\left|\frac{x(t+\Delta t)-x(t)}{\Delta t}\right| > k\right) &= \int_{|x|>k\Delta t}\frac{\mathrm{d}x}{\sqrt{2\pi\Delta t}}e^{-\frac{1}{2}\frac{x^2}{\Delta t}} \\
&= 2\int_{k\Delta t}^{\infty}\frac{\mathrm{d}x}{\sqrt{2\pi\Delta t}}e^{-\frac{1}{2}\frac{x^2}{\Delta t}} \\
&\overset{x'=x/\Delta t}{=} 2\int_{k\sqrt{\Delta t}}^{\infty}\frac{\mathrm{d}x'}{\sqrt{2\pi}}e^{-\frac{1}{2}x'^2} \overset{\Delta t\to 0}{\to} 1
\end{aligned}$$

This means, though the propagator is continuous, the trajectories are not differentiable at any point. As $\Delta x \sim \sqrt{\Delta t}$, typical velocities scale as $\frac{\Delta x}{\Delta t} \sim \frac{1}{\sqrt{\Delta t}}$ and become infinite large for small $\Delta t$.

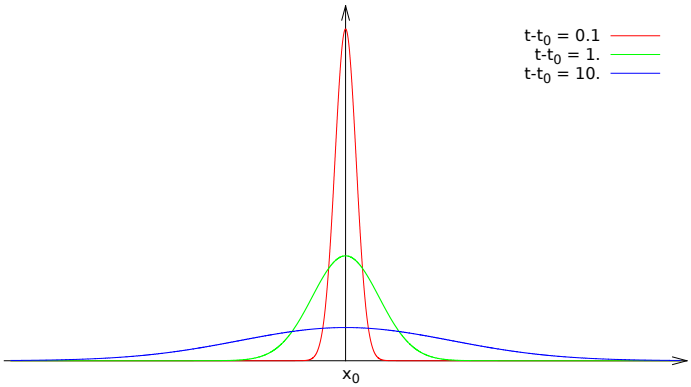



Figure 7.1: Wiener process: The figure shows the spreading of an initially sharp peaked distribution.

## 7.2 Stochastic differential equations

For the master equation, we specified $w(x, x')$. The equation depends on the full $w(x, x')$ but we knew how to generate sample paths.

For continuous processes, the first two moments of the propagator were sufficient. But how to generate trajectories?

*Langevin equation*:

$$\boxed{\partial_t x = a(x(t)) + b(x(t))\xi(t)} \tag{7.9}$$

where $\xi(t)$ is a random ("stochastic") function, $a(x(t))$ denotes the drift, $b(x(t))$ the diffusion. Two properties of $\xi(t)$ are:

- $\langle \xi(t) \rangle_\xi = 0 \; \forall t$
- $\langle \xi(t)\xi(t') \rangle_\xi = \delta(t - t')$

Apparently, $\langle \xi^2 \rangle_\xi \to \infty$. The infinite variance of $\xi(t)$ is key to diffusion. For the Wiener process $(a(x) = 0, b(x) = 1)$

$$\partial_t x = \xi(t) \Rightarrow x(t) = \int_0^t \mathrm{d}s \, \xi(s)$$

It follows:

$$\langle x(t) \rangle_\xi = \int_0^t \mathrm{d}s \; \langle \xi(s) \rangle_\xi = 0$$

$$\langle x(t)^2 \rangle = \langle \int_0^t \mathrm{d}s \int_0^t \mathrm{d}s' \, \xi(s)\xi(s') \rangle = \int_0^t \mathrm{d}s \, \mathrm{d}s' \; \langle \xi(s)\xi(s') \rangle = \int_0^t \mathrm{d}s \, \mathrm{d}s' \, \delta(s - s') = \int_0^t \mathrm{d}s = t$$

This shows that the infinite variance of $\xi$ produces the correct variance of $x$.

Interpret equation 7.9 as difference equation

$$(x_{i+1} - x_i) = \Delta t a(x_i) + b(x_i)\xi_i \Delta t \tag{7.10}$$

The variance of the second term is $(\frac{T}{\Delta t}) b^2 \sigma(\xi) (\Delta t)^2$. Thus, $\sigma(\xi)$ has to be $\propto 1/\Delta t$ (and have an infinite variance for $\Delta t \to 0$).

$\xi(t)$ with its infinite variance and no correlation remains a strange object. We investigate

$$W(t) = \int_0^t \mathrm{d}s \, \xi(s)$$

Assume this integral exists and is continuous. For some $t' > t$, we split the integral:

$$W(t') = \int_0^t \mathrm{d}s\, \xi(s) + \int_t^{t'} \mathrm{d}s\, \xi(s) = \lim_{\epsilon \to 0} \int_0^{t-\epsilon} \mathrm{d}s\, \xi(s) + \int_t^{t'} \mathrm{d}s\, \xi(s)$$

Now, the second term is statistically independent of the first term. ($W(t)$ is a Markov process.)

$$A(w_0, t) = \lim_{\Delta t \to 0} \langle w(t + \Delta t) - w_0 \rangle = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \langle \int_t^{t+\Delta t} \mathrm{d}s\, \xi \rangle = 0$$
$$B(w_0, t) = \lim_{\Delta t \to 0} \langle (w(t + \Delta t) - w_0)^2 \rangle = 1$$

This means, $W(t)$ is a Wiener process, yet $\xi(t) = \frac{\mathrm{d}W}{\mathrm{d}t}$ does not exist ($W(t)$ is nowhere differentiable). Thus, the Langevin equation must be interpreted in a difference form as equation 7.10 or in an integral form

$$\boxed{x(t) - x(0) = \int_0^t \mathrm{d}s\, a(x(s), s) + \int_0^t \mathrm{d}s\, \xi(s) b(x(s), s)} \tag{7.11}$$

### 7.2.1 Scaling of the Wiener process

Let us start with a Wiener process with explicit diffusion constant $D/2$:

$$\partial_t p(x, t) = \frac{D}{2} \partial_x^2 p(x, t) \tag{7.12}$$

Consider a simultaneous rescaling of length and time, $\tilde{x} = bx, \tilde{t} = b^z t$.

$$x(\tilde{x}) = \frac{\tilde{x}}{b}$$
$$t(\tilde{t}) = \frac{\tilde{t}}{b^z}$$

As well, $\partial_t = \frac{\mathrm{d}\tilde{t}}{\mathrm{d}t} \frac{\mathrm{d}}{\mathrm{d}\tilde{t}} = b^z \partial_{\tilde{t}}$, analogously for $\partial_x$. We then yield:

$$b^z \partial_{\tilde{t}} \tilde{p}(\tilde{x}, \tilde{t}) = \frac{D}{2} b^2 \partial_{\tilde{x}}^2 \tilde{p}(\tilde{x}, \tilde{t})$$
$$\Rightarrow \partial_{\tilde{t}} \tilde{p}(\tilde{x}, \tilde{t}) = \frac{D}{2} b^{2-z} \partial_{\tilde{x}}^2 \tilde{p}(\tilde{x}, \tilde{t})$$

with $\tilde{p}(\tilde{x}, \tilde{t}) = p(x(\tilde{x}), t(\tilde{t}))$. $\tilde{D} = b^{2-z} D$ is the new effective diffusion constant. For the choice $z = 2$, $D = \tilde{D}$.

## 7.3 Growth of interfaces

Some examples are

- snow flakes
- ballistic deposition of molecules on a surface
- flame front as a 2d sheet of paper is burnt
- water (coffee?) on a table cloth
- 1d flux lines in a semiconductor

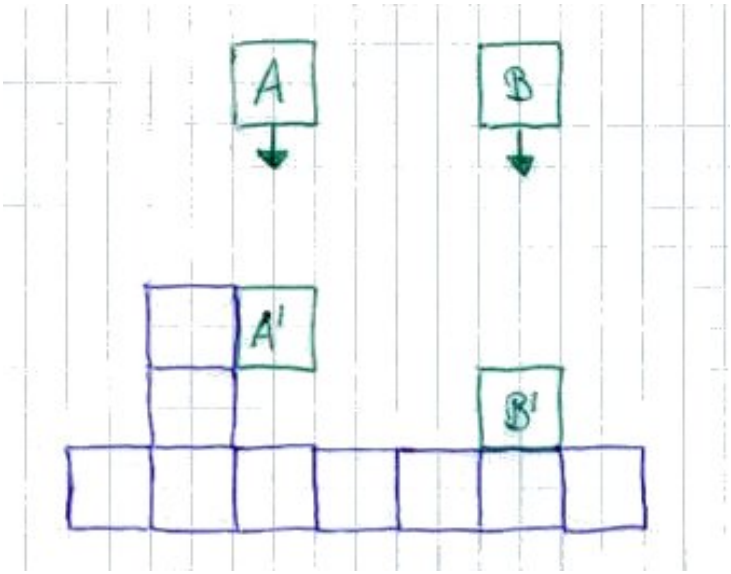



Figure 7.2: One example for the ballistic deposition model. Particle A is released randomly and sticks at the first site where it has an occupied nearest neighbor. Particle B can fall to the ground.

### 7.3.1 A simple model: Ballistic deposition (BD)

The following growth models work a bit as tetris: We release a particle from a randomly chosen point above the surface. In the BD-model with the nearest neighbor sticking rule, the particle sticks to the first site along its trajectory that has an occupied nearest neighbor. This is demonstrated in figure 7.2.

To describe the growth quantitatively, we introduce two functions:

- $h_i(t)$ denotes the height of the interface at time $t$ and lattice site $i$. The average height is given by

$$\overline{h}(t) = \frac{1}{L}\sum_{i=1}^{L} h_i(t) \tag{7.13}$$

- The standard deviation

$$w(t) = \sqrt{\frac{1}{L}\sum_{i=1}^{L}\left(h_i(t) - \overline{h}(t)\right)^2} \tag{7.14}$$

  characterizes the "roughness" of the surface.

Numerical experiments show that $w$ is a function of the systemsize $L$ and time $t$ and scales

$$w(L,t) \propto \begin{cases} t^{\beta} & t \ll t_x \\ L^{\alpha} & t \gg t_x \end{cases}$$

with $t_x \propto L^z$. This leads us to

$$w(L,t) \approx L^{\alpha}\Phi\left(\frac{t}{L^z}\right)$$

with a function $\Phi$ that becomes constant for $t \gg L^z$ and dominates otherwise (see figure 7.3). This function can be fitted. However, the exponents are not independent: E.g. at $t \approx t_x$, we have $w \propto t_x^{\beta} \propto L^{z\beta}$ and $w \propto L^{\alpha}$ what implies $\alpha = z\beta$.

The existence of power laws and a scaling function strongly reminds us of RG!

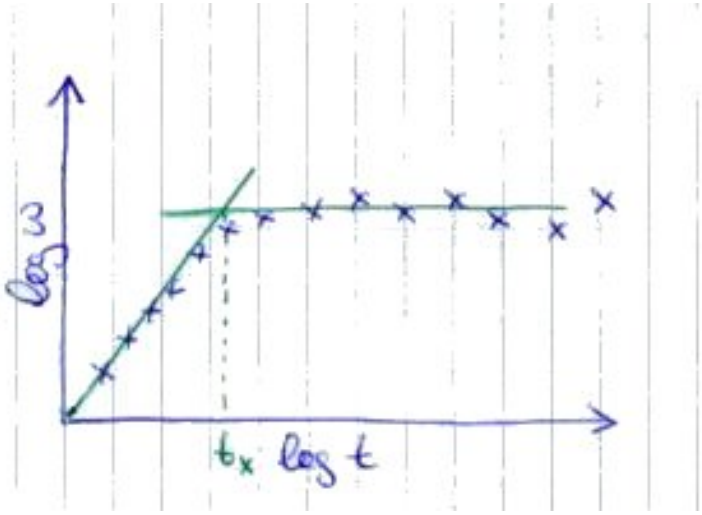



Figure 7.3: Growth of $w$ with time for the BD-model. We can distinguish two regimes: For $t \ll t_x$ we have a power law behavior, for $t \gg t_x$ we have saturation.

### 7.3.2 An even simpler model: Random deposition

Modify the BD rule so that a particle moves down until it touches a particle (or surface) below it. Then, there are no correlations of $h_i(t)$ accross the columns. $h_i$ grows by 1 when a particle is dropped above column $i$. The probability $P(h, N)$ that a given column has height $h$ after $N$ particles have been dropped follows a binomial distribution

$$P(h, N) = \left( \begin{array}{c} N \\ h \end{array} \right) p^h (1-p)^{N-h}$$

with $p = 1/L$, the probability that a given particle is released above a certain column.

Now, average height and surface roughness can be derived from the properties of the binomial distribution as

$$\overline{h} = \sum_{h=0}^{N} P(h, N) h = Np = \frac{N}{L} \equiv t \tag{7.15}$$

$$w^2 = \overline{(h - \overline{h})^2} = Np(1-p) = \frac{N}{L}\left(1 - \frac{1}{L}\right) \propto t = t^{2\beta} \tag{7.16}$$

Such an exact solution is not always possible.

The Langevin equation for $h_i(t)$ is

$$\partial_t h_i(t) = F_i(t) + \xi_i(t) \tag{7.17}$$

where $f_i(t)$ is the deterministic growth and $\xi_i(t)$ some noise (that will determine the statistical properties) with properties:

$$\langle \xi_i(t) \rangle = 0$$
$$\langle \xi_i(t) \xi_j(t') \rangle = 2D\delta_{ij}\delta(t - t')$$

For a simple example $F_i(t) = F$, we get

$$\begin{aligned}
h_i(t) &= Ft + \int_0^t \mathrm{d}s\, \xi(s) \\
\langle h_i(t) \rangle &= Ft + \int_0^t \mathrm{d}s\, \langle \xi_i(s) \rangle = Ft \\
w^2 &= \frac{1}{L} \sum_i h_i^2(t) - \underbrace{\left( \frac{1}{L} \sum_i h_i(t) \right)^2}_{\overline{h}^2} \\
\langle w^2 \rangle &= \frac{1}{L} \sum_i \left\langle \left( Ft + \int_0^t \mathrm{d}s\, \xi(s) \right)^2 \right\rangle - \overline{h}^2 \\
&= \frac{1}{L} \sum_i \left\langle \left( \int_0^t \mathrm{d}s\, \xi_i(s) \int_0^t \mathrm{d}s'\, \xi_i(s') \right) \right\rangle \\
&= \frac{1}{L} \sum_i \left( \int_0^t \mathrm{d}s \int_0^t \mathrm{d}s'\, 2D\delta(s - s') \right) \\
&= 2Dt \propto t^{2\beta}
\end{aligned}$$

what is the exact result from before. However, the Langevin equation is a flexible framework and can accomodate cases for which no exact solution exists.

### 7.3.3 Random deposition with surface relaxation

Relaxation of the surface was ignored in the RD model. We try to construct a simple Langevin equation including surface relaxation, compatible with the fundamental symmetries of the problem.

1. invariance under time translation $t \to t + \delta t$.
2. invariance under translation in the growth direction $h \to h + \delta h$.
3. invariance under translation, rotation, inversion in the plane perpendicular to the growth direction $x \to x + \delta x$, $\mathbf{x} \to \mathbf{R}\mathbf{x}$, $\mathbf{x} \to -\mathbf{x}$.
4. up/down symmetry of $h$ about the surface mean $h \to -h$ $(F = 0)$.

In general, we would expect a Langevin equation

$$\partial_t h(x,t) = \text{terms in } (x, t, h, h^2, h^3, ..., \nabla h, (\nabla h)^2, \nabla^2 h, ...) + \xi(x,t)$$

but terms linear in $x$ are incompatible with 3., $t$ with 1., $h$-terms violate 2., $\nabla h$ violates 3., $(\nabla h)^2$ violates 4. However, $\nabla^2 h$-terms are surviving! The resulting growth-equation is (in lowest order) the "Edwards-Wilkinson equation"

$$\partial_t h(x,t) = \nu \nabla^2 h(x,t) + \xi(x,t) \tag{7.18}$$

The statistics of the $\xi$ are the same as before but in a continuous form

$$\begin{aligned}
\langle \xi(x,t) \rangle &= 0 \\
\langle \xi(x,t) \xi(x',t') \rangle &= 2D\delta(t - t')\delta(x - x')
\end{aligned}$$

As the equation is linear in $h$, we will perform a Fourier transformation in space and time. Using

$$h(k,\omega) = \int \mathrm{d}x \int \mathrm{d}t\, e^{ikx-i\omega t} h(x,t)$$

we get

$$-i\omega h(k,\omega) = -\nu k^2 h(k,\omega) + \xi(k,\omega)$$

and

$$h(k,\omega) = \frac{\xi(k,\omega)}{\nu k^2 - i\omega} \tag{7.19}$$

We now calculate:

$$\begin{aligned}
\langle \xi(k,\omega)\rangle_\xi &= \int \mathrm{d}x \int \mathrm{d}t\, e^{ikx-i\omega t} \langle \xi(x,t)\rangle = 0 \\
\langle \xi(k,\omega)\xi(k',\omega')\rangle_\xi &= \int \mathrm{d}x\, \mathrm{d}x' \int \mathrm{d}t\, \mathrm{d}t'\, e^{ikx+ik'x'-i\omega t-i\omega' t'} \langle \xi(x,t)\xi(x',t',)\rangle \\
&= 2D \int \mathrm{d}x \int \mathrm{d}t\, e^{ix(k+k')-it(\omega+\omega')} \\
&= 2D(2\pi)^2 \delta(k+k')\delta(\omega+\omega') \\
\langle h(k,\omega)h(k',\omega')\rangle &= \frac{\langle \xi(k,\omega)\xi(k',\omega')\rangle}{(\nu k^2 - i\omega)(\nu k'^2 - i\omega')} \\
&= \frac{2D(2\pi)^2\delta(k+k')\delta(\omega+\omega')}{(\nu k^2 - i\omega)(\nu k'^2 - i\omega')}
\end{aligned}$$

Inverse FT yields:

$$\langle h(x,t)h(x',t')\rangle = \frac{D}{2\nu}|x-x'|\Phi\left(\frac{\nu|t-t'|}{|x-x'|^2}\right)$$

with the property that

$$\Phi(u) \begin{cases} \to \text{const} & u \to \infty \\ \to u^{1/2} & u \to 0 \end{cases} \tag{7.20}$$

that means at long times $\langle h(x,t)h(x',t')\rangle \propto |x-x'|$ and at small times $\langle h(x,t)h(x',t')\rangle \propto |t-t'|^{1/2}$ ($w \propto t^\beta$ with $\beta = 1/4$).

**Scaling of the EW equation** We perform a change of variables

$$\begin{aligned}
x &\to \tilde{x} = bx \\
h &\to \tilde{h} = b^\alpha h \\
t &\to \tilde{t} = b^z t
\end{aligned}$$

How does the EW-equation change?

$$b^{2-\alpha}\partial_{\tilde{t}}\tilde{h}(\tilde{x},\tilde{t}) = \nu b^{2-\alpha}\partial_{\tilde{x}}^2\tilde{h}(\tilde{x},\tilde{t}) + \ldots$$

What about the noise term $\xi$? Look at the autocorrelation

$$\begin{aligned}
\langle \xi(x,t)\xi(x',t')\rangle &= 2D\delta(x-x')\delta(t-t') \\
\Rightarrow \langle \xi(b^{-1}\tilde{x}, b^{-z}\tilde{t})\xi(b^{-1}\tilde{x}', b^{-z}\tilde{t}')\rangle_\xi &= 2D\delta(b^{-1}(\tilde{x}-\tilde{x}'))\delta(b^{-z}(\tilde{t}-\tilde{t}')) \\
&= 2Db^{1+z}\delta(\tilde{x}-\tilde{x}')\delta(\tilde{t}-\tilde{t}')
\end{aligned}$$

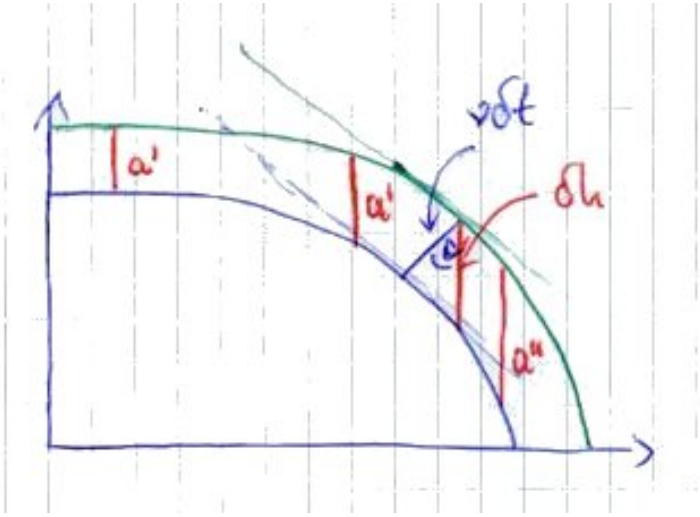



Figure 7.4: Origin of the non-linear term in the KPZ-equation

This motivates the definition

$$\tilde{\xi}(\tilde{x},\tilde{t})b^{\frac{1+z}{2}} = \xi(b^{-1}\tilde{x}, b^{-z}\tilde{t}) \tag{7.21}$$

and we can complete the EW-equation

$$b^{z-\alpha}\partial_{\tilde{t}}\tilde{h}(\tilde{x},\tilde{t}) = \nu b^{2-\alpha}\partial^2_{\tilde{x}}\tilde{h}(\tilde{x},\tilde{t}) + b^{\frac{1+z}{2}}\tilde{\xi}(\tilde{x},\tilde{t}) \tag{7.22}$$

$$\partial_{\tilde{t}}\tilde{h}(\tilde{x},\tilde{t}) = \nu b^{2-z}\partial^2_{\tilde{x}}\tilde{h} + b^{\frac{1-z}{2}+\alpha}\tilde{\xi}(\tilde{x},\tilde{t}) \tag{7.23}$$

We are free to choose $\alpha$ and $z$. Choosing $z = 2$ and $\alpha = 1/2$,

$$\partial_{\tilde{t}}\tilde{h} = \nu\partial^2_{\tilde{x}}\tilde{h} + \tilde{\xi} \tag{7.24}$$

leaves the EW equation invariant under scaling. Hence, the statistics of the surface on scales $h, x, t$ is the same as of $\tilde{h}, \tilde{x}, \tilde{t}$. $\Rightarrow$ The EW equation shows scale invariance and self-similarity.

Experiments give the exponents

$$\alpha = 0.47 \pm 0.02$$
$$\beta = 0.33 \pm 0.006$$

that deviate from our results $\alpha = 1/2$ and $\beta = 1/4$.

The deviations from the predicted critical exponents imply that there is something missing.

### 7.3.4 Kardar-Parisi-Zhang equation

We need to add a term to the EW-equation to describe the surface growth (see figure 7.4).

$$\tan\theta = \frac{a\nu\delta t}{\nu\delta t} = a$$
$$\tan\theta = \frac{\nabla h\delta x}{\delta x} = \nabla h$$
$$\delta h = \sqrt{(v\delta t)^2 + (\nabla h v\delta t)^2} = v\delta t(1 + (\nabla h)^2)^{1/2} \approx v\delta t(1 + \frac{1}{2}((\nabla h)^2 + ...)$$

This leads to the new equation

$$\partial_t h(x,t) = v\nabla^2 h + \frac{\lambda}{2}(\nabla h)^2 + \xi(x,t)$$

called KPZ (1986). This equation leads to the right exponents. However, things are not as easy as in the EW-model, as the equation is not linear anymore. The non-linear term $(\nabla h)^2$ puts the model into a universality class different from EW! Simple scaling ansatz fails. Here, we have to consider the cut-off and therefore have to integrate out degrees of freedom.