PHYSICS OF COMPLEX SYSTEMS

— LECTURE NOTES —

PRELIMINARY VERSION

APRIL 27, 2020

PROF. DR. HAYE HINRICHSEN

LEHRSTUHL FÜR THEORETISCHE PHYSIK III
FAKULTÄT FÜR PHYSIK UND ASTRONOMIE
UNIVERSITÄT WÜRZBURG, GERMANY

SUMMER TERM 2019
Contents

1. Epidemic spreading ........................................... 1
   1.1. The SIR model ........................................... 2

2. Many-particle systems on a lattice ......................... 5
   2.1. Classical cartoon of complex systems ..................... 5
      2.1.1. Configuration space .................................. 5
      2.1.2. Stochastic dynamics .................................. 7
      2.1.3. Probability distribution and Master equation ........ 8
      2.1.4. Example: Biased random walk on a one-dimensional chain 13
   2.2. The exclusion process on a one-dimensional lattice ....... 15
      2.2.1. The Exclusion Process ................................ 16
      2.2.2. Systematic construction and analysis of 1D lattice models 19
      2.2.3. From product states to matrix product states .......... 24
      2.2.4. Matrix product states ................................ 29
      2.2.5. Other algebraic structures ............................ 34

3. Entropy and Information .................................... 41
   3.1. Entropy as an information measure ........................ 41
   3.2. Entropy in Statistical Physics ............................ 52
   3.3. Thermostatics with conserved quantities and reservoirs .... 58
   3.4. Conserved quantities and external reservoirs ............. 61

4. Equilibrium phase transitions ................................ 65
   4.1. Ising model: Phenomenology ............................... 66
      4.1.1. Definition of Ising model ............................ 67
      4.1.2. Thermostatics of the Ising model .................... 67
      4.1.3. Ising mean field theory ............................... 70
      4.1.4. Scale invariance and critical behavior ............... 73
      4.1.5. Numerical simulation of the Ising model ............... 77
   4.2. Ising model: Transfer matrix and quantum limit .......... 81
      4.2.1. One-dimensional Ising model ........................ 81
      4.2.2. Two-dimensional Ising model ........................ 86
   4.3. Exact solution of the quantum XY chain ................... 89
   4.4. Conformal invariance .................................... 96
   4.5. Continuum limit of the Ising model ...................... 97

5. Nonequilibrium phase transitions .......................... 99
   5.1. Directed percolation ..................................... 99
      5.1.1. Directed bond percolation on a lattice ................ 99
      5.1.2. Absorbing states and critical behavior ............... 101
      5.1.3. The Domany-Kinzel cellular automaton ............... 104
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1.4</td>
<td>The contact process</td>
<td>106</td>
</tr>
<tr>
<td>5.1.5</td>
<td>The critical exponents of directed percolation</td>
<td>107</td>
</tr>
<tr>
<td>5.1.6</td>
<td>Mean field approximation</td>
<td>109</td>
</tr>
<tr>
<td>5.1.7</td>
<td>Phenomenological scale invariance</td>
<td>110</td>
</tr>
<tr>
<td>5.1.8</td>
<td>Universality</td>
<td>112</td>
</tr>
<tr>
<td>5.2.</td>
<td>Field-theoretic approach to Directed Percolation</td>
<td>113</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Langevin equation</td>
<td>113</td>
</tr>
<tr>
<td>5.2.2</td>
<td>The integral equation for DP</td>
<td>115</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Field-theoretic formulation of DP</td>
<td>116</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Multifractal properties of currents on directed percolation clusters</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td>Characterizing non-equilibrium transition by Yang-Lee zeroes in the complex plane</td>
<td>119</td>
</tr>
<tr>
<td>5.3.</td>
<td>Other classes of absorbing phase transitions</td>
<td>120</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Parity-conserving particle processes</td>
<td>121</td>
</tr>
<tr>
<td>5.3.2</td>
<td>The voter universality class</td>
<td>122</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Absorbing phase transitions with a conserved field</td>
<td>123</td>
</tr>
<tr>
<td>5.3.4</td>
<td>The diffusive pair contact process</td>
<td>124</td>
</tr>
<tr>
<td>5.4.</td>
<td>Epidemic spreading with long-range interactions</td>
<td>126</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Immunization and mutations</td>
<td>127</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Long-range infections</td>
<td>127</td>
</tr>
<tr>
<td>5.4.3</td>
<td>Incubation times</td>
<td>128</td>
</tr>
<tr>
<td>5.5.</td>
<td>Surface growth and non-equilibrium wetting</td>
<td>129</td>
</tr>
<tr>
<td>6.</td>
<td>Systems out of equilibrium</td>
<td>131</td>
</tr>
<tr>
<td>6.1.</td>
<td>Dynamics of subsystems</td>
<td>131</td>
</tr>
<tr>
<td>6.2.</td>
<td>Entropy production</td>
<td>137</td>
</tr>
<tr>
<td>6.3.</td>
<td>Fluctuation theorem</td>
<td>140</td>
</tr>
<tr>
<td>6.4.</td>
<td>Heat and work</td>
<td>144</td>
</tr>
<tr>
<td>7.</td>
<td>Spin glasses</td>
<td>147</td>
</tr>
<tr>
<td>7.1.</td>
<td>Glass phases</td>
<td>147</td>
</tr>
<tr>
<td>8.</td>
<td>Neural networks</td>
<td>161</td>
</tr>
<tr>
<td>8.1.</td>
<td>Biological background</td>
<td>161</td>
</tr>
<tr>
<td>8.2.</td>
<td>Magnetically inspired neural networks</td>
<td>164</td>
</tr>
<tr>
<td>8.3.</td>
<td>Hierarchical networks</td>
<td>170</td>
</tr>
<tr>
<td>A.</td>
<td>Appendices</td>
<td>173</td>
</tr>
<tr>
<td>A.</td>
<td>Perron-Frobenius Theorem</td>
<td>173</td>
</tr>
<tr>
<td>B.</td>
<td>Tensor products</td>
<td>174</td>
</tr>
<tr>
<td>C.</td>
<td>Properties of the Shannon entropy</td>
<td>177</td>
</tr>
<tr>
<td>D.</td>
<td>Moments and Cumulants</td>
<td>179</td>
</tr>
<tr>
<td>E.</td>
<td>The Schnakenberg Formula for Entropy Production</td>
<td>181</td>
</tr>
<tr>
<td></td>
<td>List of Acronyms</td>
<td>185</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>187</td>
</tr>
</tbody>
</table>
1. Epidemic spreading

This year is special. Universities all over the world have been closed. Entire countries are under lockdown. SARS-CoV-2 brings the world into a state of emergency. Therefore, in this year, our lecture notes start with a special chapter related to the pandemic.

Epidemic spreading is a very complex phenomenon, especially if the mode of transmission, the fraction of unreported infections, the incubation time and the immunization efficiency is only poorly understood. That makes it very difficult to predict the course of the coronavirus pandemic. As of March 31 2020, while I am writing this chapter, we have only partial data at our disposal.

A essential insight for understanding epidemic data is its multiplicative nature: Without control, a virus spreads exponentially in a fully susceptible population, that is, the number of infections \( I(t) \) grows as

\[
I(t) \propto e^{\beta t},
\]

where \( \beta \) plays the role of the spreading rate. When plotted on a linear scale, exponentially growing curves look horrible. But when plotted on a logarithmic scale, that is \( \ln I(t) \) versus \( t \), an exponential increase manifests itself as a straight line with slope \( \gamma \). Therefore, epidemic data should always be analyzed visually on a logarithmic scale.

**Remember:** Epidemic data should always be analyzed in a log-lin representation.

Have a look at Fig. 1.1 on the next page. It displays the number of confirmed infections and the number of deaths in three countries and worldwide on a log-lin plot. This is what we can see immediately:

- The curves for the number of confirmed infections and confirmed deaths (shown in the same color) have roughly the same shape and differ essentially by a vertical displacement, corresponding to a constant factor.
- The vertical distance differs from country to country, depending on the intensity of testing and the performance of the health system.
- Initially the curves display an almost exponential (straight) increase with the slope (growth rate)

\[
\beta \approx 0.3.
\]

- After some time, the slope of the curves decreases. This is caused by social distancing measures but it is also due to the increasing immunization of the population. For China, the curves even saturate at a constant value (and it is interesting to see whether this state will be stable in the future).
Epidemic spreading

In order to forecast the length and intensity of epidemic outbreaks, researchers have developed an uncountable number of models for epidemic spreading which capture the reality to different extent. Generally we expect that the accuracy of a model increases as we add more parameters, taking more and more available data into account. But this is not the approach favored by physicists. We as physicists are particularly fascinated by simple models of a complex reality, so simple that we can still make powerful predictions, but on the other hand complicated enough so that it captures the essential elements of the reality it want to describe.

In the following we want to discuss two very simple models for epidemic spreading, namely, the SIR model and the contact process. They are, of course, too simplistic to make realistic quantitative predictions, but nevertheless they can teach us a lot about the nature of epidemic spreading.

1.1. The SIR model

The Susceptible-Infected-Recovered model (SIR model) is a simple model for the dynamics of an epidemics by rate equations. It involves three degrees of freedom:

- $S(t) \in [0, 1]$ is the fraction of susceptible individuals in the population which are not yet immune and therefore susceptible to the infection.
- $I(t) \in [0, 1]$ is the fraction of the infected individuals in the population
- $R(t) \in [0, 1]$ is the fraction of recovered individuals in the population, assuming
1.1 The SIR model

perfect immunity.

The model is optimistic in so far as there are no deaths, that is, it assumes that all individuals recover in the end. In the case of SARS-CoV-2, where the mortality rate is of the order of a few percent, this is a good approximation.

Since there are no deaths and the three states, susceptible-infected-recovered, are mutually excluding, it is clear that we have the conservation law

\[ S(t) + I(t) + R(t) = 1. \]  

(1.2)

The SIR model is then defined by the following set of non-linear differential equations:

\[
\begin{align*}
\dot{S}(t) &= -\beta I(t)S(t) \\
\dot{I}(t) &= +\beta I(t)S(t) - \gamma I(t) \\
\dot{R}(t) &= +\gamma I(t).
\end{align*}
\]  

(1.3)

As can be checked easily, we have \( \dot{S}(t) + \dot{I}(t) + \dot{R}(t) = 0 \), compatible with the conservation law (1.2).

The three equations in (1.3) are so-called rate equations that you may have encountered in your courses on chemistry. The quantities depend only on time, i.e. the ansatz neglects any spatial information.

The three equations can be explained as follows:

* Infections occur on contact of a susceptible and an infected individual. Therefore the probability of a new infection is proportional to \( I(t) \) but also proportional to \( S(t) \). Each infection decreases \( S(t) \) and increases \( I(t) \). This explains the red terms in the equation. Note that the factor in front ensures that initially, when \( S(t) \approx 1 \), the fraction of infected individuals grows exponentially as \( e^{\beta t} \), where \( \beta \approx 0.3 \).

* The green terms account for the process of recovery. Compared to infections, this process is slow. For Corona patients the time scale for recovery is about three weeks. Therefore \( \gamma \approx 0.05 \) seems to be a reasonable choice.

Note that the first two equations are autonomous (they do not contain \( R(t) \)), so they can be solved independently and finally we can compute \( R(t) \) simply by means of the conservation law (1.2).

Although the first two differential equations look so simple, they are nonlinear [because they involve products of \( I(t) \) and \( S(t) \)] and thus an exact solution is difficult if not impossible. Nevertheless, it is possible to compute the height of the peak if infected individuals exactly and to integrate the equation numerically (see first exercise sheet). For example, for \( \epsilon = 0.01, \beta = 0.3 \) and \( \gamma = 0.05 \) we get the curves which are shown in the figure on the right.

The media usually report the total number of confirmed infections, irrespective of
whether these individuals have already recovered. In our model, this quantity has to be compared with \( I(t) + R(t) \). Fitting the numerically determined prediction to the corresponding empirical data for China and Germany measured at the beginning of the Corona crisis (until March 30, 2020), we get roughly the following prediction, as indicated by the dashed lines. When this course starts in summer 2020, we will see how good this prediction was.

Concerning the quality of the prediction, we should not be too optimistic. For example, it is assumed that all individuals finally recover, ignoring the potential loss of lives. The incubation time (time between receiving and spreading the infection) is ignored as well as the impact of immunization.

But most importantly, as mentioned before, the equations ignore any spatial resolution and possible regional inhomogeneities. Nevertheless this type of models in terms of non-linear differential rate equations is very popular among scientist. For example, a more elaborate version of a rate equation simulator can be found on the website [covidsim.eu](http://covidsim.eu) developed by ExploSys GmbH.

Figure 1.2.: Fitting the SIR prediction to the empirical data during the Corona crisis (March 30). Here it is assumed that the number of infections (upper curves) are roughly proportional to the number of lost lives (lower curves) so that we can apply the same type of fit.

Figure 1.3.: CovidSim Simulator based on rate equations
2. Many-particle systems on a lattice

2.1. Classical cartoon of complex systems

2.1.1. Configuration space

In the following let us consider an arbitrary physical system which consists of a large number of interacting building blocks. There is a large variety of such systems, including multi-particle systems such as gases and solids as well as more macroscopic phenomena such as granular flow or the dynamics of brain cells.

Ultimately every physical system follows the laws of quantum physics, which is the most fundamental physical theory of today. However, as we cannot even solve the Helium atom exactly within the framework of quantum mechanics, such a task will be practically impossible in the case of interacting complex systems. Therefore, we need a simplified description which is still able to account for the most salient features of complex systems.

It is well-known that the laws of quantum mechanics have far-reaching consequences. This includes the quantization of states in bound systems and the emergence of new phenomena such as quantum entanglement. However, usually these quantum features are not visible in our macroscopic world. The deep reason for the apparent classical behavior on macroscopic scales is the phenomenon of decoherence caused by ongoing interaction of the system with the environment. Roughly speaking, the environment permanently ‘measures’ the system, carrying away its quantum information and thereby destroying quantum effects in the system itself.

In many cases it is therefore sufficient to model a complex quantum-mechanical system as a classical one. Nevertheless this classical cartoon inherits some of the quantum-mechanical features. One of them is the assumption that the states of the system, which are now interpreted as classical configurations, are in some sense quantized so that they can be thought of as being discrete. Another one is that the dynamics between these states is of stochastic nature, i.e. the system jumps spontaneously from one state to the other, just as it would happen in quantum mechanics according to Fermi’s golden rule.

As an example let us consider molecular beam epitaxy (MBE), an experimental technique which is frequently used e.g. in Prof. Molenkamps lab. In such experiments one exposes a solid-state surface in a UV chamber to a beam of incident particles evaporating from a thermal source. Some of these atoms land on the surface, forming a deposition layer. The actual microscopic processes depend on various parameters such as the temperature and the involved materials. Typically the deposited atoms (called adatoms) diffuse for some time on the surface until they find another adatom, forming
the nucleus of an immobile deposition layer. This is shown schematically on the left side of Fig. 2.1.

With advanced microscopy techniques it is possible to track the motion of individual adatoms in real time. It turns out that the motion is discontinuous, i.e. the adatoms jump instantaneously from a given position on the lattice to a neighboring one. Moreover, these jumps occur spontaneously, similar to the clicks of a Geiger counter, indicating that the events of jumping are totally random. In fact, the jumps are not caused by quantum-mechanical tunneling, rather they are thermally induced by lattice vibrations. Since thermal fluctuations are fully chaotic, they can be considered as some kind of random noise, triggering diffusive moves of the adatoms every now and then.

In a minimalistic model one would of course not incorporate the chaotic thermal vibrations of the substrate in all detail, instead one would restrict the description of the model to the specific configurations of the adatoms with certain probabilistic transition rules. To this end we first have to specify the set of all possible configurations of the system, which will be denoted as $\Omega_{\text{sys}}$. This is shown schematically in the right panel of Fig. 2.1, where the configurations are represented as red dots.

The example of MBE nicely illustrates that the precise meaning of a configuration depends on the chosen level of abstraction in a given model. With respect to this level of abstraction, a configuration specifies the actual state of the system at a given time $t$ in all detail. Such a configuration, which accounts for all details on the chosen level, is often referred to as a microstate of the system. However, many authors simply use the term ‘state’ instead of ‘microstate’ or ‘configuration’. This can lead to confusion since the term ‘state’ is also used for probability distributions and sometimes for macroscopic states such as $(p, T)$ in thermodynamics. Therefore, we prefer to use the term ‘configuration’ throughout these lecture notes.

**Summary:** In these lecture notes we use the following terms in the following sense:

- **Configuration:** All details about a microscopic configuration of a model
- **Microstate:** Same as ‘configuration’
- **Configuration space $\Omega_{\text{sys}}$:** Set of all possible configurations
- **State:** Probability distribution (measure) on configurations.
- **Thermodyn. state:** Set of macroscopic thermodynamic variables such as $p, T$. 
2.1.2. Stochastic dynamics

Having characterized the configuration space, we have to find a suitable formulation for the dynamics of the system, i.e. we need appropriate rules how the system evolves in time. As outlined before, a large class of complex system evolves by instantaneous jumps from one configuration to the other. Denoting the individual configurations by \( c \in \Omega_{\text{sys}} \) such a jump (sometimes also called microscopic transition) can be denoted by \( c \rightarrow c' \).

The microscopic transitions give rise to a transition network in the configuration space. Note that in realistic systems this network is far from being fully connected because jumps between very different configurations are usually impossible. For example, in MBE it will never happen that ten adatoms hop simultaneously in one direction. In fact, assuming instantaneous jumps it is clear that only a single atom can move at a given time since the probability that two random time intervals of size zero coincide vanishes. This means that the transition network is usually very sparse.

Moreover, the likelihood of different microscopic transitions may be different, meaning that any microscopic transition \( c \rightarrow c' \) occurs randomly with a certain individual rate \( w_{c \rightarrow c'} \), where a zero rate indicates that the transition does not take place at all. Note that in general rates may be different in opposite directions, i.e.

\[
w_{c \rightarrow c'} \neq w_{c' \rightarrow c}.
\]  

At this point it is important to understand the difference between a probability \( p \) and a rate \( w \). While a probability is a numerical value \( p \in [0,1] \), a rate \( w \in \mathbb{R}^+_0 \) is defined as a probability per unit time and thus carries the dimension \([\text{time}]^{-1}\). A rate can be interpreted as follows: If \( dt \) denotes an infinitesimal time span, the probability for the transition \( c \rightarrow c' \) to happen just within this time span is given by \( w_{c \rightarrow c'} \, dt \).

In principle the rates could vary as a function of time and they could also depend on the specific history of the evolution. However, in what follows we will assume that the system under consideration has no memory. This is the so-called Markov assumption, stating that the future evolution of the system does not (statistically) depend on the history but only on the actual configuration of the system. Unless stated otherwise, all systems considered in this lecture are Markov processes which fulfill this assumption. Moreover, we will usually assume that the rates do not depend on time.

Summary: The definition of a complex stochastic system requires:

- the definition of a configuration space \( \Omega_{\text{sys}} \)
- the definition of a set of transition rates \( w_{c \rightarrow c'} > 0 \).
Ergodicity

A system with a transition network, where each configuration can be reached from any other configuration, is called ergodic. This means that the transition network does not decompose into separate disconnected pieces.

Such a decomposition emerges naturally in systems with conserved quantities. For example, let us consider a system that conserves the total number of particles (see Fig. 2.2). This means that transitions are only possible between configurations with the same particle number. Obviously such a system is not fully ergodic, instead it decomposes into a collection of dynamical sectors labeled by different particle numbers.

Another situation emerges in the presence of one-way transitions, i.e. \( w_{c \rightarrow c'} > 0 \) while \( w_{c' \rightarrow c} = 0 \). The presence of such transitions does not automatically break ergodicity. However, it may happen that certain configurations (or sets of configurations) can be reached but cannot be left. Such trapping states or sectors, which will also play an important role later in these lectures, are called absorbing.

Finally, it is possible that a transition network decomposes only effectively into several parts. In this case the parts are still mutually connected, but the likelihood for the system to go from one part to the other tends to zero. This happens, for example, in a ferromagnet: Here the macroscopic states of positive and negative magnetization are fully symmetric to each other. However, the system stays in one of the magnetization directions because it can only flip if all internal degrees of freedom flip almost simultaneously. This means that a magnetization flip is in principle possible but highly unlikely so that the two subspaces of the configuration space are effectively disconnected. This happens only in systems with a virtually infinite configuration space (the so-called thermodynamic limit and plays an important role in the context of spontaneous symmetry breaking.

2.1.3. Probability distribution and Master equation

So far we have seen that a classical stochastic complex system is given in terms of a set \( \Omega_{\text{sys}} \) of possible configurations \( c \in \Omega_{\text{sys}} \). The system evolves in time by instantaneous
2.1 Classical cartoon of complex systems

transitions $c \to c'$ which occur spontaneously like a radioactive decay with certain rates $w_{c \to c'} \geq 0$. The set of all configurations, the transition rates, and the initial configuration fully define the stochastic process under consideration.

As the time evolution is stochastic, it is of course impossible to predict the actual sequence of transitions, i.e., the stochastic trajectory is completely unpredictable. In fact, the only quantities that can be predicted in a stochastic process are probabilities. In this context an important object to study is the probability $P_c(t)$ to find the system at time $t$ in a certain configuration $c$. Obviously this probability distribution evolves continuously in time. Moreover it has to be normalized at any time, i.e.,

$$\sum_c P_c(t) = 1 \quad \forall t. \tag{2.2}$$

For a given set of rates one can easily figure out how this probability distribution evolves in time. On the one hand, the probability $P_c(t)$ will decrease due to the outgoing microscopic transitions from $c$ to somewhere else, and this loss will be proportional to the sum of the corresponding rates $\sum_{c'} w_{c \to c'}$ (the so-called escape rate) and proportional to the probability $P_c(t)$ itself. On the other hand, the probability $P_c(t)$ will increase due to the incoming microscopic transitions, and this gain will be proportional to the corresponding rates $w_{c' \to c}$ times the probability $P_{c'}(t)$ to find the system in the configuration from where the transition originates. Collecting all these loss and gain contributions, one arrives at a system of linear differential equations. This system of equations, called master equation, reminds of a continuity equation and describes the probability flow between different configurations in terms of gain and loss contributions:

$$\frac{\partial}{\partial t} P_c(t) = \sum_{c'} w_{c' \to c} P_{c'}(t) - \sum_{c'} w_{c \to c'} P_c(t). \tag{2.3}$$

In this equation the gain and loss terms balance one another so that the normalization condition (2.2) is preserved as time proceeds. Note that the factor $P_c(t)$ in the second term can be pulled out in front of the sum.

It is important to note that the rates $w_{c \to c'}$ carry the unit $[\text{time}]^{-1}$. Unlike probabilities, the numerical value of a rate depends on the unit of time and may by larger than 1. Rescaling all rates of a system by the same factor simply results in a change of the time scale. For example, multiplying all rates by 2 would mean that the whole process is simply running twice as fast.

**Formal properties of the master equation**
**Dirac notation:**

Let $N = |\Omega|$ be the number of configurations and let us enumerate the configurations by $c_1, \ldots, c_N$ in a specific order. The probability distribution may be thought of as a list $\{P_{c_1}(t), \ldots, P_{c_N}(t)\}$ of $N$ time-dependent non-negative functions which sum up to 1.

Let us now interpret this list of probabilities as a column vector in some linear vector space $V$, using the Dirac notation $|P_t\rangle \in V$:

$$|P_t\rangle = (P_{c_1}(t), \ldots, P_{c_N}(t))^T.$$ (2.4)

Since the probability distribution $P_c(t)$ is referred to as the *state* of the system, the vector $|P_t\rangle$ is called *state vector*. Likewise, the vector space $V$ is denoted as the *state space* of the system. Note that a state (i.e. vector in $V$) represents an ensemble of a large number of configurations with individual probabilities.

The dimension of the vector space $V$ is equal to the number of configurations of the system, i.e. it is generally an extremely high-dimensional space, comparable with a Hilbert space in quantum physics. However, in contrast to quantum mechanics, $V = \mathbb{R}^N$ is a real vector space. In addition, its vectors are only physically meaningful if all components in the canonical representation are positive and sum up to 1. This means that the actual state space is actually a small subset of $\mathbb{R}^N$ in the positive hyper-quadrant which has the form of a convex simplex.

Since the master equation is a linear system of differential equations, it is possible to write it in the compact form

$$\partial_t |P_t\rangle = -\mathcal{L}|P_t\rangle,$$ (2.5)

with a certain linear operator $\mathcal{L}$, where the minus sign is introduced as a matter of convenience, as will be explained below. The operator $\mathcal{L}$ is the so-called *Liouville operator* or *Liouvillian* which generates the temporal evolution of the system.

In order to represent the Liouvillian as a matrix we have to define a suitable basis. The most natural choice is the so-called *canonical configuration basis*, defined by unit vectors

$$|c_1\rangle = (1, 0, 0, \ldots, 0)^T$$
$$|c_2\rangle = (0, 1, 0, \ldots, 0)^T$$
$$\ldots$$
$$|c_N\rangle = (0, 0, 0, \ldots, 1)^T.$$ (2.6)

In the canonical configuration basis the Liouvillian is defined by the matrix elements

$$\langle c'|\mathcal{L}|c\rangle = -w_{c\rightarrow c'} + \delta_{c,c'} \sum_{c''} w_{c\rightarrow c''}.$$ (2.7)
2.1 Classical cartoon of complex systems

**Eigenmode decomposition:**

Obviously, a formal solution of this first-order differential equation is given by

$$|P_t\rangle = \exp(-\mathcal{L}t)|P_0\rangle,$$  \hspace{1cm} (2.8)

where $\exp(-\mathcal{L}t)$ is the **matrix exponential function** and $|P_0\rangle$ denotes the initial probability distribution at $t = 0$, the so-called **initial state**.

**Recall:**

Matrix exponential function:
The exponential function $e^A$ of a matrix $A$ is defined by the usual power series

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k$$

or by the product representation

$$e^A = \lim_{n \to \infty} \left(1 + \frac{A}{n}\right)^n.$$ Technically the matrix exponential function is most easily computed if one chooses the eigenbasis in which the operator $A$ is diagonal. In this basis one can compute $e^A$ simply by applying the exponential function to each diagonal element separately.

This allows us to express the solution $|P_t\rangle$ as a sum over exponentially varying eigenmodes. To this end we diagonalize the Liouville operator, solving the eigenvalue problem

$$\mathcal{L}|k\rangle = \lambda_k|k\rangle.$$ \hspace{1cm} (2.9)

Expanding the initial state as a linear combination of these eigenvectors by

$$|P_0\rangle = \sum_k a_k|k\rangle$$ \hspace{1cm} (2.10)

with certain coefficients $a_k$, the formal solution can be written as

$$|P_t\rangle = e^{-\mathcal{L}t} \sum_k a_k|k\rangle = \sum_k a_k e^{-\lambda_k t}|k\rangle = \sum_k a_k e^{-\lambda_k t}|k\rangle.$$ \hspace{1cm} (2.11)

This is the so-called **eigenmode decomposition** of the master equation.

**Remark:** Comparison with quantum mechanics

In quantum mechanics the situation is similar. Here we also have a linear evolution equation, namely, the Schrödinger equation $i\hbar \partial_t |\psi_t\rangle = H|\psi_t\rangle$. This equation can be solved formally by

$$|\psi_t\rangle = \exp(-\frac{i}{\hbar} H t)|\psi_0\rangle.$$ Diagonalizing the Hamiltonian by $H|n\rangle = E_n|n\rangle$ and expanding the initial state by $|\psi_0\rangle = \sum_n a_n|n\rangle$, the general solution can be written as $|\psi_t\rangle = \sum_n a_n e^{-\frac{i}{\hbar} E_nt}|n\rangle$, known as the **eigenmode decomposition** of the Schrödinger evolution equation.

**Probability conservation:**

The gain and loss terms in the master equation correspond to the non-diagonal and diagonal matrix elements of the Liouville operator, respectively. As the Liouville operator was introduced together with a minus sign in front, its non-diagonal elements are always negative while its diagonal elements are positive (c.f. Eq. (2.7)).

As mentioned above, the gain and loss terms balance one another in order to preserve the normalization of the probability distribution. To see this in the vector notation, let us introduce the row vector

$$|\Sigma\rangle = (1, 1, 1, \ldots, 1).$$ \hspace{1cm} (2.12)

Using this vector the normalization condition (2.2) can then be rewritten in the simple form

$$\langle \Sigma | P_t \rangle = 1 \hspace{1cm} \forall t.$$ \hspace{1cm} (2.13)
This immediately implies that
\[ \langle \Sigma | \mathcal{L} = 0, \] (2.14)
i.e. the sum over all columns in the matrix of the Liouville operator is zero. In the mathematical literature such matrices are known as stochastic operators or intensity matrices, having the property that all off-diagonal entries are real and positive and that the sum over each column of the matrix vanishes.

As can be seen, unlike a quantum-mechanical Hamiltonian, the Liouville operator of a stochastic system is in general neither Hermitean nor symmetric. Consequently the eigenvalues of an intensity matrix may be complex, indicating oscillatory behavior, but one can show that their real part is always non-negative. This ensures that all eigenmodes are either stationary or decaying exponentially with time:
\[ |P(t)\rangle = \sum_k a_k e^{-\lambda_k t} |k\rangle, \quad \text{Re}(\lambda_k) \geq 0. \] (2.15)

Moreover, the eigenvectors of a non-symmetric matrix are not necessary pairwise orthogonal, yet they can be used as a basis, justifying the decomposition in Eq. (2.10). In addition, the left and right eigenvectors (row and column eigenvectors) of \( \mathcal{L} \) do not have the same components. In quantum mechanics we are used to obtain the adjoint left eigenvector simply by taking the complex conjugate components of the right eigenvector. For a Liouvillian, these components of left and right eigenvectors are generally unrelated.

Intensity matrices obey the Perron-Frobenius theorem (see Appendix A). This theorem tells us that an intensity matrix has at least one eigenvector with the eigenvalue zero, and that the components of this vector behave like probabilities. If the eigenvalue zero is non-degenerate, this is the only state in the expansion (2.15) which survives in the limit \( t \to \infty \). This means that any system with a finite configuration space relaxes exponentially into a time-independent state \( |P_{\text{stat}}\rangle \) with
\[ \partial_t |P_{\text{stat}}\rangle = -\mathcal{L} |P_{\text{stat}}\rangle = 0. \] (2.16)

This is the so-called stationary state, which is also denoted as \( |P_0\rangle, |P_{\text{stat}}\rangle \) or \( |P_\infty\rangle \). Note that the term ‘stationary’ does not mean that the dynamics of the system is frozen, rather it may continue to jump between different configurations. Stationarity rather means that our knowledge about the system, the probability distribution \( P_c(t) \), does no longer depend on time. As we will discuss in the following chapter, stationarity must not be confused with thermal equilibrium, which turns out much a stronger constraint.

Solving a stochastic process basically means to diagonalize its Liouvillian. To obtain such a complete solution is often very difficult. Sometimes it is already useful to determine only the lowest-lying eigenvector which is the stationary state. In fact, most of the exact solution methods presented throughout the remainder of this chapter are solely concerned with finding the stationary state. In addition, it is sometimes of interest to find the second eigenstate with the smallest non-vanishing real part. The corresponding eigenvalue can be interpreted as the longest time scale, determining the asymptotic

---

1As will be discussed in Chapter xxx, the possibility of complex eigenvalues plays an important role in the context of chemical oscillations.
2.1 Classical cartoon of complex systems

relaxation properties.

As mentioned before the Liouvillian is generally not symmetric and thus the components of its left and right eigenvectors are generally different. The example of the stationary state nicely illustrates this difference between right and left eigenvectors. The stationary distribution of the system is given by the components of the right eigenvector, which is usually non-trivial to compute. How does the corresponding left eigenvector look like? The answer is very simple: If the spectrum is non-degenerate, it has to be the vector $\langle \Sigma |$ because this is by definition a left eigenvector with eigenvalue zero:

$$\mathcal{L} | P_\infty \rangle = 0 \iff 0 = \langle \Sigma | \mathcal{L} .$$

(2.17)

**Remark:** In comparing quantum mechanics with stochastic Markov processes, the following table may be helpful:

<table>
<thead>
<tr>
<th>Quantum theory</th>
<th>Stochastic Markov processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>complex Hilbert space $\mathbb{C}^n$</td>
<td>real probability space $\mathbb{R}^n$</td>
</tr>
<tr>
<td>complex amplitude vectors $</td>
<td>\psi\rangle$</td>
</tr>
<tr>
<td>quadratic normalization $\langle \psi</td>
<td>\psi \rangle = 1$</td>
</tr>
<tr>
<td>unitary evolution $\partial_t</td>
<td>\psi\rangle = \frac{i}{\hbar} \mathbf{H}</td>
</tr>
<tr>
<td>Hermitian evolution generator $\mathbf{H} = \mathbf{H}^\dagger$</td>
<td>vanishing column sum $\langle \Sigma</td>
</tr>
<tr>
<td>eigenvalue problem $\mathbf{H}</td>
<td>\phi\rangle = E</td>
</tr>
<tr>
<td>energy $E \in \mathbb{R}$</td>
<td>relaxation time $1/\lambda$, $\Re(\lambda) \geq 0$</td>
</tr>
<tr>
<td>ground state $\Leftrightarrow$ lowest $E$</td>
<td>stationary state $\Leftrightarrow \lambda = 0$</td>
</tr>
</tbody>
</table>

2.1.4. Example: Biased random walk on a one-dimensional chain

The probably simplest example of a stochastic Markov process in the framework described above is a (symmetric or biased) random walk on a one-dimensional lattice. In this model a single particle (the random walker) is located at site $k \in \mathbb{N}$ of a one-dimensional lattice. As time evolves it jumps spontaneously to site $k + 1$ with rate $w_R$ and to site $k - 1$ with rate $w_L$ (see Fig. 2.3). If $w_L = w_R$ the random walk is symmetric, otherwise it is said to be biased in one direction.

By constraining the dynamics to non-negative integers we introduce some kind of “wall” left from $k = 1$, where the random walker is reflected. As we will see below, this guarantees the existence of a stationary state, provided that the walk is biased towards the wall.

**Remark:** Simulation on a computer:

The total rate for a jump in either direction is $w_L + w_R$. Therefore, the average time elapsing between consecutive jumps is on average given by $\tau = 1/(w_L + w_R)$. Actually these events do not happen at regular time intervals, rather they occur randomly as in a radioactive decay. This means that the time intervals $\Delta t$ between consecutive events are randomly distributed according to a Poisson distribution (also known as shot noise), i.e.

$$P(\Delta t) \propto \frac{1}{\tau} e^{-\Delta t/\tau}.$$  

(2.18)

On a computer (see lecture on “Computational Physics”) such time intervals can simply be generated by setting

$$\Delta t := -\tau \ln(r),$$  

(2.19)

where $r$ is a uniformly distributed random number between 0 and 1. The leads us to the following update procedure (written here in C/C++ style):

if (rnd() < wr/(wr+wl)) k++;
else if (k>1) k--;  
t += -log(rnd())/(wl+wr);

Here double rnd(void) is a standard random number generator, int k is the position,  
double t is the actual physical time, and double wL,wR are the rates for jumps to the left  
and to the right. Take care that the random numbers equal to zero are excluded since other-  
wise the logarithm would diverge.

The master equation for this process takes the form

$$\partial_t P_k(t) = \frac{w_R P_{k-1}(t)}{\text{gain coming from left}} + \frac{w_L P_{k+1}(t)}{\text{gain coming from right}} - \frac{(w_L + w_R)P_k(t)}{\text{loss terms}}.$$  

(2.20)

As for the compact vector notation, we first note that there are infinitely many possible  
configurations $k \in \mathbb{N}$, hence the corresponding vector space of probability vectors is  
infinite-dimensional. In the canonical configuration basis $\{1, 2, 3, \ldots\}$ the Liouvillian  
has the matrix elements

$$L_{k,k'} = w_L \delta_{k,k'} - w_L \delta_{k,k'+1} - w_L \delta_{k,k'-1}.$$  

(2.21)

or, in matrix notation

$$L = \begin{pmatrix}
w_R & -w_L & \cdots & \cdots \\
-w_R & w_L + w_R & -w_L & \cdots \\
& -w_R & w_L + w_R & -w_L & \cdots \\
& & & & \ddots & \ddots & \ddots \\
& & & & & \ddots & \ddots & \ddots \\
& & & & & & \ddots & \ddots & \ddots \\
\end{pmatrix}.$$  

(2.22)

By construction, the vector $\langle \Sigma \rangle$ is a left eigenvector to the eigenvalue zero. As discussed  
above, the components of the corresponding right eigenvector are different and can be  
computed recursively as follows. Assume that the first component is given by some  
number $P_1$. Then the first line of the matrix tells us that

$$w_R P_1 - w_L P_2 = 0,$$  

(2.23)

hence $P_2 = bP_1$ with $b = w_R/w_L$. As can be verified easily, the following lines provide  
a recursion relation

$$P_k = bP_{k-1}.$$  

(2.24)

with the closed solution

$$P_k = P_1 b^{k-1}.$$  

(2.25)

In order to interpret these components as probabilities, they need to be normalized.
This can be done by computing the geometric series

\[ 1 = \sum_{k=1}^{\infty} p_k = p_1 \sum_{k=1}^{\infty} b^{k-1} = p_1 \frac{1}{1-b}, \tag{2.26} \]

determining the value of \( p_1 \). This lead to the result

\[ |P\rangle = (p_1, p_2, p_3, \ldots)^T, \quad p_k = (1-b)b^{k-1}. \tag{2.27} \]

The right eigenvector with these components just describes the stationary state of the process. As expected, this result is only meaningful for \( b < 1 \), where the random walk is biased to move towards zero.

**Remark:** For \( b = 1 \) (unbiased case) one obtains one half of a Gaussian distribution which spreads continuously so that the width is increasing as \( \sqrt{t} \). For \( b > 1 \) (biased to the right) one obtains a full Gaussian which is spreading as \( \sqrt{t} \) and moving away from the origin at constant velocity. In both cases the solution is not stationary.

### 2.2. The exclusion process on a one-dimensional lattice

In this lecture we are mainly concerned with stochastic many-particle systems on a discrete lattice. The lattice consists of a finite or an infinite number of sites which are arranged in a certain lattice geometry (see examples in Fig. 2.4). For the purpose of this lecture, we will be mostly concerned with one-dimensional chains of sites.

Each site can be either vacant or occupied by one or several particles. Models in which the number of particles per site is unrestricted are often referred to as *fermionic*. Contrarily, models with a restricted number of particles per site are called *fermionic*.

The particles residing on the lattice sites may be of different type. The particle species are usually enumerated by capital letters \( A, B, C \), and so forth. In the simplest case, there is only one species of particles involved, denoted by \( A \), and the lattice sites are restricted to carry at most one particle. This means that each site can be only in one of two possible local states, namely, vacant (\( \emptyset \)) or occupied by a single particle (\( A \)). This allows us to encode the presence of the particle by a binary digit, known as *bit*.
2.2.1. The Exclusion Process

Many-particle diffusion

The probably simplest example of a non-trivial many-particle system is the so-called \textit{simple exclusion process}. In contrast to the random walk model discussed above, where we had only a single random walker, this model describes a random walk of many particles on a lattice. It is assumed that each lattice site can be occupied by at most one particle. In other words, the particles ‘exclude each other’, which explains the name of the model. Thus the exclusion process falls into the simplest class of lattice models, where each site can be in only two different configurations, namely, vacant or occupied by an particle, denoted as ‘∅’ and ‘A’.

On a finite one-dimensional chain with \(L\) sites, the exclusion principle ensures that the configuration space of the model is finite, comprising \(2^L\) different configurations which may be represented like binary numbers:

\[
\begin{align*}
\text{∅∅...∅∅∅} & \quad \text{empty lattice} \\
\text{∅∅...∅∅A} & \\
\text{∅∅...∅A∅} & \\
\text{∅∅...∅AA} & \\
\text{∅∅...A∅∅} & \\
\vdots & \\
\text{AAA...AAA} & \quad \text{fully occupied lattice}
\end{align*}
\]

\textbf{Remark: Numerical bit coding techniques:}

The possibility to enumerate all configurations by binary numbers can be exploited as a very efficient implementation on a computer. For example, an \texttt{unsigned long int} allows us to describe a chain with \(L = 64\) sites. This so-called bit-coding is very efficient since the processor can handle 64 bits in parallel. However, the dynamical rules have to be implemented by bit manipulations which are not so easy to code for inexperienced programmers. For the purpose of this lecture it is much simpler to use a simple static array \texttt{int s[L]}, storing the values 0 (vacant) and 1 (occupied).

The dynamics of the exclusion process is defined in such a way that every particle jumps to the right (left) with rate \(w_R \ (w_L)\), provided that the target site is empty. This may be written symbolically as the following microscopic transition rules:

\[
\begin{align*}
A∅ & \xrightarrow{w_R} ∅A \\
∅A & \xrightarrow{w_L} A∅
\end{align*}
\]

Depending on whether the hopping rates are equal or unequal, the following acronyms are used in the literature:

\begin{itemize}
  \item \textbf{SEP} \quad \text{Symmetric exclusion process } w_L = w_R
  \item \textbf{ASEP} \quad \text{Asymmetric exclusion process } w_L \neq w_R
  \item \textbf{PASEP} \quad \text{Partially asymmetric exclusion process } w_L \neq w_R, \ w_R > 0, w_L > 0
  \item \textbf{TASEP} \quad \text{Totally asymmetric exclusion process } w_L \neq w_R \\text{with } w_R > 0, w_L = 0 \text{ or } w_R = 0, w_L > 0.
\end{itemize}
2.2 The exclusion process on a one-dimensional lattice

Although the model describes a simple many-particle random walk, it is important to note that the individual random walks of these particles are not totally independent since it is assumed that each site is occupied by at most one particle, effectively leading to some kind of repulsive interaction between the particles. As we will see, this is the reason why the exclusion process non-trivial.

**Special case of an exclusion process with only two sites:**

For a system with only two sites \((L = 2)\) the Liouville-Operator, represented in the configuration basis \((\emptyset \emptyset, \emptyset A, A \emptyset, AA)\), would be given by

\[
L^{(2)} = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & w_L & -w_R & 0 \\
0 & -w_L & w_R & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}.
\]  

(2.28)

This matrix has a three-fold degenerate eigenvalue 0, meaning that the stationary state of the system is not unique. This is of course not very surprising since the dynamical rules conserve the total number of particles. Therefore, in a system with only two sites the configuration space decomposes into three decoupled dynamical sectors with zero, one, and two particles.

In addition to the zero eigenvalues, the matrix \((2.28)\) has a single non-vanishing eigenvalue \(\lambda = w_L + w_R\), describing a relaxation mode. The corresponding eigenstate belongs to the sector with one particle because the empty and the fully occupied lattice are frozen configurations. In the one-particle sector the dynamics relaxes into a stationary state where the system flips randomly between \(A \emptyset\) and \(\emptyset A\). The normalized probability to find the system in one of these configurations is given by

\[
P_{\text{stat}}(A \emptyset) = \frac{w_L}{w_L + w_R}, \quad P_{\text{stat}}(\emptyset A) = \frac{w_R}{w_L + w_R}.
\]  

(2.29)

Obviously, if we decided to multiply both rates by the same number, then the system would simply switch faster back and forth between the two configurations, but the stationary probabilities given in \((2.29)\) will remain unaffected. In fact, rescaling all rates of a given model by the same common factor just changes the relaxation time scale while the stationary properties remain invariant. In the literature this freedom is often used to set one of the rates to 1. Another common choice is to choose reciprocal rates, in our case for example:

\[
w_R = q, \quad w_L = q^{-1}.
\]  

(2.30)

**ASEP with external particle input and output:**

Let us now couple the system to two external particle reservoirs, as sketched in Fig. 2.5. The left reservoir acts as a particle source, providing as many particles as needed at the leftmost site at rate \(\alpha\). Similarly, the rightmost reservoir acts like a sink, removing particles from the rightmost site at rate \(\beta\). Depending on the values of the rate, we intuitively expect a certain particle current flowing from left to right.
For simplicity let us first consider the special case of only two sites (see Fig. 2.6). In a two-site system, particle injection and removal at the boundaries amount to adding the transitions

- $\emptyset \emptyset \rightarrow A \emptyset$ with rate $\alpha$
- $\emptyset A \rightarrow AA$ with rate $\alpha$
- $\emptyset A \rightarrow \emptyset \emptyset$ with rate $\beta$
- $AA \rightarrow A \emptyset$ with rate $\beta$,

or, written more compactly using a ‘wildcard notation’:

- $\emptyset * \rightarrow A *$ with rate $\alpha$
- $* A \rightarrow * \emptyset$ with rate $\beta$.

In order to incorporate these rules into the Liouvillian, the two-site interaction matrix has to be extended by appropriate boundary terms, namely by

$$L = L^{(2)} + A^{(2)} + B^{(2)}$$

where

$$A^{(2)} = \begin{pmatrix} \alpha & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ -\alpha & 0 & 0 & 0 \\ 0 & -\alpha & 0 & 0 \end{pmatrix}, \quad B^{(2)} = \begin{pmatrix} 0 & -\beta & 0 & 0 \\ 0 & 0 & \beta & 0 \\ 0 & 0 & 0 & -\beta \\ 0 & 0 & 0 & \beta \end{pmatrix}$$

describe particle entry and exit at the left and right boundary, respectively. Note that each process appears twice in these matrices since the site which is not involved can be either vacant or occupied. Adding all contributions, the Liouvillian for the system shown in figure 2.6 reads

$$L = \begin{pmatrix} \alpha & -\beta & 0 & 0 \\ 0 & w_L + \alpha + \beta & -w_R & 0 \\ -\alpha & -w_L & w_R & -\beta \\ 0 & -\alpha & 0 & \beta \end{pmatrix}.$$  

(2.33)

Unfortunately, the eigenvalues of this $4 \times 4$ matrix cannot be computed easily. In particular, there are no longer three different stationary states because the ongoing particle influx from the left reservoir and the exit into the right one break particle conservation, thereby mixing the three sectors mentioned above. Consequently, the system has only...
2.2 The exclusion process on a one-dimensional lattice

![Diagram of simple exclusion process with two lattice sites. Site 1 is vacant while site 2 is occupied by a particle A, i.e., the system is currently in the configuration "∅A". In addition, the system may be connected to two external reservoirs. At site 1 particles may enter from the left reservoir at rate α while particles at site 2 may leave the system at rate β, moving to the right reservoir.]

one stationary state. As can be verified by inspection, this state is given by

$$|P_{\text{stat}}\rangle = \frac{1}{N}(\beta^2 w_R, \alpha \beta w_R, \alpha \beta (\alpha + \beta + w_L), \alpha^2 w_R),$$

where $N$ is a normalization constant given by

$$N = \alpha \beta (\alpha + \beta + w_L) + w_R (\alpha^2 + \alpha \beta + \beta^2).$$

Similarly we could construct the Liouvillian of a three-site chain by hand. Using the canonical configuration basis

$$\emptyset \emptyset \emptyset, \emptyset \emptyset A, \emptyset A \emptyset, \emptyset AA, A \emptyset \emptyset, A \emptyset A, AA \emptyset, AAA$$

we would arrive at the matrix

$$L^{(3)} = \begin{pmatrix}
\alpha & -\beta & 0 & 0 & 0 & 0 & 0 \\
0 & w_L + \alpha + \beta & -w_R & 0 & 0 & 0 & 0 \\
0 & -w_L & w_L + w_R + \alpha & -\beta & -w_R & 0 & 0 \\
0 & 0 & 0 & w_L + \alpha + \beta & 0 & -w_R & 0 \\
-\alpha & 0 & -w_L & 0 & w_R & -\beta & 0 \\
0 & -\alpha & 0 & -w_L & 0 & w_L + w_R + \beta & -w_R \\
0 & 0 & -\alpha & 0 & 0 & -w_L & w_R & -\beta \\
0 & 0 & 0 & -\alpha & 0 & 0 & 0 & \beta \\
\end{pmatrix}.$$  

Increasing $L$ further, a manual setup of the Liouvillian is certainly painful. Therefore, we need a mathematical structure which allows us to construct such matrices systematically.

2.2.2. Systematic construction and analysis of 1D lattice models

The tensor product $\otimes$

In practice it is rather inconvenient to construct such a matrix by hand. Therefore, it is useful to understand the general structure of a Liouvillian on a 1D chain in terms of tensor products and then to automatize its construction, using e.g. Mathematica®.

The tensor product of vector spaces, which is introduced in detail in Appendix B, work as follows. Suppose that $U$ and $V$ are two vector spaces with dimensions $n$ and $c$ respectively.

$$U \otimes V,$$
The tensor product combines these two vector spaces in such a way that the resulting vector space \( U \otimes V \) is \( n \cdot m \)-dimensional, i.e., the dimension is multiplicative under the formation of the tensor product. This operation, denoted as \( \otimes \), can be applied to both vectors and matrices. In Euclidean coordinate systems, the components of the resulting object are given by the product of all combinations of the components of the tensor factors. For example, for two vectors we have

\[
\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix},
\]

and likewise for matrices

\[
\begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae & af & be & bf \\ ag & ah & bg & bh \\ ce & cf & de & df \\ cg & ch & dg & dh \end{pmatrix}.
\]

In the context of one-dimensional lattices, we will associate with each lattice site a separate tensor factor. Let us, for example, reconsider the boundary terms of the Liouvillian discussed above. These boundary terms are represented by \( 4 \times 4 \) matrices, but it is clear that the actual process takes place at only one of the two lattice sites while the other lattice site remains unaffected. This circumstance is reflected by the fact that the boundary operators can be written as tensor products in the following form:

\[
\mathcal{A}^{(2)} = \left( \begin{array}{ccc} \alpha & 0 \\ -\alpha & 0 \end{array} \right) \otimes \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right), \quad \mathcal{B}^{(2)} = \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) \otimes \left( \begin{array}{ccc} 0 & -\beta \\ 0 & \beta \end{array} \right).
\]

or in short

\[
\mathcal{A}^{(2)} = \mathcal{A}^{(1)} \otimes 1, \quad \mathcal{B}^{(2)} = 1 \otimes \mathcal{B}^{(1)}.
\]

Here the left tensor factor \( \mathcal{A}^{(1)} \) describes a single-site particle creation process \( \emptyset \rightarrow A \), whereas the meaning of the second tensor factor, which is a \( 2 \times 2 \) unit matrix, is just that of “doing nothing”.

The insight that any local operation can be applied to a chain with many sites simply by forming a suitable tensor product with the unit matrices in all places that are not involved, allows us to systematically construct the Liouvillian for any lattice size, as will be explained in the following.

Remark: In practice it is useful to automatize the tensor product on an algebraic computer system (CAS). A very simple Mathematica® function, which can form the tensor product of both tensors and matrices, takes only a few lines:

```mathematica
Attributes[CircleTimes] = {Flat, OneIdentity};
CircleTimes[a_List /; VectorQ[a], b_List /; VectorQ[b]] := Flatten[KroneckerProduct[a, b]]; 
CircleTimes[a_List /; MatrixQ[a], b_List /; MatrixQ[b]] := KroneckerProduct[a, b];
```

For the tensor product \( |c\rangle = |a\rangle \otimes |b\rangle \) one simply writes
cvec = \{a_1, a_2, a_3\} \otimes \{b_1, b_2\}

where the symbol $\otimes$ can be obtained by typing \texttt{ESC} c * \texttt{ESC}.

### Setting up the Liouvillian:

In order to construct the Liouvillian formally, let us again consider a chain with 3 sites. Note that the elementary microscopic diffusion event $A \leftrightarrow \otimes A$ only involves a pair of adjacent sites along the chain. As for the Liouvillian, which plays the role of a time evolution generator, this means that $L$ is given by a sum of two-site operators:

$$L^{(3)} = L^{(2)\downarrow_{12}} + L^{(2)\downarrow_{23}} = L^{(2)\otimes 1 + 1 \otimes L^{(2)}}. \quad (2.41)$$

Here $L^{(2)\downarrow_{i,i+1}}$ is a $4 \times 4$ matrix describing hopping from site $i$ to $i + 1$ and vice versa which has exactly the same form as in Eq. \((2.28)\), while the ‘total’ Liouvillian $L$ on the left hand side is an $8 \times 8$ matrix acting on three sites.

Including the external reservoirs, we have to add boundary contributions of the form

$$L^{(3)} = L^{(2)\otimes 1 + 1 \otimes L^{(2)} + A^{(1)} \otimes 1 \otimes 1 + 1 \otimes 1 \otimes B^{(1)}}. \quad (2.42)$$

It is a straight-forward exercise to verify that $L^{(3)}$ coincides exactly the matrix given in Eq. \((2.36)\).

In the professional literature, nobody would write this in such a complicated way. Instead, the prevailing lazy notation is:

$$L^{(3)} = \sum_{i=1}^{2} L_{i,i+1} + A_1 + B_3 \quad (2.43)$$

with

$$L_{1,2} = L^{(2)\otimes 1}$$
$$L_{2,3} = 1 \otimes L^{(2)}$$
$$A_1 = A^{(1) \otimes 1 \otimes 1}$$
$$B_3 = 1 \otimes 1 \otimes B^{(1)}. \quad (2.44)$$

Now we can easily extend this formalism to an arbitrary number of sites. The Liouvillian can be expressed as

$$L = \sum_{i=1}^{L} L_{i,i+1} + A_1 + B_L. \quad (2.45)$$

with the bulk interaction

$$L_{i,i+1} := 1 \otimes (i-1) \otimes L^{(2)} \otimes 1 \otimes (L-i-1) \quad (2.46)$$
and the boundary matrices

\[ A_1 := A^{(1)} \otimes 1^{(L-1)} \]
\[ B_L := 1^{(L-1)} \otimes B^{(1)}. \]  

(2.47)

In full form the Liouvillian is given by

\[ \mathcal{L} = \sum_{i=1}^{L-1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \ldots \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & w_L & -w_R & 0 \\ 0 & -w_L & w_R & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \ldots \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \]

acting on sites 1, \ldots, i-1

\[ \rightarrow \]

acting on sites \(i+2, \ldots, L\)

\[ = \mathcal{L}^{(2)} \text{acting on sites } i,i+1 \]

\[ + \begin{pmatrix} \alpha & 0 \\ 0 & -\alpha \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \ldots \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \]

acting on the leftmost site

\[ + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \ldots \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & -\beta \\ \beta & 0 \end{pmatrix} \]

acting on the rightmost site

**Observables**

As in quantum mechanics, we are interested in measuring certain observables of interest. In quantum physics, observables are represented by Hermitean operators \( M = M^\dagger \) whose expectation value is given by

\[ \langle M \rangle = \langle \psi | M | \psi \rangle, \]

(2.48)

where \( | \psi \rangle \) is the actual quantum state normalized by \( \langle \psi | \psi \rangle = 1 \). Note that the state vector enters twice, both in the expectation value and the normalization, meaning that in this sense quantum mechanics has a quadratic structure. In stochastic Markov processes, however, the situation is different. Here the actual state of the system is described by a probability vector \( |P\rangle \). This vector is normalized linearly by \( \langle \Sigma | P \rangle = 1 \), and therefore it is no surprise that the same applies to measurements. More specifically, if \( M \) is a measurement operator, its expectation values given by

\[ \langle M \rangle = \langle \Sigma | M | \psi \rangle. \]

(2.49)

Opposed to quantum theory, the left vector \( \langle \Sigma \rangle \) is constant in this case. Therefore, as indicated by the curly bracket in the expression above, we do not need measurement operators in the framework of this theory, rather it is fully sufficient to work with measurement vectors.

A measurement vector is a row vector where each component corresponds to a particular configuration of the system. The meaning of these components is very simple: it just contains the value of what the measurement apparatus would measure in the respective configuration. For example, on a chain with two lattice sites, the probability of
finding a particle at the left site can be measured by applying the measurement vector

\[ \langle M | = (0, 0, 1, 1) = (0, 1) \otimes (1, 1). \] (2.50)

As shown above, the vector can be written as a tensor product of two local vectors. The left local vector \((0, 1)\) represents a measurement which responds with '1' to a particle and with a '0' otherwise. The right local vector \((1, 1)\), on the other hand, is neutral and has the simple meaning of 'measuring nothing'.

As a second example, let us consider a vector measuring the mean number of particles on the chain:

\[ \langle M | = (0, 1, 1, 2) = (0, 1) \otimes (1, 1) + (1, 1) \otimes (0, 1). \] (2.51)

As can be seen, this is simply the sum of the particle occupancy vector \((0, 1)\) at the leftmost and the rightmost site.

**Correlation functions:**

In the literature the notion of measurement vectors is very uncommon. Instead one uses *diagonal* measurement operators, where the diagonal elements are just the components of the corresponding measurement vector. These matrices are usually constructed as sums of the tensor products of the form

\[ \chi_i = 1 \otimes (i-1) \otimes (L-i), \quad \chi = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \] (2.52)

which gives the probability of finding a particle at site \(i\) (while the meaning of identity matrices is again that of "measuring nothing"). For example, the average density of particles \(\rho\) can be measured by applying the measurement operator

\[ \rho(t) = \frac{1}{L} \langle \Sigma |N|P(t) \rangle, \quad N = \sum_{i=1}^{L} \chi_i. \] (2.53)

Another important example is that of a two-point correlation function

\[ C_{ij}(t) = \langle \Sigma |\chi_i \chi_j|P(t) \rangle \]

\[ \chi_i \chi_j = 1 \otimes \ldots \otimes 1 \otimes \chi_i \otimes 1 \otimes \ldots \otimes 1 \otimes \chi_j \otimes 1 \otimes \ldots \otimes 1 \] (2.54)

which measures the probability that the lattice sites \(i\) and \(j\) are both occupied at the same time. However, finding positive values \(C_{ij}(t) > 0\) does not necessarily mean that the two sites are really correlated. For example, in a fully occupied system we have \(C_{ij} = 1\) although the two sites do not communicate with each other. For this reason it is meaningful to subtract the product of the expectation values at each site. This is known
as the “connected part” of the correlation function:

\[ C_{ij}^{\text{conn}}(t) := \langle \Sigma | \chi_i \chi_j | P(t) \rangle - \langle \Sigma | \chi_i | P(t) \rangle \cdot \langle \Sigma | \chi_j | P(t) \rangle. \]  

(2.55)

Note that the connected part of a correlation function is non-linear in the state vector \( |P(t)\rangle \). Therefore, it cannot be expressed in terms of a single observable \( C_{ij}^{\text{conn}} \) such that \( C_{ij}^{\text{conn}}(t) = \langle \Sigma | C_{ij}^{\text{conn}} | P(t) \rangle \).

### 2.2.3. From product states to matrix product states

**Product States:**

A (stationary) state \( |P\rangle \) is called *factorizable* or *product state* if it can be written as a tensor product of the form

\[ |P\rangle = \left( \frac{e_1}{d_1} \right) \otimes \left( \frac{e_2}{d_2} \right) \otimes \cdots \otimes \left( \frac{e_L}{d_L} \right). \]  

(2.56)

Note that if each of the tensor factors is rescaled individually, the scale factors can be pulled out in front of the expression, i.e.,

\[ \left( \frac{\lambda_1 e_1}{\lambda_1 d_1} \right) \otimes \left( \frac{\lambda_2 e_2}{\lambda_2 d_2} \right) \otimes \cdots \otimes \left( \frac{\lambda_L e_L}{\lambda_L d_L} \right) = \lambda_1 \lambda_2 \cdots \lambda_L |P\rangle. \]  

(2.57)

Therefore, without loss of generality, each of these factors can be normalized in such a way that its components add up to 1, i.e., \( d_i + e_i = 1 \). With this convention the resulting vector is already properly normalized, and \( d_i = p_i \) may be interpreted as the probability of finding a particle at site \( i \):

\[ |P\rangle = \left( \frac{1 - p_1}{p_1} \right) \otimes \left( \frac{1 - p_2}{p_2} \right) \otimes \cdots \otimes \left( \frac{1 - p_L}{p_L} \right). \]  

(2.58)

This argument shows that a product state involves only \( L \) degrees of freedom. Since a general normalized vector is characterized by \( 2^L - 1 \) independent degrees of freedom, it is clear that product spaces form only a small subspace in the full state space. To find out in what sense product states are special, we note that the connected part of the two-point correlation function vanishes:

\[ C_{ij}^{\text{conn}} = \langle \Sigma | \chi_i \chi_j | P(t) \rangle - \langle \Sigma | \chi_i | P(t) \rangle \cdot \langle \Sigma | \chi_j | P(t) \rangle = p_i p_j - p_i p_j = 0. \]  

(2.59)

The same applies to any connected \( n \)-point correlation function. Thus, we can conclude that

**Product states have no correlations.**

However, in reversed direction we have to be very careful. For example, if all connected two-point correlation functions vanish, this does not imply that the system is necessarily in a product state. In fact, there could be hidden three-point correlations although all two-point correlations vanish.
2.2 The exclusion process on a one-dimensional lattice

Remark: Note that product states in classical stochastic systems are analogous to non-entangled states in quantum physics, which also factorize into tensor products.

Stationary product states:

Since product states are simple but also very special, the question arises under which conditions a given system has a factorizing stationary state without any correlations. As an example let us again consider the asymmetric simple exclusion process (ASEP) without boundary terms. As discussed above, the time evolution of this model is generated by the Liouvillian

\[
\mathcal{L} = \sum_{i=1}^{L-1} \mathcal{L}^{(2)}_{i,i+1},
\]

where \( \mathcal{L}^{(2)} \) is the \( 4 \times 4 \) matrix defined in Eq. (2.28). Stationarity means that

\[
\mathcal{L}|P\rangle = 0.
\]

Obviously, there are two possibilities how this can happen:

- It may happen that each of the summands in Eq. (2.60) applied to the product state vanishes individually, i.e. \( \mathcal{L}^{(2)}_{i,i+1}|P\rangle = 0 \) for all \( i = 1, \ldots, L - 1 \).

- It may happen that the individual summands do not vanish separately, but the terms cancel each other so that the total sum gives zero. This requires a nontrivial cancellation mechanism between adjacent pairs of sites.

Let us start with the first case, where each term is supposed to vanish separately. This means that

\[
\mathcal{L}^{(2)} \left[ \begin{pmatrix} 1 - p_i \\ p_i \end{pmatrix} \otimes \begin{pmatrix} 1 - p_{i+1} \\ p_{i+1} \end{pmatrix} \right] = 0 \quad \forall i = 1, \ldots, L - 1
\]

or equivalently

\[
\begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & w_L & 0 & 0 \\
0 & -w_L & w_R & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
(1 - p_i)(1 - p_{i+1}) \\
(1 - p_i) p_{i+1} \\
p_i (1 - p_{i+1}) \\
p_i p_{i+1}
\end{pmatrix}
= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.
\]

Since \( \mathcal{L}^{(2)} \) has four lines one obtains in principle four different equations. In the present case, however, the first and the last line vanish identically while the second and third line are linearly dependent. Therefore, in the case of the ASEP one obtains only a single equation, namely

\[
w_L (1 - p_i) p_{i+1} = w_r p_i (1 - p_{i+1}),
\]

leading to the recursion relation

\[
p_{i+1} = f(p_i) = \frac{w_r p_i}{w_L (1 - p_i) + w_r p_i}.
\]
With the help of Mathematica®, we can easily convince ourselves that this recursion relation is solved by the exact solution

$$p_i = \frac{w_R p_1}{w_L (1 - p_1) + w_R p_1}. \quad (2.65)$$

As shown in Fig. 2.7, this allows us to compute density profiles, describing the stationary state of the system. However, we should keep in mind that the dynamics of the ASEP conserves the number of particles in the system, decomposing the configuration space into several independent sectors with a fixed number of particles, but the product state computed above actually describes a probabilistic ‘superposition’ over many of such sectors. In order to compute the stationary state for a given number of particles, it is therefore necessary to project this solution onto the corresponding sector (left as an exercise to the reader).

Let us now turn to the generalized model with external reservoirs and let us find out under which conditions it possesses a factorizable stationary state. If all terms in the sum of the Liouvillian vanish separately, this would imply that its action on the first two sites of the system obeys the equation

$$(L^{(2)} + A \otimes \mathbb{1} + \mathbb{1} \otimes B) \left[ \begin{pmatrix} 1 - p_1 \\ p_1 \end{pmatrix} \otimes \begin{pmatrix} 1 - p_2 \\ p_2 \end{pmatrix} \right] = 0, \quad (2.66)$$

or, equivalently

$$\begin{pmatrix} \alpha & -\beta & 0 & 0 \\ 0 & w_L + \alpha + \beta & -w_R & 0 \\ -\alpha & -w_L & w_R & -\beta \\ 0 & 0 & 0 & \beta \end{pmatrix} \left( \begin{pmatrix} 1 - p_1 \\ p_1 \end{pmatrix} \otimes \begin{pmatrix} 1 - p_2 \\ p_2 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \quad (2.67)$$

As can be seen, we now have three independent lines in the matrix while the vector involves only two degrees of freedom (namely, $p_1$ and $p_2$). This means that for generic values of the rates the system of equations is over-determined. This demonstrates that...
it is no longer possible to find stationary states for which all summands of the Liouvillian vanish separately. However, as we will see in the next paragraph, stationary products states are still possible thanks to an elaborate compensation mechanism between adjacent terms.

**Zipper-like compensation mechanism for stationary product states:**

In the following we demonstrate that the asymmetric exclusion process (ASEP) coupled to external reservoirs still admits factorizable solutions thanks to a compensation mechanism between adjacent sites. Such products states are homogeneous, i.e., they consist of identical tensor factors:

\[
|P_{\text{stat}}\rangle = \frac{1}{\mathcal{N}} (e_d) \otimes (e_d) \otimes \ldots \otimes (e_d),
\]

where \( \mathcal{N} = (d + e)^L \) is a normalization factor. The notation with letters \( e \) and \( d \) is frequently used in the literature and corresponds to the probabilities \( 1 - p \) and \( p \).

First, we note that the application of \( A \) at the leftmost site yields the vector

\[
A \begin{pmatrix} e_d \\ d \end{pmatrix} = \begin{pmatrix} \alpha & 0 \\ -\alpha & 0 \end{pmatrix} \begin{pmatrix} e_d \\ d \end{pmatrix} = +\alpha e \begin{pmatrix} 1 \\ -1 \end{pmatrix}.
\]

Remarkably, the application of \( B \) at the rightmost site yields the same vector, although with a different prefactor:

\[
B \begin{pmatrix} e_d \\ d \end{pmatrix} = \begin{pmatrix} 0 & -\beta \\ 0 & \beta \end{pmatrix} \begin{pmatrix} e_d \\ d \end{pmatrix} = -\beta d \begin{pmatrix} 1 \\ -1 \end{pmatrix}.
\]

Clearly, the two vectors on the right hand side coincide up to a minus sign, provided that \( e\alpha = d\beta \). Thus let us from now on choose \( e \) and \( d \) in such a way that that

\[
e\alpha = d\beta = 1.
\]

As an example, let us investigate a system with only two sites \( (L = 2) \). With the choice \( e\alpha = d\beta = 1 \), the action of the boundary terms is given by

\[
A_1 \left[ \begin{pmatrix} e_d \\ d \end{pmatrix} \otimes \begin{pmatrix} e_d \\ d \end{pmatrix} \right] = + \begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} e_d \\ d \end{pmatrix}.
\]

\[
B_2 \left[ \begin{pmatrix} e_d \\ d \end{pmatrix} \otimes \begin{pmatrix} e_d \\ d \end{pmatrix} \right] = - \begin{pmatrix} e_d \\ d \end{pmatrix} \otimes \begin{pmatrix} 1 \\ -1 \end{pmatrix}.
\]

For establishing a zipper-like compensation mechanism, the vector \( (1, -1)^T \) marked in red color has to be “commuted” from the leftmost to the rightmost site by means of the remaining contributions in the bulk of the chain. In the case of a two-site system, there
is only one such contribution, namely
\[ L^{(2)} \left[ \left( \frac{e}{d} \otimes \frac{e}{d} \right) \right] = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & w_L & -w_R & 0 \\ 0 & -w_L & w_R & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} e^2 \\ ed \\ ed \\ d^2 \end{pmatrix} = \begin{pmatrix} 0 \\ -de(w_R - w_L) \\ +de(w_R - w_L) \\ 0 \end{pmatrix}. \] (2.74)

This can be rewritten in the form
\[ L^{(2)} \left[ \left( \frac{e}{d} \otimes \frac{e}{d} \right) \right] = de(w_R - w_L) \left[ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} e \\ d \end{pmatrix} \right] + \begin{pmatrix} e \\ d \end{pmatrix} \otimes \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right]. \] (2.75)

In fact, adding the boundary contributions (2.72) and (2.73) and the bulk contribution (2.75), it is easy to see that the sum is zero provided that the pre-factor in (2.75) is tuned in such a way that
\[ \frac{de(w_R - w_L)}{d + e} = 1. \] (2.76)

Because of \( \alpha e = \beta d = 1 \) this means that \( \alpha + \beta = w_R - w_L \). Thus, with this particular tuning of the rates, the two-site product state is indeed stationary:
\[ \left( L^{(2)} + A_1 + B_2 \right) \left[ \left( \frac{e}{d} \otimes \frac{e}{d} \otimes \frac{e}{d} \right) \right] = 0. \] (2.77)

The same applies to chains with more than two sites. For example, for a chain with three sites, we obtain the contributions
\[ A_1 \left[ \left( \frac{e}{d} \otimes \frac{e}{d} \right) \right] = + \begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} e \\ d \end{pmatrix} \otimes \begin{pmatrix} e \\ d \end{pmatrix} \] (2.78)
\[ L^{(2)}_{12} \left[ \left( \frac{e}{d} \otimes \frac{e}{d} \otimes \frac{e}{d} \right) \right] = - \begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} e \\ d \end{pmatrix} \otimes \begin{pmatrix} e \\ d \end{pmatrix} + \begin{pmatrix} e \\ d \end{pmatrix} \otimes \begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} e \\ d \end{pmatrix} \] (2.79)
\[ L^{(2)}_{23} \left[ \left( \frac{e}{d} \otimes \frac{e}{d} \otimes \frac{e}{d} \right) \right] = - \begin{pmatrix} e \\ d \end{pmatrix} \otimes \begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} e \\ d \end{pmatrix} + \begin{pmatrix} e \\ d \end{pmatrix} \otimes \begin{pmatrix} e \\ d \end{pmatrix} \otimes \begin{pmatrix} 1 \\ -1 \end{pmatrix} \] (2.80)
\[ B_3 \left[ \left( \frac{e}{d} \otimes \frac{e}{d} \right) \right] = - \begin{pmatrix} e \\ d \end{pmatrix} \otimes \begin{pmatrix} e \\ d \end{pmatrix} \otimes \begin{pmatrix} 1 \\ -1 \end{pmatrix} \] (2.81)

which, adding them all up, cancel mutually on the r.h.s., hence
\[ \left( L^{(2)}_{12} + L^{(2)}_{23} + A_1 + B_3 \right) \left[ \left( \frac{e}{d} \otimes \frac{e}{d} \otimes \frac{e}{d} \right) \right] = 0. \] (2.82)

Thus, we have shown that the ASEP coupled to external reservoirs possesses a stationary product state on a particular line in the phase diagram given by the condition
\[ \alpha + \beta = w_R - w_L. \] (2.83)

Obviously this implies the inequality \( w_R > w_L \) as a necessary condition. This is physically reasonable since a homogeneous stationary state can only be established if the chain transports particles preferentially to the right since otherwise there would be some kind of “traffic jam” emerging at one of the boundaries.
Note that in the case of particle input and output there is no particle conservation any more, hence the system is ergodic and the stationary state is unique.

Since the tensor factors are identical along the chain, the particle density will be constant. Nevertheless, the random walk of the particles is biased to the right, generating a constant flux of particles. As will be discussed later, this is a simple example of a so-called non-equilibrium steady-state (NESS).

### 2.2.4. Matrix product states

So far we have seen that the stationary state of the asymmetric exclusion process is given by a product state, provided that \( \alpha + \beta = w_R - w_L \). This means that we have solved the stationary problem along a particular line in the phase diagram, where the lattice sites are uncorrelated. Is it possible to solve the same problem for arbitrary rates \( \alpha, \beta, w_L, \) and \( w_R \), where the system is expected to exhibit non-trivial correlations?

A very elegant solution to this problem was proposed by Bernhard Derrida and coworkers in the middle of the 90s [7, 8]. They realized that the main restriction with ordinary product states comes from the fact that the product of two tensor factors always gives identical numbers in the second and third component of the resulting vector on the right-hand side. As a way out they proposed to replace the numbers \( e \) and \( d \) by non-commutative objects \( \tilde{E} \) and \( \tilde{D} \). These objects may be thought of as nontrivial operators acting in some fictitious auxiliary space. This space must not be confused with the configuration space of the system, rather it is an additional space on top of that whose only purpose is to establish a certain type of non-commutativity between \( \tilde{E} \) and \( \tilde{D} \). Once a suitable representation of the commutation relations is found, \( \tilde{E} \) and \( \tilde{D} \) can be expressed as matrices acting in the auxiliary space. In order to avoid confusion, we mark all quantities acting in the auxiliary space by a tilde.

Before entering the problem of finding such operators, let us first consider the consequences of this approach. Replacing the numbers \( e \) and \( d \) in the product state

\[
\left| P_{\text{stat}} \right> = \begin{pmatrix} e \\ d \end{pmatrix} \otimes \begin{pmatrix} e \\ d \end{pmatrix} \otimes \ldots \otimes \begin{pmatrix} e \\ d \end{pmatrix}
\] (2.85)

literally by non-commutative matrices \( \tilde{E} \) and \( \tilde{D} \), we would obtain a vector with entries consisting of matrices instead of numbers. For example, on a chain with two sites we would get

\[
\left| P_{\text{stat}} \right> = \begin{pmatrix} \tilde{E} \\ \tilde{D} \end{pmatrix} \otimes \begin{pmatrix} \tilde{E} \\ \tilde{D} \end{pmatrix} = \begin{pmatrix} \tilde{E} \tilde{E} \\ \tilde{E} \tilde{D} \\ \tilde{D} \tilde{E} \\ \tilde{D} \tilde{D} \end{pmatrix}.
\] (2.86)
In this vector the components would be matrices acting in the auxiliary space. However, in the end what we need is a vector of real-valued probabilities. Therefore, another operation is needed to map these matrices back to numbers. This mechanism depends on the boundary conditions under consideration. For example, if the system is open (i.e. coupled to reservoirs) we may take the expectation value of the matrix-valued components between two vectors $|\tilde{\alpha}\rangle$ and $|\tilde{\beta}\rangle$ living in the auxiliary space, i.e.

$$\langle \tilde{\alpha} | \tilde{E} \tilde{D} \otimes \tilde{E} \tilde{D} \otimes \ldots \otimes \tilde{E} \tilde{D} | \tilde{\beta} \rangle,$$

(2.87)

where $\mathcal{N}$ is a normalization factor given by

$$\mathcal{N} = \langle \tilde{\alpha} | \tilde{C} L | \tilde{\beta} \rangle,$$

(2.88)

For example, in the case of $L = 2$ sites the explicit vector reads

$$|P_{stat}\rangle = \frac{1}{\langle \tilde{\alpha} | \tilde{C}^2 | \tilde{\beta} \rangle} \begin{pmatrix}
\langle \tilde{\alpha} | \tilde{E} \tilde{E} | \tilde{\beta} \rangle \\
\langle \tilde{\alpha} | \tilde{E} \tilde{D} | \tilde{\beta} \rangle \\
\langle \tilde{\alpha} | \tilde{D} \tilde{E} | \tilde{\beta} \rangle \\
\langle \tilde{\alpha} | \tilde{D} \tilde{D} | \tilde{\beta} \rangle
\end{pmatrix}. $$

(2.89)

Thus, if we had found a valid set of matrices, this would allow us to compute any physical quantity, unfolding the full power of the approach. For example, the probability of finding the configuration $\emptyset A \emptyset A$ in a four-site system would be given by

$$P_{stat}(\emptyset A \emptyset A) = \frac{\langle \tilde{\alpha} | \tilde{E} \tilde{D} \tilde{E} \tilde{D} | \tilde{\beta} \rangle}{\langle \tilde{\alpha} | \tilde{C}^4 | \tilde{\beta} \rangle}$$

(2.90)

That is, to get the probability of a given configuration, we simply we have to replace a vacancy by the operator $\tilde{E}$ and a particle by the operator $\tilde{D}$ and then compute the product of these matrices. Note that the string $\tilde{E} \tilde{D} \tilde{E} \tilde{D}$ is an ordinary operator product (matrix product) in auxiliary space and should not be confused with the tensor product in configuration space.

**Remark:** Another important case is that of periodic boundary conditions. Here the most natural choice for the reduction of matrices to numbers would be a trace operation

$$|P_{stat}\rangle = \frac{1}{\mathcal{N}} \text{Tr} \begin{pmatrix}
\tilde{E} \\
\tilde{E} \\
\ldots \\
\tilde{E}
\end{pmatrix} \otimes \begin{pmatrix}
\tilde{D} \\
\tilde{D} \\
\ldots \\
\tilde{D}
\end{pmatrix},$$

(2.91)

with the normalization

$$\mathcal{N} = \text{Tr} [\tilde{C}^4],$$

(2.92)

where the trace is carried out in auxiliary space. Likewise, the probability of the configuration $\emptyset A \emptyset A$ in the case of periodic boundary conditions would be given by

$$P_{stat}(\emptyset A \emptyset A) = \frac{\text{Tr} [\tilde{E} \tilde{D} \tilde{E} \tilde{D}]}{\text{Tr} [\tilde{C}^4]}.$$
2.2 The exclusion process on a one-dimensional lattice

As an example, let us find out under which conditions the asymmetric exclusion process (ASEP) with open boundaries admits a matrix representation obeying the same zipper-like compensation mechanism as outlined above for ordinary product states. In analogy to Eqs. 2.75–2.76 let us postulate the bulk relations

\[ \mathcal{L}^{(2)} \left[ \begin{pmatrix} \tilde{E} \\ \tilde{D} \end{pmatrix} \otimes \begin{pmatrix} \tilde{E} \\ \tilde{D} \end{pmatrix} \right] = - \begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} \tilde{E} \\ \tilde{D} \end{pmatrix} + \begin{pmatrix} \tilde{E} \\ \tilde{D} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (2.94) \]

or, explicitly:

\[ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & w_L & -w_R & 0 \\ 0 & -w_L & w_R & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{E} \\ \tilde{D} \\ \tilde{E} \\ \tilde{D} \end{pmatrix} = \begin{pmatrix} 0 \\ w_L \tilde{E} \tilde{D} - w_R \tilde{D} \tilde{E} \\ w_R \tilde{D} \tilde{E} - w_L \tilde{E} \tilde{D} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -D - \tilde{E} \\ D + \tilde{E} \\ 0 \end{pmatrix}. \quad (2.95) \]

Again, the second and the third line are linearly dependent, giving a single algebraic relation for the operators \( \tilde{E} \) and \( \tilde{D} \):

\[ w_R \tilde{D} \tilde{E} - w_L \tilde{E} \tilde{D} = \tilde{D} + \tilde{E}. \quad (2.96) \]

This is the so-called matrix algebra induced by the Liouvillian. If the two operators were just numbers, this relation would be equivalent to that of Eq. 2.76. In other words, the product state solution discussed before is nothing but the special case of a one-dimensional representation of the matrix algebra.

Let us now turn to the boundary equations. Following the same strategy of literally replacing numbers by matrices in Eqs. 2.69 and 2.70 we would postulate the relations

\[ A \begin{pmatrix} \tilde{E} \\ \tilde{D} \end{pmatrix} = + \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{and} \quad B \begin{pmatrix} \tilde{E} \\ \tilde{D} \end{pmatrix} = - \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad (2.97) \]

giving

\[ \alpha \tilde{E} = \beta \tilde{D} = 1. \quad (2.98) \]

But obviously, this result is too restrictive because it would force us to use the one-dimensional representation, reproducing again the case of a product state. However, at this point we can exploit the fact that the boundary matrices at the leftmost and the rightmost site are always contracted with the boundary vectors \( \langle \tilde{\alpha} | \) and \( | \tilde{\beta} \rangle \). That is, instead of Eqs. 2.97 it rather suffices to have

\[ \langle \tilde{\alpha} | A \begin{pmatrix} \tilde{E} \\ \tilde{D} \end{pmatrix} = + \langle \tilde{\alpha} | \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{and} \quad B \begin{pmatrix} \tilde{E} \\ \tilde{D} \end{pmatrix} | \tilde{\beta} \rangle = - \begin{pmatrix} 1 \\ -1 \end{pmatrix} | \tilde{\beta} \rangle, \quad (2.99) \]

turning the scalar equation 2.98 into two eigenvalue problems

\[ \langle \tilde{\alpha} | \tilde{E} = \frac{1}{\alpha} \langle \tilde{\alpha} | \]
\[ \tilde{D} | \tilde{\beta} \rangle = \frac{1}{\beta} | \tilde{\beta} \rangle. \quad (2.101) \]

These eigenvalue equations are in fact more general and allow for a nontrivial matrix.
Finding matrix representations:

With the ansatz described above, the problem of calculating the stationary state of the ASEP with open boundaries has now been shifted to the problem of finding a representation of the matrix algebra, i.e., identifying two matrices $\hat{E}$ and $\hat{D}$ as well as two boundary vectors $\langle \hat{\alpha} \rangle$ and $|\hat{\beta}\rangle$ obeying the relations

$$w_R \hat{D} \hat{E} - w_L \hat{E} \hat{D} = \hat{D} + \hat{E}, \quad \langle \hat{\alpha}| \hat{E} = \frac{1}{\alpha} \langle \hat{\alpha}|, \quad \hat{D}|\hat{\beta}\rangle = \frac{1}{\beta}|\hat{\beta}\rangle. \quad (2.102)$$

Finding such matrix representation is not easy at all, and, if successful, can be considered as a little breakthrough. A general guide for finding matrix representations was written by Blythe and Evans [?], where one can find several representations of the quadratic algebra given above. It turns out that in this case, any matrix representation has to be infinite-dimensional. One of them is:

$$D = \frac{1}{w_R - w_L} \begin{pmatrix} 1 + b & \sqrt{c_0} & 0 & 0 & \cdots \\ 0 & 1 + bq & \sqrt{c_1} & 0 & \\ 0 & 0 & 1 + bq^2 & \sqrt{c_2} & \\ \vdots & \vdots & \vdots & \ddots & \end{pmatrix}, \quad \hat{E} = \frac{1}{w_R - w_L} \begin{pmatrix} 1 + a & 0 & 0 & 0 & \cdots \\ \sqrt{c_0} & 1 + aq & 0 & 0 & \\ 0 & \sqrt{c_1} & 1 + aq^2 & 0 & \\ 0 & 0 & \sqrt{c_2} & 1 + aq^3 & \\ \vdots & \vdots & \vdots & \ddots & \end{pmatrix} \quad (2.103)$$

$$\langle \hat{\alpha}| = (1 \ 0 \ 0 \ \cdots), \quad |\hat{\beta}\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix},$$

where

$$q = \frac{w_L}{w_R}, \quad a = \frac{w_R - w_L}{\alpha} - 1, \quad b = \frac{w_R - w_L}{\beta} - 1, \quad c_n = (1 - q^{n+1})(1 - abq^n). \quad (2.104)$$

This representation is not unique since it can be mapped to other ones by similarity transformation. Let us again remind the reader that these matrices live in an infinite-dimensional auxiliary space, which must not be confused with the (finite-dimensional) configuration space.

Interestingly, there are exceptions: As one sees from the matrices, for certain choices of the parameters, namely for

$$1 - abq^n = 0, \quad (2.105)$$

the representation becomes finite-dimensional since $c_n = 0$ and thus the upper left
corner of the matrices becomes disconnected from the rest (we will study the special case of a two-dimensional representation in the tutorial). These special representations exist along certain submanifolds in the parameters space (see Fig. 2.8). But apart from these special cases one can show that generically an infinite-dimensional representation is needed.

Is this really useful? Is it really an advantage to compute the probability of a configuration with say four sites in terms of the product of four infinite-dimensional matrices? It turns out that for finite chains it is in fact possible to truncate the matrices. For example, for a system with only four sites, it suffices to consider only the first four rows and columns of the matrix. This is because the matrices are nonzero only along their diagonal and one of its neighboring diagonals while the boundary factors have only a single entry in the first component. This means that the matrices \( \tilde{D} \) and \( \tilde{E} \) can be thought of as some kind of ladder operators, moving forward and backward by one component in the auxiliary space.

### Using matrix product states:

Having determined matrix representation it is possible to compute almost any quantity of interest in the stationary state. The most important one is \( \rho_{i}^{\text{stat}} \), the density of particles at site \( i \). For example, in a system with only three sites, the density of particles at the leftmost site would be given by

\[
\rho_{1}^{\text{stat}} = \rho_{\text{stat}}^{(A\emptyset\emptyset)} + \rho_{\text{stat}}^{(A\emptyset A)} + \rho_{\text{stat}}^{(AA\emptyset)} + \rho_{\text{stat}}^{(AAA)} = \rho_{\text{stat}}^{(A\star\star)}
\]

(2.106)

with \( \rho_{\text{stat}} \) given by Eq. (2.87), that is, we have to “integrate out” the degrees of freedom of the other two sites which are ignored. In the matrix product formalism this process of “integrate out” is accomplished by the matrix \( \tilde{C} = \tilde{D} + \tilde{E} \) in that we simply put this matrix in all places which are ignored. In the three-site example, this can be accounted for by writing

\[
\rho_{1}^{\text{stat}} = \frac{\langle \tilde{\alpha} \mid \tilde{D} \tilde{C} \tilde{C} \mid \tilde{\beta} \rangle}{\langle \tilde{\alpha} \mid \tilde{C} \tilde{C} \tilde{C} \mid \tilde{\beta} \rangle}.
\]

(2.107)
More generally, on a chain with \( L \) sites, the probability to find a particle at site \( i \) is given by
\[
\rho_{i}^{\text{stat}} = \frac{\langle \tilde{\alpha} | \tilde{C}^{-i-1} \tilde{D} \tilde{C}^{L-i} | \tilde{\beta} \rangle}{\langle \tilde{\alpha} | \tilde{C}^{L} | \tilde{\beta} \rangle}.
\] (2.108)

The density profile can now be computed by simply inserting the matrices Eq. (2.103) and evaluating the corresponding matrix products. Similarly, it is in principle possible to compute any type of \( n \)-point correlation function. For example, two point-function is given by
\[
C_{ij}^{\text{stat}} = \frac{\langle \tilde{\alpha} | \tilde{C}^{-i-1} \tilde{D} \tilde{C}^{j} \tilde{D}^{-i-1} \tilde{C}^{L-j} | \tilde{\beta} \rangle}{\langle \tilde{\alpha} | \tilde{C}^{L} | \tilde{\beta} \rangle}.
\] (2.109)

**How useful is the matrix product technique?**

The matrix product method is a powerful technique which allows certain problems (such as the ASEP with open boundaries) to be solved exactly. As quantum mechanics, which relies on the idea of replacing real-valued phase space variables by non-commutative operators, this method replaces real-valued probabilities by non-commutative matrices. Physical quantities such as density profiles and correlation functions can be expressed conveniently in terms of simple matrix products.

Nevertheless, the range of applicability seems to be limited. In non-equilibrium statistical physics, the range of soft models seems to be restricted to systems describing various kinds of diffusion. So far only very few systems with particle reactions can be described in terms of matrix products.

Recently, the matrix product technique regained enormous importance in the field of quantum information theory, where it is used to solve chains of interacting qubits. This is the reason why we devoted so much attention to this method in this lecture.

### 2.2.5. Other algebraic structures

**Symmetries and conserved quantities**

From quantum mechanics we know that a **conserved quantity** is represented by an operator \( A \) which commutes with the Hamiltonian \( H \). This is plausible since \( H \) is the generator of translations in time. Therefore, the condition \([H, A] = 0\) implies that it does not make a difference whether we first measure \( A \) and then evolve the system in time or the other way round, meaning that the measurement statistics does not depend on time. The conserved quantity \( A \) usually reflects a certain symmetry property of the system. For example, if the angular momentum is conserved, the system is expected to be invariant under rotations.

The same applies to the formalism of continuous-time Markov processes: An operator \( A \) which commutes with the Liouvillian \( \mathcal{L} \) represents a symmetry associated with a conserved quantity. As an example let us again consider the symmetric exclusion process with 2 sites and without particle input and output at the boundaries. Setting
\( w_L = w_R = 1 \), the Liouvillian is given by the matrix

\[
\mathcal{L} = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 1 & -1 & 0 \\
0 & -1 & 1 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}.
\tag{2.110}
\]

Using the Pauli matrices

\[
\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\tag{2.111}
\]

it is easy to show that the Liouvillian (2.110) can be rewritten as

\[
\mathcal{L} = \frac{1}{2} \left( \mathbb{1} \otimes \mathbb{1} - \sigma^x \otimes \sigma^x - \sigma^y \otimes \sigma^y - \sigma^z \otimes \sigma^z \right),
\tag{2.112}
\]

so up to a shift (the constant \( \mathbb{1} \otimes \mathbb{1} \) which commutes with everything) and a prefactor \(-\frac{1}{2}\), the Liouvillian is essentially given by the square of the Pauli vector operator, namely,

\[
\sigma^2 = (\vec{\sigma} \otimes \mathbb{1}) \cdot (\mathbb{1} \otimes \vec{\sigma}) = \sigma^x \otimes \sigma^x + \sigma^y \otimes \sigma^y + \sigma^z \otimes \sigma^z
\]

and – recalling what we have learned about two spin-1/2 particles in quantum mechanics – we expect it to be an invariant under the symmetry group \( SU(2) \).

**Remember:** The symmetry group \( SU(2) \) describes special (that is without reflections) unitary (i.e. probability-conserving) complex transformations in two dimensions and it is known to describe the phenomenon ‘spin’ in quantum mechanics. It consists of all transformations of the form

\[
T(\vec{a}) = \exp\left(-\frac{\vec{a} \cdot \vec{\sigma}}{2}\right), \quad \vec{a} \in \mathbb{R}^3,
\]

where \( \vec{a} \) plays the role of a rotation axis. The \( SU(2) \) is a Lie-group, i.e. it can be Taylor-expanded around the identity admitting infinitesimal transformations. The corresponding generators are the Pauli matrices \( \vec{\sigma} = (\sigma^x, \sigma^y, \sigma^z) = (\sigma^1, \sigma^2, \sigma^3) \) which obey the \( su(2) \)-Lie algebra (= set of commutation relations)

\[
[\sigma^i, \sigma^j] = 2i\epsilon_{ijk}\sigma^k.
\]

The ordinary rotations \( SO(3) \) in three dimensions obey the same algebra (called angular momentum algebra) wherefore \( SO(3) \) and \( SU(2) \) are locally isomorphic, giving spin its interpretation of an intrinsic angular momentum.

When dealing with the \( SU(2) \), it is convenient to introduce raising and lowering operators

\[
\sigma^\pm = \frac{1}{2}(\sigma^x \pm i\sigma^y)
\tag{2.113}
\]
as well as the number operator

\[
n = \sigma^+ \sigma^-.
\tag{2.114}
\]

The three operators \( \sigma^\pm, n \) can be used instead of \( \sigma^x, \sigma^y, \sigma^z \) as an alternative basis of the \( su(2) \) Lie algebra with the well-known commutation relations

\[
[n, \sigma^\pm] = \pm \sigma^\pm.
\tag{2.115}
\]

So far we have summarized how the symmetry group \( SU(2) \) can be realized on a single vector space \( \mathbb{C}^2 \), i.e., on a single site. When dealing with two sites as in the example
above, we have to employ as in many-particle quantum mechanics. That is, we define the operators

\[ N = n \otimes \mathbb{1} + \mathbb{1} \otimes n, \quad S^\pm = \sigma^\pm \otimes \mathbb{1} + \mathbb{1} \otimes \sigma^\pm. \]  

(2.116)

As can be verified easily, these operators obey the same algebra, namely

\[ N = S^+ S^-, \quad [N, S^\pm] = \pm S^\pm, \]  

(2.117)

hence they are just a two-site (reducible) representation of the su(2) algebra. The generalization to \( L > 2 \) sites is straightforward by adding more and more summands in Eq. (2.116). It is also easy to check that this \( L \)-site representation of the SU(2) is a symmetry group of the symmetric exclusion process, simply by checking that all generator commute with the Liouvillian:

\[ [S^\pm, \mathcal{L}] = [N, \mathcal{L}] = 0. \]  

(2.118)

Nevertheless, the representations \((\sigma^\pm, N)\) and \((S^\pm, N)\) are different in so far as they exhibit a different multiplet structure. For example, on a single site we have \((\sigma^+)^2 = 0\), while on two sites \((S^+)^2 \neq 0\).

**Interpretation of the SU(2) in the exclusion process**

The interpretation of the symmetry group in the context of the exclusion process is very simple: The symmetry is just the conserved number of particles. The operator \( N \) just counts the number of particles, while \( S^\pm \) add and remove a particle, respectively. Therefore, the interpretation of the SU(2) in the context of the exclusion process is much simpler than in the context of quantum spins.

How can we ‘detect’ a symmetry in a given Markov process? The answer is:

> In Markov processes, symmetries and conserved quantities imply degeneracies in the spectrum of the Liouvillian.

The multiplet structure allows us to identify the symmetry group.

Let us verify this claim in the example of the 2-site exclusion process discussed above. The eigenvalues of the Liouvillian (2.110) are

\[ (0, 0, 0, 2). \]

This is a triplet and a singlet, a typical hallmark of SU(2) which we also get when adding two spin-1/2 by Clebsch-Gordon coefficients. The triplet represents three different stationary states with 0, 1, and 2 particles, respectively. The operator \( N \) distinguishes the states while the operators \( S^\pm \) allow us to switch between them. In addition, there is only one excited mode in the sector with one particle since the other two sectors (empty of fully occupied) are frozen in the stationary state.

Moving to three sites, we expect a degeneracy pattern of two doublets and one quartet.
2.2 The exclusion process on a one-dimensional lattice

Quantum groups *

In the 90’s so-called quantum groups were extremely fashionable. Roughly speaking, a quantum group and the associated quantum Lie algebra are a kind of continuously deformed group (together with a corresponding deformed Lie algebra), typically in terms of a continuous deformation parameter $q$ defined in such a way that the non-deformed standard case is recovered for $q = 1$.

In the spectrum of the time evolution operator, the effect of changing the deformation parameter amounts to shifting the eigenvalues while the degeneracies are preserved. In other words, the system is deformed without destroying its symmetry.

Again, the exclusion process is probably the simplest example where we can study and understand a quantum group. To this end we simply consider asymmetric hopping rates

$$w_R = q, w_L = q^{-1}.$$  

It is trivial to see that this modification does not destroy the symmetry of particle conservation, hence it preserves the multiplet structure of the levels in the spectrum of the Liouvillian. However, it turns out that the definition of the two-site generators in Eq. (2.116) has to be replaced by tensor products with different weights.

However, the real fun starts only in the quantum case, when the deformation parameter $q$ is complex. Here it can happen (especially if $q$ is a root of unity) that different multiplet merge into a single larger one, hence leading to a new degeneracy pattern and therewith to a higher degree of symmetry. This is a mathematically fascinating subject on its own, which we will not address here in further detail.

Spectrum-generating algebra

So far we were concerned with symmetry groups and their Lie algebras, which determine the degeneracies in the spectrum of the Liouvillian. But what about the actual values of the levels, which determine the inverse time scale of the corresponding relaxation modes? It turns out that under certain circumstances it is also possible to explain the eigenvalues by means of algebraic methods.

To this end let us consider the symmetric exclusion process with three sites by introducing the following notations:

$$\mathcal{L}^{(2)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad e_1 := \mathcal{L}^{(2)} \otimes \mathbb{1}, \quad e_2 := \mathbb{1} \otimes \mathcal{L}^{(2)}$$

(2.119)
so that $\mathcal{L} = e_1 + e_2$. It is easy to check that

$$e_1^2 = 2e_1, \quad e_2^2 = 2e_2, \quad e_1 e_2 e_1 = e_1, \quad e_2 e_1 e_2 = e_2. \tag{2.120}$$

Now the idea is to consider these equations as algebraic relations independent of the underlying matrix representation and to compute the spectrum algebraically in a representation-free framework.

**Temperley-Lieb algebra**

An *algebra* is not a secret science, on the contrary. It is a kind of construction kit which works as follows:

- **Elementary objects:** Starting point is a set of ‘letters’, say $A, B, C$, or $\hat{q}$ and $\hat{p}$ in quantum mechanics, $a, a^\dagger$ in the harmonic oscillator or, as in our case, the interaction operators $e_1, e_2, \ldots, e_{L-1}$. These letters are the elementary objects of the algebra.

- **Monomials:** As in any language, the letters can be used to form words, the so-called *monomials* of the algebra, such as $ABCBA, \hat{q}\hat{p}, a^\dagger a$, or $e_1 e_2 e_7 e_4$. There are of course infinitely many possibilities to form words. For multiple letters like $AAA$, one uses a power notation $A^3$.

- **Polynomials:** The next step is to introduce a linear space over the words, i.e. we can add different words and multiply them by a (real or complex) scalar. For example, we may consider the linear combination $e_1 e_2 - 5e_6 e_7$. Linear combinations of more than one monomial are called *polynomials* and the corresponding space is the *word space* or *polynomial space*.

- **Algebraic relations:** The core of an algebra is a set of relations which restrict the word space. They often come in the form of commutation relations. The algebraic relations establish linear dependencies between different words.

- **Identity:** Algebras may or may not contain the identity $1$ as an elementary object. The identity is a special ‘letter’ which is restricted by the algebraic relations $1e = e1 = e$ for all other elements $e$ of the algebra.

- **Monomial basis:** A maximal set of monomials (words) which spans the whole word space by linear combinations (constrained by the algebraic relations) is called a *monomial basis*. Each polynomial can be represented as a vector of numbers in terms of these basis elements.

- **Closed algebra:** An algebra is said to close if the monomial basis is finite.

- **Lie groups and Lie algebras:** A *Lie group* is a group $G$ which can be Taylor-expanded around the identity. A well-known example are the rotations in $\mathbb{R}^n$, forming the special orthogonal group $SO(n)$. The group elements $g \in G$ can be expressed as $g = \exp(\lambda h)$, where $h$ is the generator of $g$ and $\lambda$ is a scalar parameter. The generators form an algebra, called the corresponding *Lie algebra*, where the group structure induces certain commutation relations on the generators.
2.2 The exclusion process on a one-dimensional lattice

Example: Group of phase shifts $U(1)$
In quantum mechanics and elsewhere, phase shifts are operations of the form $|\psi\rangle \rightarrow e^{i\theta}|\psi\rangle$. The corresponding Lie algebra contains only the identity and another a single elementary element, say $a$, as well as a single algebraic relation, namely, $a^2 = -\mathbb{1}$. This restricts the word space to linear combinations of the monomial basis $\{\mathbb{1}, a\}$. The group elements of the $U(1)$ are of the form

$$
\gamma = \exp(\lambda a) = \sum_{k=0}^{\infty} a^k \lambda^k = \mathbb{1} \sum_{j=0}^{\infty} (-^j/\lambda^{j+1}) + a \sum_{j=0}^{\infty} (-^j/\lambda^{j+1}) = \mathbb{1} \cos \lambda + e \sin \lambda.
$$

Note that the above equation is abstract and does not use an explicit representation. If we want to find a representation of the Lie algebra and the corresponding Lie group, we have to find objects which precisely obey the correct commutation relations. In the present case possible representations include:

$$
a = i \quad \text{(the imaginary unit), and} \quad a = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.
$$

Returning to the Temperley-Lieb algebra

$$
e_1^2 = 2e_1, \quad e_2^2 = 2e_2, \quad e_1e_2e_1 = e_1, \quad e_2e_1e_2 = e_2
$$

we first note that the elementary letters are $e_1$ and $e_2$ and that the words may never contain the same letter in a row. Hence the words can only be made up of alternating letters, e.g. $e_1e_2e_1e_2\ldots$, but those are in turn constrained by the cubic algebraic relations. Thus the only remaining words, which serve as a monomial basis, are $e_1, e_2, e_1e_2, e_2e_1$. Thus the Temperley-Lieb algebra closes and its word space on two sites is four-dimensional.

All elements $E$ of the algebra can thus be expressed as linear combinations

$$
E = a e_1 + b e_2 + \gamma e_1 e_2 + \delta e_2 e_1 \quad \leftrightarrow \quad (a, \beta, \gamma, \delta)
$$

and may therefore be understood as a vector of coefficients $\vec{E} = (a, \beta, \gamma, \delta)$. Two elements of the algebra $E$ and $E'$ can be combined, giving the result

$$
EE' = (a a' e_1 + b b' e_2 + \gamma a' e_1 e_2 + \delta a' e_2 e_1)(a' e_1 + b' e_2 + \gamma' e_1 e_2 + \delta' e_2 e_1)
$$

$$
= aa' e_1 + ab' e_1 e_2 + a\gamma' e_1 e_1 e_2 + a\delta' e_1 e_2 e_1
+ \beta a' e_2 e_1 + \beta b' e_2 e_2 + \beta\gamma' e_2 e_1 e_2 + \beta\delta' e_2 e_2 e_1
+ \gamma a' e_1 e_2 e_1 + \gamma b' e_1 e_2 e_2 + \gamma\gamma' e_1 e_2 e_2 e_1 + \gamma\delta' e_1 e_2 e_2 e_1
+ \delta a' e_2 e_1 e_1 + \delta b' e_2 e_1 e_2 + \delta\gamma' e_2 e_1 e_2 e_1 + \delta\delta' e_2 e_1 e_2 e_1
= 2aa' e_1 + ab' e_1 e_2 + 2a\gamma' e_1 e_2 + a\delta' e_1
+ \beta a' e_2 e_1 + 2\beta b' e_2 e_2 + 2\beta\gamma' e_2 e_1 e_2 + 2\beta\delta' e_2 e_2 e_1
+ \gamma a' e_1 + 2\gamma b' e_1 e_2 + 2\gamma\gamma' e_1 e_2 e_1 + 2\gamma\delta' e_1 e_1 e_2
+ \delta a' e_1 e_1 + \delta b' e_1 e_2 + 2\delta\gamma' e_1 e_2 + \delta\delta' e_1 e_2
= (2aa' + a\delta' + \gamma a' + 2\gamma\delta') e_1 + (2\beta b' + \beta\gamma' + \delta b' + 2\delta\gamma') e_2
+ (a b' + 2a\gamma' + 2\gamma b' + \gamma\gamma') e_1 e_2 + (b a' + 2b\delta' + 2\delta a' + \delta\delta') e_2 e_1
$$
meaning that the corresponding ‘vectors’ are combined by the rule

\[
\vec{E} \circ \vec{E}' = \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} \circ \begin{pmatrix} \alpha' \\ \beta' \\ \gamma' \\ \delta' \end{pmatrix} = \begin{pmatrix} 2\alpha \alpha' + \alpha \delta' + \gamma \alpha' + 2\gamma \delta' \\ 2\beta \beta' + \beta \gamma' + \delta \beta' + 2\delta \gamma' \\ \alpha \beta' + 2\alpha \gamma' + 2\gamma \beta' + \gamma \gamma' \\ \beta \alpha' + 2\beta \delta' + 2\delta \alpha' + \delta \delta' \end{pmatrix}
\]

(2.123)

Obviously, this rule captures the entire structure of the algebra.

**Eigenvalue problem on the abstract level of the algebra**

Let us now return to the exclusion process without particle input and output on \( L = 3 \) lattice sites. As for the spectrum-generating algebra, it is important to note that the Liouvillian itself is an element of the Temperley-Lieb algebra, namely, \( \mathcal{L} = e_1 + e_2 \). This allows us to set up an ‘eigenword problem’ of the form

\[
\mathcal{L} \vec{E} = \lambda \vec{E} \quad \Rightarrow \quad \vec{L} \circ \vec{E} = \lambda \vec{E},
\]

(2.124)

where \( \vector{E} \) is the ‘eigenword’ and \( \lambda \) the corresponding eigenvalue. In practice we can simply solve the equation \( \vec{L} \circ \vec{E} = \lambda \vec{E} \), which is an ordinary linear eigenvalue problem in the unkown coefficients \( \alpha, \beta, \gamma, \delta \) and the eigenvalue \( \lambda \).

We leave it as an exercise to show that on the one hand the eigenvalues \( \lambda \) exactly reproduce the eigenvalues of the Liouvillian, although the degeneracies are not necessarily the same. The ‘eigenwords’ \( W \), on the other hand, when inserting the matrices for \( e_1 \) and \( e_2 \), turn out to be projectors onto the subspace corresponding to that eigenvalue. Note that the algebraic eigenvalue problem has the size \( 4 \times 4 \) which is smaller than the original eigenvalue problem \( \mathcal{L}|P\rangle = \lambda|P\rangle \) of size \( 8 \times 8 \).

**Difficult? Not at all! Algebra is great fun!**
3. Entropy and Information

3.1. Entropy as an information measure

In statistical physics and the theory of complex systems, *Entropy* is one of the most fundamental concepts. But entropy differs from other physical quantities in many respects. On the one hand, there is a large variety of possible definitions of entropy. On the other hand, unlike other physical quantities such as energy and momentum that can be measured and have a clear intuitive meaning, entropy measures the degree of “uncertainty” is conceptually much more difficult to understand. In 1948, when Claude Shannon found a lower bound for the required bandwidth of telephone lines for which he needed an appropriate name, it is said that J. v. Neumann gave him the advice to

“...call it entropy. [...] Nobody knows what entropy really is, so in a debate you will always have the advantage.” (see e.g. [?])

Shannons discovery was actually a rediscovery, but it changed the point of view fundamentally. Traditionally, entropy was introduced in the context of thermodynamics. Even today, many scientist argue that there are several kinds of entropy around, such as for example information entropy and thermodynamic entropy. Personally I never understood this debate. For me there exists (up to some freedom regarding the contextuality) only one kind of entropy, which is just the same as information. This is the underlying assumption in all what follows: *Entropy* and *Information* are just two words for the same thing.

**Verbal definition of information**

Information is a relative concept which requires a subject and an object and deals with a situation where the subject wants to know something about the object. Roughly speaking, the information of the object is what the subject does *not yet* know about the object. Or, in other words, information is something like the amount of data that the subject would need to gain maximal knowledge about the object. Obviously, this amount of data depends on the already existing knowledge of the subject about the object.
The object could be anything, for example a physical system, a digital storage, a news channel, a black hole, or the text of the bible. In addition, we have to specify the level of the description we are interested in. For example, in the case of a SD card, it makes a difference whether we are interested in the digital information stored on the chip or in the actual position of all molecules of the SD card.

In what follows we will mostly use the example of a classical Markov process introduced in the previous chapter, choosing the microscopic configurations as the level of description. For simplicity let us assume that the set of possible configurations, the configuration space $\Omega$, is finite and let us denote by $N = |\Omega|$ the number of its elements. A light switch, for example, possesses a configuration space $\Omega = \{\text{on, off}\}$ with $N = 2$ states while a die has the configuration space $\Omega = \{1, 2, 3, 4, 5, 6\}$ with $N = 6$ states. Within this framework we start out by defining verbally:

**The information or entropy $S$ of a system is the minimal number of bits needed to specify its configuration.**

In other words, entropy or information can be thought of as the minimal length of a file on a computer that is needed to describe a system in detail. It is important that this length is minimal, i.e., one has to compress the file by removing all redundancies in the description of the configuration. Alternatively, one may define entropy as the number of binary yes-no questions which is necessary to fully identify the microscopic configuration of the system.

As the entropy of a system is the minimal length of a describing file, any composite system can be characterized by first describing its parts and then concatenating the files to a single one. If the components of the system were uncorrelated, it would be impossible to compress the concatenated file even further. It is therefore obvious that entropy is an extensive, i.e., additive quantity.

It should be emphasized that the term ‘information’ used in the present context must not be confused with the notion of information in everyday life. According to the definition given above, a meaningless sequence of random numbers contains a lot of information since a large amount of data is needed to specify all random numbers in detail. Nevertheless these random numbers are meaningless, i.e., they have no ‘information’ in the usual sense.

**The ‘bit’ as elementary unit of information**

If the configuration space of a system contains only one element, meaning that the system is always in a unique configuration, it is already fully characterized and therefore it has the entropy zero. Contrarily, a binary system such as a light switch can be in two possible states, hence it is characterized by a single binary digit, called bit. A bit is the smallest possible portion of information and plays the role as a fundamental information unit, from which other commonly used units of information are derived (see Table 3.1).

It should be pointed out that ‘bit’ is not a physical unit like ‘meter’, which needs to be gauged by an international prototype in Paris. Since the unit ‘bit’ is not scalable,
3.1 Entropy as an information measure

<table>
<thead>
<tr>
<th></th>
<th>Byte (B)</th>
<th>Kilobyte (kB)</th>
<th>Megabyte (MB)</th>
<th>Gigabyte (GB)</th>
<th>Terabyte (TB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>= 8 bit</td>
<td>= 8 \cdot 10^3 \text{ bit}</td>
<td>= 8 \cdot 10^6 \text{ bit}</td>
<td>= 8 \cdot 10^9 \text{ bit}</td>
<td>= 8 \cdot 10^{12} \text{ bit}</td>
</tr>
</tbody>
</table>

Table 3.1.: Commonly used units of information.

it is rather defined in itself. In fact, aliens on a different planet will use the same unit of information as we do. For this reason the universal unit ‘bit’ is often suppressed, treating entropy as a dimensionless quantity.

**Non-integer number of bits:**

Since \( n \) bits can form \( 2^n \) different combinations, it is immediately clear that a system with \( |\Omega| = 2^n \) configurations can be characterized by \( n \) bits. But what happens if the number of state is not a power of 2? For example, two bits with \( 2^2 = 4 \) configurations are not sufficient to encode the six faces of a die, but with three bits two of the \( 2^3 = 8 \) possibilities would be wasted. This suggests that the actual information content of a die is somewhere between 2 and 3 bits.

In order to define such a non-integer information in cases where \( |\Omega| \) is not a power of 2, let us now consider \( M \) (uncorrelated) copies of the system. Since each of the copies can be in \( |\Omega| \) possible configurations, the total system of \( M \) copies can be in \( |\Omega|^M \) possible configurations. Let \( k \) be the minimal number of bits used to encode one of them, i.e.

\[
|\Omega|^{M-1} < 2^k \leq |\Omega|^M.
\]  

(3.1)

To determine the average information content per copy \( k/M \), we first take the logarithm to base 2 and then divide by \( M \). This turns the inequality given above into

\[
\frac{M - 1}{M} \log_2 |\Omega| < \frac{k}{M} \leq \log_2 |\Omega|.
\]  

(3.2)

When taking \( M \to \infty \), this inequality becomes sharper and sharper, and it is easy to see that the information content per copy tends to

\[
S = \lim_{M \to \infty} \frac{k}{M} = \log_2 |\Omega|.
\]  

(3.3)

Following this argument it is well motivated that the information or entropy of the system is given by

\[
S = \log_2 |\Omega|.
\]  

(3.4)

With this formula we can easily confirm that the entropy is additive (extensive) on uncorrelated subsystems: Since the total number of configurations of a system composed of independent subsystems is obtained by multiplying the number of configurations of the respective subsystems, the total entropy is just the sum of their entropy of the parts.
Table 3.2: The six most and least frequently used characters of the English alphabet, their probabilities and the corresponding Morse code (see text).

<table>
<thead>
<tr>
<th>letter</th>
<th>E</th>
<th>T</th>
<th>A</th>
<th>...</th>
<th>X</th>
<th>Q</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>frequency</td>
<td>12.70%</td>
<td>9.06%</td>
<td>8.17%</td>
<td>...</td>
<td>0.15%</td>
<td>0.09%</td>
<td>0.07%</td>
</tr>
<tr>
<td>morse code</td>
<td>.</td>
<td>-</td>
<td>.</td>
<td>-</td>
<td>-..-</td>
<td>-.-</td>
<td>-..</td>
</tr>
</tbody>
</table>

Conventions and scientific cultures

The definition of information-theoretic quantities depends on the community. Basically, there are three major conventions, namely

- Information theory: $S = \log_2 |\Omega|$
- Mathematics: $S = \ln |\Omega| \quad \leftarrow$ preferred choice
- Physics: $S = k_B \ln |\Omega|$

Information scientists use $\log_2$ because they prefer to read information in units of bits. Mathematicians prefer the natural logarithm because it is, as the name tells us, the most natural one and can be inverted by the exponential function. Physicists have the worst definition: It comes with an additional factor $k_B$, which gives entropy the unit of energy per Kelvin, thus masking its real meaning.

The Boltzmann constant $k_B$ is a historical accident due to the fact that entropy was discovered first in the thermodynamic context. We decided to give temperature its own unit, namely, Kelvin. However, as we will see below, this unit is not really needed, and the price we have to pay for a superfluous unit is a superfluous conversion factor, called $k_B$.

From now on we use the convention $S = \ln |\Omega|$, setting $k_B = 1$.

Information with previous knowledge in form of a probability distribution

In defining entropy we assumed that the observer, who is informed about the actual configuration of the system, does not have any previous knowledge about the system. However, if the observer possesses some partial information in advance, it is intuitively clear that less information is needed to characterize a specific configuration of the system. For example, if the observer already knows that a die is manipulated in such a way that the result is always even, then the accessible configuration space is reduced from six to three possibilities, decreasing the necessary information by one bit.

In many cases an observer has a previous knowledge in form of a probability distribution. For example, in the English alphabet the letters ‘E’ and ‘T’ are much more frequent than the letters ‘X’ and ‘Q’. Samuel Morse was one of the first to recognize that it is then more efficient to represent frequent characters by short and seldom ones by longer codes (see Table 3.2). The Morse alphabet is in fact a very early example of what is known today as entropy-optimized coding.
Individual entropy of a specific configuration:

To understand the reduction of information by previous knowledge in the form of a probability distribution, let us consider an arbitrary space $\Omega$ of configurations $c \in \Omega$ occurring with the probabilities $p_c \in [0, 1]$. This probability distribution has to be normalized, i.e.

$$
\sum_{c \in \Omega} p_c = 1 .
$$

For simplicity, let us assume that these probabilities are rational numbers which we will express in terms of their common denominator $m \in \mathbb{N}$ by $p_c = m_c / m$. With these numbers let us now construct a fictitious set with $m$ elements, where the configuration $c$ occurs $m_c$ times. In this fictitious set the relative frequency of the configuration is exactly equal to the given probabilities.

Let us then choose one element from this fictitious set. To specify which of the elements was selected, we would need an information of $\log_2 m$ bit. However, since the set may contain several copies of the same configuration which cannot be distinguished, the information which of the copies was selected is not of interest and has to be subtracted. Therefore, the configuration $c$ has the information content

$$
S_c = \log_2 m - \log_2 m_c = \log_2 \left( \frac{m}{m_c} \right)
$$

in units of bits, or, using the mathematical convention:

$$
S_c = \ln \left( \frac{m}{m_c} \right) .
$$

Because of $m_c / m = p_c$, we arrive at the main result that the configurational or individual entropy $S_c$ of configuration $c$ with respect to a previously known probability distribution is given by

$$
S_c = - \ln p_c
$$

For example, in the English alphabet the most frequent character E, which occurs with the probability $p_E = 0.127$, carries an information of approximately three bit, while the rarely used letter Z carries an information of roughly 10 bits. The less likely a configuration is, the greater its information.

**Example:** Entropy of an information channel:

Let us consider an information channel transmitting three different characters $A, B, C$. Furthermore let it be known that these characters occur with the probabilities $p_A = 1/2$, $p_B = 1/3$, and $p_C = 1/6$. According to the prescription above, we construct a fictitious set $\{A, A, A, B, B, C\}$ in which the relative frequency of the characters equals these probabilities. In this set the configurational entropies are given (in units of bits) by

$$
S_A = - \log_2 (1/2) = 1 \text{ bit},
$$

$$
S_B = - \log_2 (1/3) \approx 1.585 \text{ bit},
$$

$$
S_C = - \log_2 (1/6) \approx 2.585 \text{ bit}
$$

---

1This restriction is not very severe, since any real number can be approximated by a rational number to arbitrary precision.
Average entropy:

In many cases, one is not interested in each of the individual entropies $S_c$ but only in their arithmetic average

$$S := \langle S \rangle_\Omega = - \sum_{c \in \Omega} p_c \ln p_c .$$ (3.9)

This expectation value is the celebrated Gibbs-Boltzmann or Shannon entropy. In the case of data transmission, it just gives the minimal information capacity of the transmission channel that is required to transmit the signal. Note that configurations with a vanishing probability $p_c = 0$, for which the logarithm in the above expression diverges, do not occur and can be excluded. Mathematically this can be accounted for by using the convention

$$0 \ln 0 = 0 .$$ (3.10)

Example: In the previous example of an information channel transmitting three letters A,B,C with different probabilities, the Shannon entropy would be given by

$$S = \sum_{c=A,B,C} p_c S_c = 1.459 \text{ bit} .$$

Without previous knowledge the entropy of the system (information channel) would have been $S = \log_2 3 \approx 1.585$ which is larger. This demonstrates that previous knowledge in form of a probability distribution reduces the amount of information.

Numerical estimation of entropy by sampling on a computer

Everyone knows that the average number of dots on a die can be estimated by throwing the die many times and computing the arithmetic mean of the results. The reason is that the probabilities $p_c$ can be estimated by the relative frequency $n_c / n$ of the results in $n$ repeated experiments, where $n_c$ of them give the result $c$. According to the central limit theorem we know that the error of such an estimate goes down with $1/\sqrt{n}$.

The entropy of a system can simply be estimated in the same way by replacing the probabilities with the relative frequencies, i.e.

$$S \approx - \sum_{c \in \Omega} \frac{n_c}{n} \ln \frac{n_c}{n} .$$ (3.11)

As any average, this estimator converges to the exact value in the limit $n \to \infty$. However, as an important difference we encounter systematic errors on top of the statistical ones. In order to demonstrate these errors, Fig. 3.1 shows the actual value of the estimator for the entropy of a die plotted against the number of experiments (red curve). As expected, the data converges to the theoretical value $S = \ln 6$, which is indicated as a green horizontal line. However, the red data points are not symmetrically scattered around the expected limit, instead they approaches the green line from below, indicating the presence of systematic corrections. To rule out that this effect is just a fluctuation, we averaged over many such sequences, plotting the result as a red dashed line.

These systematic corrections can be traced back to the nonlinearity of the logarithm
3.1 Entropy as an information measure

Figure 3.1.: Numerical estimation of the entropy of a die (red data) depending on the number of trials $n$ together with the average over many repetitions (dashed line). This systematic bias of the data can be compensated by applying Millers correction method (blue data, see text).

in the entropy. Even many experts do not know that various correction methods have been developed which can compensate these systematic deviations to a different extent. The simplest one is the correction term

$$S \approx \frac{1}{2n} |\Omega| - \sum_{c \in \Omega} \frac{n_c}{n} \ln \frac{n_c}{n}$$

(3.12)

introduced by Miller [?]. As is shown in Fig. 3.1 in blue color, this simple $1/n$ correction improves the entropy estimates significantly.

An even better estimator was suggested by Grassberger in 2003, who added another term, reading [?]

$$S \approx \frac{1}{2n} |\Omega| - \frac{1}{n} \sum_{c \in \Omega} \left( -1 \right)^{n_c} \frac{n_c}{n_c + 1} - \sum_{c \in \Omega} \frac{n_c}{n} \ln \frac{n_c}{n} .$$

(3.13)

Deformed measures of information *

Having understood that entropy is basically same as the information content of an object, it may be surprising that several other “entropies” are discussed in the literature. If there are several entropies, what is there meaning with respect to information? In most cases such alternative entropies involve a continuous “deformation parameter”. Since deformed entropies play an important role in current research, we will shortly address the most important examples.

Deformations

In physics, a deformation of a physical law is a modification controlled by a continuously varying parameter in such a way that the original law is retrieved in a particular limit of
the control parameter. A famous example is the formula for the non-relativistic kinetic energy $E_{\text{kin}} = \frac{1}{2}mv^2 = p^2/2m$. In special relativity, this formula is replaced by the deformed variant

$$E = \sqrt{m^2c^4 + p^2c^2}$$  \hspace{1cm} (3.14)

which depends on a deformation parameter, namely, the velocity of light $c$. For large $c$ the square root in the expression given above can be Taylor-expanded as

$$E = mc^2 \sqrt{1 + \frac{p^2}{m^2c^2}} \simeq mc^2 + \frac{p^2}{2m} - \frac{p^4}{4m^2c^2} + O(c^{-4}).$$  \hspace{1cm} (3.15)

As can be seen, one gets an additional constant offset $E_0 = mc^2$. Moreover, the non-relativistic formula $E_{\text{kin}} = p^2/2m$ is retrieved in the limit $c \to \infty$.

**Continuous deformations of the Shannon entropy**

In the following we are interested in a continuous deformation of the Shannon entropy $S = -\sum P_i \ln P_i$. We would like to deform it in a way that the deformed entropy

- is extensive on uncorrelated subsystems,
- vanishes if the actual configuration is known, and
- equals $\ln |\Omega|$ in the case of uniform distribution.

In order to find out whether such a deformed entropy exists, we first note that the Shannon entropy can be written in the form

$$S_X = f\left(\langle g(P_X) \rangle \right) = f\left(\sum_x P(x)g(P(x))\right) \quad \text{with} \quad f(x) = x \quad \text{and} \quad g(x) = -\ln x.$$

Our aim is to find a continuously parameterized family of functions $f, g$ in such a way that it meets the requirements listed above.

Let us first check under which conditions such a deformed entropy is still extensive on uncorrelated subsystems. To this end let $A, B$ be two subsystems with configurations $a \in \Omega_A$ and $b \in \Omega_B$. Let us assume that the two parts are uncorrelated, meaning that the probability distribution $P_c$ of the composite system with the configuration $c = (a, b)$ factorizes:

$$P_c = P_a P_b.$$  \hspace{1cm} (3.16)

The deformed entropy is extensive if $S_{AB} = S_A + S_B$, hence

$$S_{AB} = f\left(\sum_c P_c g(P_c)\right) = f\left(\sum_{a,b} P_a P_b g(P_a P_b)\right)$$

$$\approx S_A + S_B = f\left(\sum_a P_a g(P_a)\right) + f\left(\sum_b P_b g(P_b)\right).$$  \hspace{1cm} (3.17)

Since the product of probabilities has to be converted into sum, it is clear that this will not work without a logarithm. Turns out that there are basically two possibilities.
3.1 Entropy as an information measure

The logarithm could be either placed inside or outside the summation. If it is inside, meaning that \( g(x) \propto \ln x \), we recover the Shannon entropy. However, if the logarithm is in front of the sum \( (f(x) \propto \ln x) \), we can define a different kind of entropy, provided that

\[
\sum_{ab} P_a P_b g(P_a P_b) = \left[ \sum_a P_a g(P_a) \right] \cdot \left[ \sum_b P_b g(P_b) \right]
\]

(3.18)

Obviously, this requires the function \( g \) to be homogeneous, that is

\[
g(P_a P_b) = g(P_a) g(P_b) \Rightarrow g(P_a P_b) = (P_a P_b)^\beta,
\]

(3.19)

where \( \beta \) plays the role of a deformation parameter. In addition, extensivity requires the function \( g(z) \) to be a linear function of \( \ln z \). This linear function can be fixed by requiring that the deformed entropy vanishes in situations where we know the actual configuration, i.e., for distribution \( P_c = \delta_{c c_0} \), and that it equals \( \ln |\Omega| \) in the case of a uniform distribution \( P_c = 1/|\Omega| \). The first condition implies that the deformed entropy is of the form

\[
S_\beta = C_\beta \ln \sum_c P_c^\beta = C_\beta \ln \sum_x P_1^{1+\beta}
\]

(3.20)

with some (possibly \( \beta \)-dependent) pre-factor \( C_\beta \). The second condition means that

\[
\ln |\Omega| = C_\beta \ln \left[ |\Omega| |\Omega|^{-1-\beta} \right]
\]

(3.21)

hence \( C_\beta = -\frac{1}{\beta} \). Altogether we arrive at the conclusion that in addition to the Shannon entropy there exists another family of extensive entropies of the form

\[
S_\beta = -\frac{1}{\beta} \ln \sum_c P_1^{1+\beta}
\]

(3.22)

**Rényi entropy**

In the literature the deformed entropy derived above is known as Rényi entropy, named after the Hungarian mathematician Alfred Rényi (1921-1970). The only difference is that we have to replace \( \beta \to 1 - \alpha \). For a given probability distribution \( \{P_\omega\} \) on a sample space \( \Omega \) the Rényi entropy \( H^R \) is defined by

\[
S^R_\alpha = \frac{1}{1-\alpha} \ln \sum_c P_\omega^\alpha,
\]

(3.23)

where \( \alpha > 0 \) is the deformation parameter. The corresponding continuous version, the differential Rényi entropy reads

\[
S^R_\alpha = \frac{1}{1-\alpha} \ln \int_{-\infty}^{+\infty} \frac{dx}{|p(x)|^\alpha}.
\]

(3.24)

The Rényi entropy includes four important special cases:

\[2\text{Note that it is not allowed to commute the sum with the logarithm which is highly non-linear.}\]
\[ S_R^\alpha = \begin{cases} \ln |\Omega| & \alpha = 0 \\ -\sum_c P_c \ln P_c & \alpha = 1 \\ -\ln \sum_c P_c^2 & \alpha = 2 \\ -\ln \max_c P_c & \alpha = \infty \end{cases} \]

Hartley entropy

Shannon / information entropy

Correlation entropy

Minimal individual entropy

In particular, as can be shown easily by using the rule of l'Hôpital, the Rényi entropy reduces to the ordinary Shannon entropy in the limit \( \alpha \to 1 \):

\[
\lim_{\alpha \to 1} S_R^\alpha = S. \tag{3.25}
\]

The meaning of the Rényi entropy: Moments and cumulants of information

It is very interesting that the Rényi entropy has a clear and significant mathematical interpretation. To see this let us first recall that for a given probability distribution \( \{ P_c \} \) each configuration \( c \) carries an individual entropy (cf. Sect. 3.1 on page 45)

\[
S_c = -\ln P_c. \tag{3.26}
\]

The ordinary Shannon-Gibbs-Boltzmann entropy is in fact nothing but its expectation value (arithmetic mean)

\[
S = \langle S_c \rangle = \sum_c P_c S_c = - \sum_c P_c \ln P_c. \tag{3.27}
\]

Thus, if we know the Shannon entropy of the system, we actually know very little about the entire probability distribution, namely, only a single number which is the first moment of the individual entropies (the average information). However, it is of course also possible to define higher moments of the individual information, i.e.,

\[
m_k := \sum_c P_c S_c^k = \sum_c P_c (-\ln P_c)^k. \tag{3.28}
\]

These moments characterize the distribution of the individual information content, in particular \( m_1 = S \) is just the standard Shannon entropy. As outlined above, this allows us to define a corresponding moment-generating function (see Sect. 3 on page 45)

\[
M(t) = \sum_{k=0}^\infty \frac{M_k}{k!} t^k = \sum_{k=0}^\infty \frac{1}{k!} \sum_c P_c (-\ln P_c)^k t^k
\]

\[
= \sum_{k=0}^\infty \frac{1}{k!} \sum_c P_c ((-\ln P_c) t)^k = \sum_c P_c \sum_{k=0}^\infty \frac{1}{k!} ((-\ln P_c) t)^k \tag{3.29}
\]

\[
= \sum_c P_c \exp((-\ln P_c) t) = \sum_c P_c P_c^{-t} = \langle P_c^{-t} \rangle_c
\]

In fact, we can retrieve the standard entropy by computing the first derivative:

\[
\left. \frac{d}{dt} M(t) \right|_{t=0} = \sum_c \left. \frac{d}{dt} P_c^{1-t} \right|_{t=0} = \sum_c (-P_c^{1-t} \ln P_c) \bigg|_{t=0} = - \sum_c P_c \ln P_c = S. \tag{3.30}
\]
Moments are good, cumulants are better\footnote{Moments are good, cumulants are better. In order to obtain the cumulants of the individual information, we consider the cumulant-generating function}

\[
K(t) = \ln M(t) = \ln \sum_c p_c^{1-t}. \tag{3.31}
\]

Comparing the cumulant-generating function with the original expression for the Rényi entropy given above, we observe that they are closely related by

\[
S^R_\alpha = \frac{K(1-\alpha)}{1-\alpha}. \tag{3.32}
\]

This means that, apart from the renamed argument and an additional prefactor, the Rényi entropy is just the cumulant-generating function of the information content. Thus, knowing the Rényi entropy for all values of the deformation parameter \(\alpha\), we could in principle reconstruct all individual entropies \(S_c\) and therewith the whole probability distribution.

**Remember**: The Rényi entropy is essentially the cumulant-generating function of the individual entropy \(S_c = -\ln P_c\). It is extensive and carries in principle all information about the probability distribution.

**Tsallis entropy**

Another frequently studied deformed entropy measure is the so-called Tsallis entropy that was introduced in 1988 by Constantino Tsallis. For a given probability distribution \(\{P_c\}\) with \(\sum_c P_c = 1\) the Tsallis entropy is defined by

\[
S^T_q = \frac{1}{1-q} \left( 1 - \sum_c P_c^q \right), \tag{3.33}
\]

where \(q \in \mathbb{R}\) is a deformation parameter. As can be shown easily, when taking \(q \to 1\) the Tsallis entropy reduces to the ordinary Shannon-Boltzmann entropy:

\[
\lim_{q \to 1} S^T_q = -\sum_c P_c \ln P_c = S, \tag{3.34}
\]

meaning that the standard entropy is included as a special case \(S^T_1 \equiv S\). Likewise one can define a differential Tsallis entropy for continuous probability densities by

\[
S^R_q = \frac{1}{1-q} \left( 1 - \int [p(x)]^q \, dx \right), \tag{3.35}
\]

where \(p(x) \, dx\) is the probability to find the random variable \(x\) in the infinitesimal interval \([x, x + dx]\).

In the literature the physical relevance of Tsallis entropy has been debated controversially, mainly because there is no immediate information-theoretic interpretation.
However, the concept could be applied successfully to a large variety of physical systems and empirical data, ranging from spin glasses to high-energy experiments.

It is important to note that the deformation destroys the extensivity of the measure. In fact, if $X$ and $Y$ are two uncorrelated random variables with factorizing probability distributions, we find that

$$S^T_q(XY) = S_q(X) + S_q(Y) + (1 - q)S_q(X)S_q(Y).$$ (3.36)

Therefore, Tsallis entropy is not really suitable for uncorrelated systems. However, the hope is that it might work for correlated systems, for example, with particles interacting by long-range forces such as gravity or electromagnetic interactions.

In more recent papers, it was shown that the Tsallis and the Renyi entropy can be mapped onto each other. This means that they carry in principle the same information content.

### 3.2. Entropy in Statistical Physics

**Gibbs postulate and Second Law**

Entropy as a measure of information has no direct physical meaning in itself. It acquires a physical meaning only through the circumstance that sufficiently complex physical systems evolve chaotically. Chaotic behavior means that any kind of fluctuation is amplified by the nonlinear equations of motions, leading effectively to an apparent random behavior. As discussed before, this allows us to use the cartoon introduced in the previous chapter of a system jumping randomly in its classical configuration space $\Omega$ according to specific rates $w_{c\rightarrow c'}$. How this simple picture can be justified within the theory of quantum chaos is an interesting research topic on its own.

**Quantum reversibility:** As already mentioned in the previous chapter, to our present knowledge any physical system is ultimately described by the laws of quantum physics. In a nonrelativistic setting this means that even a very complex system evolves in time according to the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = H\psi.$$ (3.37)

This equation has the remarkable property that it is invariant under time reversal combined with complex conjugation

$$t \rightarrow -t, \quad \psi \rightarrow \psi^*.$$ (3.38)

Likewise the Hamilton equations of motion, which plays the role of a classical pendant the Schrödinger equation,

$$\dot{p} = -\frac{\partial H}{\partial q}, \quad \dot{q} = +\frac{\partial H}{\partial p}$$ (3.39)
3.2 Entropy in Statistical Physics

is invariant under time reversal:

\[ t \rightarrow -t, \quad p \rightarrow -p. \]  

(3.40)

Time reversal means the following: if a movie showing a physical time evolution is played backward it looks physically reasonable, i.e., it could be a valid solution of the evolution equation with appropriate initial conditions.

The underlying assumption of the unitary Schrödinger evolution and the Hamilton equations, which is often omitted in textbooks, is that the system under consideration is completely isolated from the environment. In other words, these equations are only valid in isolated systems.

**Equal a priori postulate:** Transferring this insight to the cartoon of a system jumping randomly in its configuration space we arrive at the conjecture that a isolated system described within this framework should be time-reversal invariant as well. This means that the probability of any stochastic sequence of transitions forward in time has to be exactly equal to the probability of the reversed sequence backward in time. This is the case if and only if their rates in forward and backward direction are identical, leading us directly to the most fundamental postulate of statistical physics:

\[ w_{c \rightarrow c'} = w_{c' \rightarrow c}. \]  

(3.41)

This axiom holds for isolated systems which are perfectly isolated so that they do not interact with their environment. Under this condition the axiom states that spontaneous jumps between a pair of states are equally probable in both directions. Therefore, the motion of the system has no preferred direction, it rather diffuses in its configuration space like a random walker.

For symmetric transition rates the Liouvillian is by construction symmetric. Hence the condition of probability conservation \( \langle 1 | \mathcal{L} = 0 \) (see Eq. (2.14)) immediately implies the equality

\[ \mathcal{L}^T |1\rangle = \mathcal{L} |1\rangle = 0, \]  

(3.42)
i.e., the column vector \( (1, 1, 1, \ldots)^T \) is a right eigenvector to the eigenvalue zero. In addition, if the system is ergodic, this eigenvector has to be proportional to the stationary state of the system, implying

\[ |p_{\text{stat}}\rangle = \frac{1}{|\Omega|} |1\rangle. \]  

(3.43)

Thus, in the long-time limit \( t \rightarrow \infty \), an isolated ergodic system will relax into a stationary state where all configurations are equally probable. This is also known as the equipartition postulate, the equal-a-priori postulate or as the Gibbs postulate:

In an isolated stationary ergodic system all states occur with the same probability \( p_i = 1/|\Omega| \).

© HAYE HINRICHSEN — PHYSICS OF COMPLEX SYSTEMS — CC BY-SA 4.0
As we will see below, the whole theory of equilibrium statistical mechanics and thermodynamics can be derived from this simple postulate.

**Second Law of thermodynamics:** Suppose that we have obtained the information that the system is in the configuration $c_0$ at time $t = 0$, meaning that the initial probability distribution is given by $P_c(t_0) = \delta_{c,c_0}$. If we let the system evolve without further measurement and if the rates are symmetric in both directions, the system will basically perform a random walk in its own configuration space. Thus, as time proceeds, it will be less and less clear where the system actually is, meaning that the information needed to characterize its configuration increases. This is the essential content of the famous Second Law of Thermodynamics:

In an isolated physical system the average entropy cannot decrease.

$$\Delta S \geq 0$$

The Second Law can be proven as follows. Since the rates are assumed to be symmetric ($w_{c\rightarrow c'} = w_{c'\rightarrow c}$, $w_{c'c}$), we simply insert the master equation $\dot{P}_c = \sum_{c'} w_{c'c} (P_{c'} - P_c)$ into the temporal derivative of the Shannon entropy:

$$\frac{d}{dt} S(t) = -\frac{d}{dt} \sum_c P_c(t) \ln P_c(t) = - \sum_c \left( \frac{d}{dt} P_c(t) \right) \ln P_c(t) - \sum_c \frac{P_c(t)}{P_c(t)} \frac{d}{dt} P_c(t)$$

$$= - \sum_c \left[ \ln P_c(t) \sum_{c'} w_{c'c} (P_{c'}(t) - P_c(t)) \right] - \sum_{c,c'} \frac{P_c(t)}{P_c(t)} \frac{d}{dt} P_c(t)$$

(3.44)

Now we use the following trick: In the last expression we can rename the indices $c \leftrightarrow c'$:

$$= - \sum_{c,c'} \ln P_c(t) w_{c'c} (P_{c'}(t) - P_c(t))$$

Since both expressions are identical, we can instead write their sum divided by 2. This leads us to

$$\frac{d}{dt} S(t) = - \frac{1}{2} \sum_{c,c'} \left[ \ln P_c(t) w_{c'c} (P_{c'}(t) - P_c(t)) + \ln P_{c'}(t) w_{c'c} (P_c(t) - P_{c'}(t)) \right]$$

(3.45)

$$= - \frac{1}{2} \sum_{c,c'} \left[ \ln P_c(t) (P_{c'}(t) - P_c(t)) - \ln P_{c'}(t) (P_c(t) - P_{c'}(t)) \right]$$

$$= + \frac{1}{2} \sum_{c,c'} \ln P_c(t) - \ln P_{c'}(t) \left( P_{c'}(t) - P_c(t) \right).$$

We now focus our attention on the two brackets in the sum. Since the logarithm is a monotonously increasing function, it is clear that both brackets, if nonzero, have always
3.2 Entropy in Statistical Physics

the same sign, hence their product is non-negative. In addition, the rates \( w_{c \rightarrow c'} \) are by definition non-negative. Hence we can conclude that \( \frac{d}{dt} S(t) \geq 0 \), completing the proof.

We will come back to the second law when studying fluctuation theorems.

Subsystems

In most physical situations the system under consideration is not isolated, instead it interacts with the environment. In this case the usual approach of statistical physics is to consider the system combined with the environment as a composite system. From an extreme point of view this could encompass the entire Universe. This superordinate total system is then assumed to be isolated, following the same rules as outlined above.

To distinguish the total system from its parts, we will use the suffixes ‘tot’ for the total system while ‘sys’ and ‘env’ refer to the embedded subsystem and its environment, respectively.

The total system, which includes the laboratory system as well as the environment, is characterized by a certain space \( \Omega^{\text{tot}} \) of classical configurations \( c \in \Omega^{\text{tot}} \). The number of these configurations may be enormous and they are usually not accessible in experiments, but in principle there should exist a corresponding probability distribution \( P_c(t) \) evolving by some master equation (cf. Eq. 2.3)

\[
\frac{d}{dt} P_c(t) = \sum_{c' \in \Omega^{\text{tot}}} \left( J_{c \rightarrow c'}(t) - J_{c' \rightarrow c}(t) \right).
\] (3.46)

Here we introduced the so-called probability current

\[
J_{c \rightarrow c'}(t) = P_c(t) w_{c \rightarrow c'}
\] (3.47)

flowing from configuration \( c \) to configuration \( c' \), where \( w_{c \rightarrow c'} \geq 0 \) denotes the corresponding time-independent transition rate.

Let us now consider a subsystem embedded into an environment. Obviously, for every classical configuration \( c \in \Omega^{\text{tot}} \) of the total system we will find the subsystem in a well-defined unique configuration, which, for the sake of clarity, we will denote as \( s \in \Omega^{\text{sys}} \). Conversely, for a given configuration of the subsystem \( s \in \Omega^{\text{sys}} \) the environment (and therewith the total system) can be in many different configurations. This relationship can be expressed in terms of a surjective map \( \pi : \Omega^{\text{tot}} \rightarrow \Omega^{\text{sys}} \) which projects every configuration \( c \) of the total system onto the corresponding configuration \( s \) of the subsystem, as sketched in schematically Fig. 3.2.

The projection \( \pi \) divides the space \( \Omega^{\text{tot}} \) into sectors \( \pi^{-1}(s) \subset \Omega^{\text{tot}} \) which consist of all configurations which are mapped onto the same \( s \). Therefore, the probability to find the subsystem in configuration \( s \in \Omega^{\text{sys}} \) is the sum over all probabilities in the
Entropy and Information

Figure 3.2: A subsystem is defined by a projection $\pi$ which maps each configuration $c \in \Omega^{\text{tot}}$ of the total system (left) onto a particular configuration $s \in \Omega^{\text{sys}}$ of the subsystem (right), dividing the configuration space of the total system into sectors. The figure shows a stochastic path in the total system together with the corresponding stochastic path in the subsystem.

Corresponding sector, i.e.

\[
P_s(t) = \sum_{c(s)} P_c(t),
\]

where the sum runs over all configurations $c \in \Omega^{\text{tot}}$ with $\pi(c) = s$, denoted as $c(s)$. Likewise, the projected probability current $J_{s \rightarrow s'}$ in the subsystem flowing from configuration $s$ to configuration $s'$ is the sum of all corresponding probability currents in the total system:

\[
J_{s \rightarrow s'}(t) = \sum_{c(s')} \sum_{c'} J_{c \rightarrow c'}(t) = \sum_{c(s)} P_c(t) \sum_{c'} w_{c \rightarrow c'}. \tag{3.49}
\]

This allows us to define effective transition rates in the subsystem by

\[
w_{s \rightarrow s'}(t) = \frac{J_{s \rightarrow s'}(t)}{P_{s}(t)} = \frac{\sum_{c(s)} P_c(t) \sum_{c'(s')} w_{c \rightarrow c'}}{\sum_{c(s)} P_c(t)}. \tag{3.50}
\]

In contrast to the transition rates of the total system, which are usually assumed to be constant, the effective transition rates in the subsystem may depend on time. With these time-dependent rates one can show that the subsystem evolves according to the master equation (see exercise)

\[
\frac{d}{dt} P_{s}(t) = \sum_{s' \in \Omega^{\text{sys}}} \left( J_{s \rightarrow s'}(t) - J_{s' \rightarrow s}(t) \right) = P_s(t) w_{s \rightarrow s'}(t). \tag{3.51}
\]

From the subsystems point of view the time dependence of the rates reflects the unknown dynamics in the environment. Moreover, ergodicity plays a subtle role: Even if the dynamics of the total system was ergodic, the dynamics within the sectors $\pi^{-1}(c)$ is generally non-ergodic and may decompose into several ergodic subsectors. As we will see later, this allows the environmental entropy to increase even if the subsystem is stationary.

**Detailed balance:** Let us now consider a situation where the total system (the whole Universe) approaches a stationary state. In this limit, all configurations of the total system are equally probable, i.e., $P_{c}^{\text{tot}} = 1/|\Omega^{\text{tot}}|$. For the laboratory system, however,
the equipartition postulate does not apply. In fact, if we evaluate Eq. (3.48) we obtain
\[P_{s}^{\text{stat}} = \sum_{c(s)} P_{c}^{\text{stat}} = \frac{|\pi^{-1}(s)|}{|\Omega|}.\] (3.52)

This means that the stationary probability of a certain configuration \(s\) of the laboratory system is proportional to the number of the corresponding configurations of the total system \(|\pi^{-1}(s)|\).

Let us now turn to the probability currents. Assuming stationarity Eq. (3.49) turns into
\[J_{s \rightarrow s'}(t) = \frac{1}{|\Omega|} \sum_{c(s)} \sum_{c'(s')} w_{c \rightarrow c'}.\] (3.53)

As the total system is isolated, we know that the transition rates of the total system have to be symmetric, i.e., \(w_{c \rightarrow c'} = w_{c' \rightarrow c} \). Inserting this symmetry into the equation above, we immediately recognize that the probability currents have to be symmetric as well:
\[J_{s \rightarrow s'} = J_{s' \rightarrow s} .\] (3.54)

In other words, in the equilibrium state the probability currents cancel one another. This is of course not very surprising: Since the probability currents in the total systems cancel by definition, we expect the same to hold in any subsystem.

Note that this cancellation takes place between all pairs of configurations and not only effectively between groups of configurations. For this reason such a situation, where all probability currents balance one another microscopically, is referred to as obeying detailed balance.

A stationary state is said to obey detailed balance if all probability currents cancel pairwise.

Obviously, detailed balance implies stationarity since there is no net probability flow in the system. However, the effective rates of the laboratory system are not necessarily symmetric:
\[w_{s \rightarrow s'} \neq w_{s' \rightarrow s} .\] (3.55)

According to the equipartition postulate the whole Universe is expected to evolve into a state where all configurations are equally probable. This so-called thermal death hypothesis has been discussed intensively around the end of the 19th century. Today we know that our expanding Universe is actually far from thermalizing into such a state. However, the immediate surrounding of the laboratory system could be locally thermalized so that the considerations made above still makes sense.

---

\[^{4}\text{As we will see in the following chapter, there are also stationary states which do not obey detailed balance. Such states are called non-equilibrium steady states (NESS).}\]
Thermodynamic equilibrium: Throughout this lecture the terms “stationarity” and “equilibrium” mean something different.

- Stationarity means that the probability distribution \( P_s(t) \) does not depend on time, i.e. \( P_s(t) = P_{s_{\text{stat}}} \).
- Equilibrium means that the system is stationary and additionally that the rates obey the condition of detailed balance.

Therefore, thermodynamic equilibrium is a much stronger condition than stationarity. This difference can also be seen by looking at the right side of the master equation

\[
\frac{d}{dt} P_s(t) = \sum_{s' \in \Omega_{\text{sys}}} \left( J_{s' \rightarrow s}(t) - J_{s \rightarrow s'}(t) \right). \tag{3.56}
\]

Stationarity requires that the sum on the right-hand side of this equation vanishes identically. Equilibrium, on the other hand, requires that each term of the sum vanishes individually. Clearly, the second condition is much more restrictive.

If the system under consideration is isolated, the Gibbs postulate tells us that in the stationary state all configurations are equally probable and that the rates have to be symmetric. Consequently, the entropy of such an isolated system attains the maximal possible value \( S = \ln |\Omega| \).

As we have shown above, the Gibbs postulate allows us to derive that an open system interacting with its environment approaches a stationary state where it obeys detailed balance. In this situation, the configurations are not necessarily equally probable, meaning that the entropy of the subsystem is not necessarily maximal. Thus, although the total system (consisting of laboratory system and environment) approaches a state of maximal entropy, this does not automatically apply to a subsystem. In fact, as will be shown, different thermodynamic potentials may become extremal in this case.

3.3. Thermo statics with conserved quantities and reservoirs

Systems with conserved quantities

Microcanonical ensemble: In an isolated system without conserved quantities, the Second Law tells us that the system would end up in a equally distributed state which is not particularly interesting. However, as in any field of physics, the situation changes completely in the presence of symmetries and the associated conserved laws.

As the most important example let us consider energy conservation. As illustrated in Fig. 3.3, this conservation law divides the configurations space \( \Omega^\text{iso} \) into sectors \( \Omega^\text{iso}(E) \) of the same energy. For an isolated system this means that spontaneous jumps are now restricted to pairs of configurations within the same sector, while jumps between different sectors are forbidden. Within these sectors the Second Law remains valid, i.e. the system will approach a stationary equi-partitioned state within this sec-
3.3 Thermostatics with conserved quantities and reservoirs

Figure 3.3: Illustration of sectors in the state space of a system with energy conservation.

$S(E) = \ln |\Omega^{\text{sys}}(E)|$. (3.57)

In statistical physics, this situation is referred to as the microcanonical ensemble.

**Systems exchanging entropy and energy:** Let us now consider the case of two weakly interacting systems $A$ and $B$ coupled by a thermal bridge, as shown in Fig. 3.4. In such a situation the energy conservation law

$E = E_A + E_B$ (3.58)

still holds for the combined total system but no longer for its parts. In fact, it may happen that subsystem $A$ jumps from one energy sector to another one, provided that subsystem $B$ jumps simultaneously to another sector in opposite direction. Obviously such an interaction leads to an exchange of energy between the systems.

Furthermore let us assume that the coupling is so weak that the thermal bridge itself does not significantly contribute to the entropy of the entire setup so that the total entropy of the system can be written as the sum of its parts, i.e.,

$S = S_A + S_B$. (3.59)

Since the composite system $AB$ can be thought of as being isolated, we can again apply the Second Law, meaning that the system approaches a stationary state in which $S = \log_2|\Omega|$ is maximal and constant. However, in the stationary state the system still fluctuates and the two parts continue to exchange energy. This means that $S_A, S_B$ as well as $E_A, E_B$ fluctuate in such a way that their sums are constant, i.e. a gain on one side implies an equal loss on the other side and vice versa:

$S = S_A + S_B = \text{const}$

$E = E_A + E_B = \text{const}$

$\Rightarrow \quad \Delta S_A = -\Delta S_B$ and $\Delta E_A = -\Delta E_B$. (3.60)

Regarding the entropies of the subsystems $S_A(E_A)$ und $S_B(E_B)$ as continuous functions of the corresponding energies, the differential relation $dE = -dE_i$ implies that

$\frac{\partial S_A}{\partial E_A} = -\frac{\partial S_A}{\partial E_B}$. (3.61)

On the other hand, since the entropy of the total system is maximal and constant we
have $dS_x = -dS_y$, leading to

$$\frac{\partial S_A}{\partial S_B} = - \frac{\partial S_B}{\partial E_B} \Rightarrow \frac{\partial S_A}{\partial E_A} = -\frac{\partial S_B}{\partial E_B}. \tag{3.62}$$

Thus, in the stationary state of the combined system, the derivative

$$\frac{1}{T} = \beta := \frac{\partial S}{\partial E} \tag{3.63}$$

is equal on both sides. Here $T$ is the so-called temperature of the system. Consequently we can interpret temperature as the change of the required information to specify the configuration of a system caused by a change of the energy. In other words, it tells us how much the logarithm of the size of the sector grows if we increase the energy of the system. According to the arguments given above, the two systems continue to exchange energy until their temperatures coincide.

The unit of temperature: In our everyday world we measure temperature in units of Kelvin, Celsius or Fahrenheit. However, from a fundamental point of view, the introduction of an independent unit of temperature is absolutely superfluous. In fact, since temperature is energy divided by information, the correct SI-unit would be $[T] = \text{Joule/bit}$. The price we pay for introducing an unnecessary unit is the necessity of an unnecessary conversion factor, the so-called Boltzmann constant $k_B = 1.380649 \cdot 10^{-23} \text{J/K}$. The existence of this constant, like the existence of vacuum permeabilities $\epsilon_0$, $\mu_0$ in electrodynamics, can be considered as a historical accident. However, of course it would be somewhat strange to specify the temperature in a weather forecast in units of electron volts per kilobyte.

Intuitive interpretation: To understand the meaning of $\beta = \frac{1}{T}$, let us consider two economies like the crude oil exchange markets in Rotterdam and New York. The traders working in these places are not interested in crude oil in itself, may be they have never seen a single barrel, rather they are interested in making money. They do so by randomly trading oil for money between the two places, causing a fluctuating physical current of oil. Theories in economy tell us that this process continues until the markets reach an equilibrium state in which the price in dollars per barrel $\beta = \frac{\partial S}{\partial \text{doll}}$ is the same on both sides. In this equilibrium state some oil will continue to fluctuate between the two places, but there will be no net flow any more, hence, both trading places will be in detailed balance.
The analogy with trading places allows us to understand intuitively the meaning of ‘hot’ and ‘cold’. If the system is hot, it can offer a lot of energy for very little entropy, i.e., the price of energy in terms of entropy is a low. Conversely, a cold system has only little energy to sell and therefore the price of energy is high. If two such systems interact with each other, the cold system will be interested in buying energy from the hot one at a very low price. As a result, energy flows from the hot to the cold system.

**Summary:** Interpretation of temperature:

\[ \beta = \frac{1}{T} \] is the price at which a system can provide energy in return of entropy. Hot systems offer energy cheaply. Contrarily, buying energy from a cold system is expensive.

### 3.4. Conserved quantities and external reservoirs

**Systems coupled to a heat bath**

**Heat baths and free energy:** If one of the two systems is so large that the smaller system cannot influence the ‘price’ of the exchanged physical quantity in terms of entropy, the larger one is called *reservoir*. In what follows, we will use the suffixes ‘sys’ for the laboratory system and ‘env’ for the reservoir or synonymously the environment.

In the case of energy exchange, the reservoir is usually referred to as a “heat bath”, which is characterized by a given constant temperature \( T_{\text{env}} \). The constant temperature allows us to relate energy and entropy changes in the heat bath without knowing its actual internal structure:

\[ dE_{\text{env}} = T_{\text{env}} \, dS_{\text{env}} . \]  
(3.64)

Since energy is conserved we have

\[ dE_{\text{sys}} = - dE_{\text{env}} = - T_{\text{env}} \, dS_{\text{env}} . \]  
(3.65)

By integrating this relation we obtain\(^5\)

\[ E_{\text{sys}} = \text{const} - T_{\text{env}} S_{\text{env}} . \]  
(3.66)

\(^5\)This integration has to be taken with a grain of salt. On the one hand, the integration is only possible in case of an exact differential. On the other hand, it is not fully clear whether the environment can be integrated at all, producing a finite integration constant, and whether there exists definite entropy of the environment. For simplicity, we ignore these questions at this point.
Defining the Helmholtz free energy

\[ F_{\text{sys}} = E_{\text{sys}} - T_{\text{env}} S_{\text{sys}}. \quad (3.67) \]

(written as \( F = U - TS \) in most textbooks) we find that

\[ F_{\text{sys}} = \text{const} - T_{\text{env}} (S_{\text{sys}} + S_{\text{env}}) = \text{const} - T_{\text{env}} S_{\text{tot}}. \quad (3.68) \]

As can be seen on the right-hand side of this equation, the maximization of the total entropy is tantamount with the minimization of \( F_{\text{sys}} \). Therefore, the potential called free energy, which looks as if it was a property of the system alone, incorporates already the interaction with the heat bath without knowing its specific internal structure. In other words, from a formal point of view, the definition of the free energy can be seen as a clever trick to write the total entropy of the combined system in form of a potential for a single system. Therefore, whenever we deal with a free energy, we have to keep in mind that the system is tacitly assumed to be coupled to an external heat bath.

**Summary:** The meaning of the “free energy”:
The free energy \( F \) is nothing but a different way to write the total entropy of the system plus heat bath (up to a minus sign and a constant) in the form of a single potential of the system alone. It may be seen as a formal trick to write the combined entropy in a more convenient form, but it is important to realize that does not introduce a new kind of Second Law.

**Individual and mean free energy:** Let us be a bit more precise in deriving the free energy: For a given configuration \( s \in \Omega^{\text{sys}} \) the system has a well-defined energy \( E_{\text{sys}}(s) \), which we will denote here as the configurational energy. The main assumption is that upon each microscopic fluctuation of the system the heat bath equilibrates so quickly that it responds to an energy change with an immediate well-defined change of its entropy:

\[ \Delta S_{\text{env}} = \beta \Delta E_{\text{env}}. \quad (3.69) \]

Thus, whenever the system undergoes a transition \( s \to s' \), the energies and entropies change instantaneously as

\[ \Delta E_{\text{env}}(s \to s') = -\Delta E_{\text{sys}}(s \to s') = -(E_{\text{sys}}(s') - E_{\text{sys}}(s)) \quad (3.70) \]
\[ \Delta S_{\text{env}}(s \to s') = -\beta (E_{\text{sys}}(s') - E_{\text{sys}}(s)) \quad (3.71) \]

implying that the entropy of the environment is a well-defined function of the system’s configuration \( s \):

\[ S_{\text{env}}(s) = \text{const} - \beta E_{\text{sys}}(s). \quad (3.72) \]
Therefore, we can express the entropy of the total system solely in terms of the configuration of the subsystem:

\[ S_{\text{tot}}(s) = \text{const} + S_{\text{sys}}(s) - \beta E_{\text{sys}}(s), \]  

(3.73)

This allows us to define the *configurational free energy*

\[ F_s = E_{\text{sys}}(s) - TS_{\text{sys}}(s). \]  

(3.74)

Note that this definition associates with each configuration \( s \in \Omega^{\text{sys}} \) of the subsystems an individual free energy. The ordinary definition of the free energy, which can be found in most textbooks, is obtained by averaging over all subsystem configurations:

\[ F = \langle F_s \rangle = \langle E_{\text{sys}}(s) \rangle - T\langle S_{\text{sys}}(s) \rangle = U - TS. \]  

(3.75)

The average energy \( U = \langle E \rangle \) is called the *internal energy* of the subsystem. Clearly, minimizing the average free energy \( F \) is equivalent to maximizing the averaged total entropy \( \langle S_{\text{tot}} \rangle \).

**Summary:** Configurational free energy:

As in the case of entropy, where we distinguish between the configurational entropy \( S_s \) and the average entropy \( S \), we can define a configurational free energy \( F_s \) for every system configuration \( s \in \Omega^{\text{sys}} \) and an average free energy \( F \), provided that the environment equilibrate almost instantaneously.
4. Equilibrium phase transitions

In this chapter we discuss an important sub-discipline of equilibrium statistical mechanics, namely, *phase transitions* and *critical phenomena* in stochastic lattice models at thermal equilibrium. Roughly speaking, the term “phase transition” stands for a sudden change of the macroscopic behavior when external parameters of the surrounding reservoir, e.g. the temperature or the pressure, are varied.

**Definition of an equilibrium system**

In contrast to a general nonequilibrium system, which is defined by a configuration space $\Omega^{\text{sys}}$ with certain transition rates $w_{s \rightarrow s'}$ and where we can compute the entire dynamics as a function of time, an equilibrium system is defined solely by its configuration space $\Omega^{\text{sys}}$ and a map which associates with each configuration a well-defined energy

$$s \mapsto E_s \quad (4.1)$$

In equilibrium statistical mechanics, we are only interested in the *stationary equilibrium state*, where the system is expected to obey *detailed balance*. Depending on the statistical ensemble (that is, the type of physical interaction with the environment which is assumed to be large), the probability distribution in the stationary state is fully determined so that we don’t have to solve the master equation explicitly. For example, for a system in contact with a thermal reservoir at temperature $T$, we already know that the stationary probability distribution is given by

$$P_s = \frac{1}{Z(\beta)} e^{-\beta E_s}, \quad (4.2)$$

where $\beta = 1/T$. Here $Z(\beta)$ is the so-called *partition sum* which serves as a normalization factor:

$$Z(\beta) = \sum_{s \in \Omega^{\text{sys}}} e^{-\beta E_s} \quad (4.3)$$

For simplicity let us begin with the simplest situation, namely, with systems exchanging energy with a heat bath. In this case the potential to be extremalized is

$$V = H^{\text{sys}} - \beta E \quad \text{or, in textbook form:} \quad F = E - TH. \quad (4.4)$$

As can be seen, the Helmholtz free energy contains two terms with different signs. One of them is the internal energy, the other one is essentially the entropy of the system. Hence for systems exchanging energy with their surrounding, internal energy and entropy compete one another. If these two components are in favor of different macroscopic situations, such a competition could lead to a phase transition when the
temperature is varied.

4.1. Ising model: Phenomenology

As we have seen in the preceding chapter, stochastic systems in contact with external reservoirs approach an equilibrium state in which the entropy of the system + environment is maximized. Moreover, we have shown that the maximization of the entropy is equivalent to finding the extremum of an appropriate thermodynamic potential $V$. The equilibrium state itself is characterized by a condition called “detailed balance”, stating that all probability currents in the system between pairs of different configurations cancel each other.

The “harmonic oscillator” of systems with an equilibrium phase transition is the celebrated Ising model, named after Ernest Ising, who investigated this model for the first time in his PhD thesis. The Ising model is a lattice model of interacting classical spins $s_i$ whose quantum nature is taken into account only in so far as to they can be either “up” or “down”, corresponding to the values $s_i = \pm 1$.

As any model in equilibrium statistical mechanics the Ising model is solely defined by its energy functional $E_s$, which associates with each system configuration $s \in \Omega$ a well-defined energy $E_s$. In the Ising model this energy functional is defined in such a way that aligned spins are favored. This introduces some kind of surface tension, having the tendency of reducing the surface between domains of equally oriented spins and therefore ordering the system. On the other hand, entropy favors disorder, trying to orient the spins randomly. Obviously these two mechanisms compete one another. As we will see in the following, this can lead to an order-disorder phase transition.
4.1 Ising model: Phenomenology

4.1.1. Definition of Ising model

The Ising model is defined on a regular \( d \)-dimensional lattice of sites enumerated by some index \( i \) (see Fig. 4.1). The lattice structure goes into the definition of the model by defining which of the sites are nearest neighbors. In order to indicate that two sites \( i \) and \( j \) are nearest neighbors, we will use the notation \( \langle i, j \rangle \). For example, a sum running over pairs of neighboring sites will be denoted by \( \sum_{\langle i, j \rangle} \). With this notation the energy functional of the Ising model is given by

\[
E_s = -J \sum_{\langle i, j \rangle} s_i s_j, \tag{4.5}
\]

where \( J \) is the coupling constant. For \( J > 0 \), the energy is least when the spins are aligned. Furthermore, we may apply an additional external magnetic field \( h \), extending the energy functional by a second term

\[
E_s = -J \sum_{\langle i, j \rangle} s_i s_j - h \sum_i s_i. \tag{4.6}
\]

This expression holds for arbitrary lattices. In the special case of the one-dimensional lattice with periodic boundary conditions, where the lattice sites are enumerated by \( i = 1, \ldots, L \), we have

\[
E_s = -J \sum_{i=1}^{L} s_i s_{i+1} - h \sum_{i=1}^{L} s_i. \tag{4.7}
\]

Note that in the case of a vanishing external magnetic field \( h = 0 \) the Ising model is symmetric under the replacement \( s_i \to -s_i \), i.e., the Ising model exhibits a \( \mathbb{Z}_2 \) symmetry under a global reversal of all spins. Clearly, this symmetry is broken if an external magnetic field \( h \) is applied.

4.1.2. Thermostatics of the Ising model

The Ising model is assumed to be in contact with a thermal reservoir at temperature \( T = 1/\beta \). This is a realistic assumptions since almost all experiments, where the spins are located on the surface of a substrate, are in thermal contact with the supporting three-dimensional solid-state body. Allowing this spin system to exchange energy with environment, its probability distribution will be given by the Boltzmann weight

\[
P_s = \frac{1}{Z} \exp(-\beta E_s). \tag{4.8}
\]

The thermodynamic potential in the canonical ensemble, which is equivalent to the Helmholtz free energy, is

\[
\mathcal{V} = H_{\text{sys}} - \beta E = \ln Z, \tag{4.9}
\]

where the partition sum

\[
Z(\beta, h) = \sum_{s \in \Omega_{\text{sys}}} \exp \left( \beta J \sum_{\langle i, j \rangle} s_i s_j + \beta h \sum_i s_i \right) \tag{4.10}
\]
Equilibrium phase transitions

depends on the inverse temperature $\beta = 1/T$ and on the external magnetic field $h$. In the ordinary notation this corresponds to the free energy

$$F = E - TH = -\frac{1}{\beta} \ln Z = -\frac{1}{\beta} V.$$  \hfill (4.11)

Sometimes it is also convenient to define the thermodynamic potential per lattice site

$$v = \frac{V}{N},$$  \hfill (4.12)

where $N$ is the total number of sites. As we will see below, this allows us to define specific (i.e. volume-independent) properties of the system.

Before proceeding, let us define the most important observables of interest.

- **Magnetization**

  The most important quantity is the magnetization, which plays the role of an order parameter at the phase transition. We distinguish between the total configurational magnetization

  $$M_s = \sum_i s_i,$$  \hfill (4.13)

the total mean magnetization

$$M(\beta, h) = \langle M_s \rangle = \sum_{s \in \Omega^\text{sys}} P_s M_s,$$  \hfill (4.14)

and the corresponding average magnetization per site

$$m(\beta, h) = \frac{1}{N} M(\beta, h).$$  \hfill (4.15)

Because of Eq. (4.10) these averages can be expressed as partial derivatives with respect to the external field:

$$M(\beta, h) = \frac{1}{\beta} \frac{\partial Z(\beta, h)}{\partial h} = \frac{1}{\beta} \frac{\partial V(\beta, h)}{\partial h}$$  \hfill (4.16)

and

$$m(\beta, h) = \frac{1}{\beta} \frac{\partial v(\beta, h)}{\partial h}.$$  \hfill (4.17)

Note that the configurational total magnetization $M_s$ appears in the energy functional $E_s = -J \sum_{ij} s_i s_j - h M_s$, where it is multiplied by $h$, the external magnetic field. The two quantities, the internal order parameter $M_s$ and the external field $h$, are said to form a conjugate pair. This is very common in physics. Whenever we have an observable $O$ of interest, it is useful to define a conjugate field or current $J$ and to add the product $\pm JO$ to the action or the energy functional. Then, the mean of $O$ can be computed by differentiating with respect to $J$.

- **Susceptibility**
Furthermore, we are interested in the susceptibility, defined as the response of the magnetization per site to a variation of the external field

\[ \chi(\beta, h) = \frac{\partial m(\beta, h)}{\partial h}, \quad (4.18) \]

which can also be expressed as

\[ \chi(\beta, h) = \frac{1}{\beta^2} \frac{\partial^2 v(\beta, h)}{\partial h^2}. \quad (4.19) \]

Usually the susceptibility increases as we approach the phase transition and finally diverges at the critical point, as shown in the figure.

- **Energy**

Likewise, another quantity of interest is the configurational energy \( E_s \), the mean energy

\[ E(\beta, h) = \langle E_s \rangle = \sum_{s \in \Omega_{\text{sys}}} P_s E_s. \quad (4.20) \]

and the corresponding average energy per site

\[ e(\beta, h) = \frac{1}{N} E(\beta, h). \quad (4.21) \]

Again, these averages may be expressed as partial derivatives of the potential

\[ E(\beta, h) = - \frac{\partial V(\beta, h)}{\partial \beta}, \quad e(\beta, h) = - \frac{\partial v(\beta, h)}{\partial \beta}. \quad (4.22) \]

- **Heat capacity**

Finally, the total and the specific heat capacity are given by

\[ C(\beta, h) = \frac{\partial^2 V(\beta, h)}{\partial \beta^2}, \quad c(\beta, h) = \frac{\partial^2 v(\beta, h)}{\partial \beta^2}. \quad (4.23) \]
4.1.3. Ising mean field theory

The Ising model is defined in such a way that each spin feels only the local magnetic field $h_i$ of its nearest neighbors and, if existent, the additional external magnetic field $h$. In fact, we can rewrite the energy functional as

$$E_s = - \sum_i (Jh_i + h) s_i , \quad h_i = \sum_{j \in \langle i \rangle} s_j .$$

(4.24)

where the sum runs over all nearest neighbors $j$ of site $i$.

The mean field limit is an approximation in which the local magnetic field is replaced by the average magnetic field $h_i \approx 2d \langle s \rangle$ (4.25) caused by all spins, where $d$ is the dimension of the system, giving

$$E_s \approx = - \sum_i \left( 2dJ \langle s \rangle + h \right) s_i .$$

(4.26)

Apart from the indirect coupling via the global magnetic mean field, $\bar{H}$, all spins are now decoupled from each other and fluctuate independently. Therefore, each of them can be seen as a single spin exchanging energy with the heat bath of temperature $T$. As such, each spin is oriented according to a probability given by the Boltzmann factor. That is, the probability distribution of a spin configuration $\{ s_i \}$ factorizes and is given by

$$P_s = \frac{\exp(-\beta E_s)}{Z} = \frac{1}{Z} \prod_i e^{-\beta E(s_i)} = \prod_i p(s_i)$$

(4.27)

with

$$p(s_i) = \frac{e^{+\beta Hs_i}}{e^{-\beta H} + e^{+\beta H}} .$$

(4.28)

Therefore, the average spin magnetization $\langle s \rangle$ is given by

$$\langle s \rangle = p(i) - (1 - p(i)) = \frac{e^{+\beta H} - e^{-\beta H}}{e^{+\beta H} + e^{-\beta H}} = \tanh(\beta H)$$

(4.29)

i.e.,

$$\langle s \rangle = \tanh\left( 2dJ\beta \langle s \rangle + \beta h \right) .$$

(4.30)

Note that there is an interesting twist in this equation. On the one hand, we use the average magnetic field to determine the Boltzmann factors for the spin orientation, but on the other hand, these Boltzmann factors in turn determine the average magnetic field. In other words, in order to compute the average magnetic field we already need to know the average magnetic field.

Assuming that the average magnetic field fluctuates only slowly compared to the individual spins, this results into a self-consistent feedback loop, allowing us to interpret Eq. (4.30) as an implicit equation. For a vanishing external magnetic field $h = 0$ this
equation has the obvious solution
\[ \langle x \rangle = 0. \quad (h = 0) \quad (4.31) \]

Furthermore, Eq. (4.30) has further nontrivial solutions which cannot be given in a closed form. However, it is possible to plot them parametrically. To this end, it is most convenient to introduce the new variables
\[ T_c = 2dJ, \quad z = \beta \bar{H} \quad (4.32) \]
such that
\[ z = \beta T_c \langle s \rangle + \beta h. \quad (4.33) \]

Then in terms of the parameter \( z \) we can compute the average magnetization
\[ \langle s \rangle = \tanh z \quad (4.34) \]
and the quotient
\[ \frac{T}{T_c} = \frac{\tanh z}{z} + \frac{h}{Jz}. \quad (4.35) \]

This allows us to draw the magnetization as a function of the temperature quotient \( T/T_c \) as a parametric plot.

\textbf{Spontaneous symmetry breaking:}

The result for vanishing external magnetic field \( h = 0 \) is shown in figure 4.3. In this figure the trivial solution \( \langle s \rangle = 0 \) is indicated as a horizontal blue line. Obviously, this solution is mapped onto itself under the \( Z_2 \) symmetry transformation \( s_i \rightarrow -s_i \).

In addition, for \( T < T_c \), there is another pair of solutions, marked by red lines. In this situation, the disordered influence of temperature is so weak that the spins order themselves due to the positive feedback via the mean field \( \bar{H} \). Due to the \( Z_2 \)-symmetry there are two symmetric solutions with positive and negative magnetization. Actually, the system selects only one of them, meaning that the \( Z_2 \)-symmetry is spontaneously

Figure 4.3.: Average magnetization of the Ising model for a vanishing external magnetic field \( h = 0 \), as predicted by mean field theory. The two solutions of the mean field equations are shown as a blue line (zero magnetization) and a red line (spontaneous magnetization).
Equilibrium phase transitions

Figure 4.4: Average magnetization of Ising model for a small positive external magnetic field $h > 0$ within mean field theory. The dashed lines indicate the previous solution for $h = 0$.

broken. The solution with vanishing magnetization (the blue line) still exists, but it becomes unstable for $T < T_c$, so that any fluctuation will drive the system away from the blue line to one of the red branches. In physics, this phenomenon is known as spontaneous symmetry breaking.

The same phenomenon happens in ferromagnetic materials. If we increase the temperature of a ferromagnet, its magnetic field will decrease and finally vanish at the so-called Curie-temperature (e.g. 1043 Kelvin for iron). Here, the material exhibits a phase transition from the ferromagnetic to the paramagnetic phase. Cooling it down again, the spins order spontaneously, magnetizing the material randomly in positive or negative direction.

**Influence of the external magnetic field:**

Let us now switch on the external magnetic field $h$. Obviously, the external magnetic field will break the $Z_2$ invariance. This can be seen very nicely when plotting the corresponding magnetization curves (see Fig. 4.4), where the red lines are distorted and no longer symmetric to each other. Moreover, the solution with zero magnetization does not exist anymore.

To understand this figure, let us assume that the system is magnetized in negative direction. For sufficiently small temperatures, the external magnetic field positive direction is not yet strong enough to flip the global orientation of the magnetization. However, as temperature is increased, there will be a certain threshold (indicated by the green arrow) where the system suddenly jumps to the upper branch. This occurs below the aforementioned critical temperature $T_c = 1/\beta_c$.

**Applicability of the mean field approximation:**

The mean field approximation relies on the assumption that each spin interacts effectively with all other spins of the system. Therefore, mean field theory ignores the notion of nearest neighbors and the lattice structure completely; the dimension of the lattice enters only as an effective number of nearest neighbors, the so-called coordination num-
ber of the interaction.

Usually, the quality of the mean field approximation increases with the dimension of the system. Roughly speaking, a high-dimensional system facilitates diffusive mixing, bringing the system closer to the mean field limit. In fact, in the limit of infinitely many dimensions, every site interacts with infinitely many other sites, and therefore mean field theory is expected to become exact.

As we will see below, there is actually a so-called upper critical dimension \( d_c < \infty \) above which mean field theory applies\(^1\). In low-dimensional systems, however, the local interactions lead to correlations among neighboring sites, resulting in a different behavior. Loosely speaking, the spins cannot be longer thought of as a structureless soup, instead they form stochastically correlated patterns which are relevant for the macroscopic behavior, by studying low dimensional, in particular one-dimensional systems, we can learn a lot about the influence of such correlation effects.

\[ T_{c}^{\text{MF}} = 2dJ. \] (4.36)

Such a transition, where the order parameter (the magnetization) vanishes continuously at the critical point, is referred to as a continuous or second-order phase transition.

Remark: In the theory of critical phenomena one distinguishes between various types of phase transitions. In experimental physics, the most frequent ones are discontinuous phase transitions, where the order parameter jumps abruptly at the critical point. Continuous phase transitions are less frequent and appear usually at the ending points of discontinuous phase transition lines. In theoretical physics continuous phase transitions are particularly interesting because they exhibit universal properties, as will be discussed below. Contrarily, discontinuous phase transitions are usually non-universal.

Scale invariance and power laws:

In physics, continuous phase transitions are usually characterized by scale invariance and long-range correlations. Scale invariance stands for a phenomenological approach based on the assumption that the model at the critical point does not exhibit any intrinsic length scale apart from the lattice spacing, provided that the system is infinite. This postulate imposes constraints on the possible functional form of physical observables. For example, it rules out that the correlation function has an exponential form

\(^1\)For \( d_c < d < \infty \) mean field theory applies in so far as it predicts the correct critical exponents, while the proportionality factors in the power laws may be different. Right at the critical dimension the power laws are superseded by logarithmic corrections.
$C(r) = e^{r/r_0}$ because an exponential function would require a certain reference scale $r_0$. Likewise, any other functional dependence represented by a power series is forbidden because it would require a reference scale as well. The only functional dependence, which is scale invariant by itself, is a power law dependence such as $C(r) = r^{-\alpha}$.

In fact, power laws are ubiquitous in the theory of continuous phase transitions and critical phenomena. However, usually such power laws do not extend over the full range, instead they hold only asymptotically, e.g. in the limit $r \to \infty$. The asymptotic limit in which the power law becomes exact is referred to as the scaling regime. Far away from the asymptotic scaling regime we expect corrections to scaling caused by the underlying lattice structure. Moreover, a power law involves an unknown proportionality factor which depends on the specific realization of the model. Ignoring this prefactor we indicate an asymptotic power law by writing e.g.

$$C(r) \sim r^{-\alpha}. \quad (4.37)$$

Here, the symbol ‘$\sim$’ has the meaning of being “asymptotic proportional to”.

**Scaling properties of the Ising model:**

In the context of the Ising model, the critical behavior is characterized by various power laws with associated critical exponents, which for historical reasons have been termed as $\alpha, \beta, \gamma, \delta, \ldots$. As we will see later, only two of them are independent while the other ones are related by so-called scaling relations.

For example, we may study how the average magnetization per site for a vanishing external field $h = 0$ scales with the temperature distance from the critical point. This means that we are looking for an exponent $\tilde{\beta}$ such that

$$m(\beta, 0) \sim (\beta - \beta_c)^{\tilde{\beta}} \sim (T_c - T)^{\tilde{\beta}}. \quad (4.38)$$

Further interesting quantities are the susceptibility $\chi$ and the specific heat $c$ in the ferromagnetic phase near the critical point for vanishing external field. These quantities scale algebraically as

$$\chi(\beta, 0) \sim (\beta - \beta_c)^{-\gamma} \sim (T_c - T)^{-\gamma} \quad (4.39)$$

$$c(\beta, 0) \sim (\beta - \beta_c)^{-\alpha} \sim (T_c - T)^{-\alpha}. \quad (4.40)$$

with negative exponents, indicating that both quantities diverge at the critical point.

Finally, an important quantity of interest is the connected part of the spin-spin correlation function

$$C(i,j,\beta,h) := \langle s_is_j \rangle - \langle s_i \rangle \langle s_j \rangle. \quad (4.41)$$

In a translationally invariant Ising model we expect this correlator to depend only on the distance $r = |\vec{r}_i - \vec{r}_j|$ between sites $i$ and $j$:

$$C(i,j,\beta,h) = C(r,\beta,h) \quad (4.42)$$

---

2The critical exponent $\tilde{\beta}$ must not be confused with the temperature parameter $\beta = 1/T$. Therefore, the exponent is marked by an additional tilde. In the literature the tilde is usually missing.
In the ferromagnetic ordered phase, this correlation function is usually found to decay in the asymptotic limit $r \to \infty$ exponentially as

$$C(r, \beta, 0) \sim e^{-r/\xi(\beta, 0)}, \quad (4.43)$$

where $\xi(\beta, h)$ is the prevailing correlation length in the system. Approaching the critical point this correlation length increases and eventually diverges algebraically at the critical point as

$$\xi(\beta, 0) \sim (\beta - \beta_c)^{-\nu} \sim (T_c - T)^{-\nu}, \quad (4.44)$$

where $\nu$ is the critical exponent associated with the correlation length.

Precisely at the transition point the correlation length $\xi$ is infinite, indicating scale invariance. At this point the correlation function does no longer decay exponentially in the asymptotic regime $r \to \infty$, instead it displays an algebraic behavior

$$C(r, \beta_c, 0) \sim r^{2-d-\eta}, \quad (4.45)$$

where $d$ is the dimension of the system and $\eta$ is another critical exponent.

Finally, an interesting situation emerges at the critical point $T = T_c$, when a small external field is applied. Here we expect that the system responds to the external field $h > 0$ with the magnetization

$$m(\beta_c, h) \sim h^{1/\delta}, \quad (4.46)$$

where $\delta$ is yet another critical exponent. Since the susceptibility diverges at the critical point, the infinitesimal response to an external field is infinite, implying that $\delta > 1$.

**Scaling relations and universality:**

So far we have introduced six different critical exponents, namely, $\alpha, \tilde{\beta}, \gamma, \delta, \eta,$ and $\nu$. As already mentioned, only two of them are independent while the other ones are related by four simple relations, namely three ordinary *scaling relations*

- **Rushbrooke:** $\alpha + 2\tilde{\beta} + \gamma = 2$ \quad (4.47)
- **Widom:** $\gamma = \tilde{\beta}(\delta - 1)$ \quad (4.48)
- **Fisher:** $\gamma = (2 - \eta)\nu$ \quad (4.49)

and the so-called *hyperscaling relation*

- **Josephson:** $2 - \alpha = \nu d$. \quad (4.50)

The difference is the following: hyperscaling relations depend explicitly on the dimension $d$, while ordinary scaling relations do not.

As we will see, the amazing finding is that the critical exponent $\alpha, \tilde{\beta}, \gamma, \delta, \eta,$ and $\nu$ depend only on the dimension of the system and the symmetry of the interaction, but not on the lattice structure and the specific realization of the Ising model. Rather they take on the same values in a large variety of models which constitute the so-called Ising universality class.
Table 4.1: Critical exponents of Ising model in various dimensions.

<table>
<thead>
<tr>
<th>d</th>
<th>α</th>
<th>β</th>
<th>γ</th>
<th>δ</th>
<th>η</th>
<th>ν</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>−</td>
<td>1/8</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1/2</td>
<td>7/4</td>
<td>15</td>
<td>1/4</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.110(1)</td>
<td>0.3265(3)</td>
<td>1.27372(5)</td>
<td>4.789(2)</td>
<td>0.0364(5)</td>
<td>0.6301(4)</td>
</tr>
<tr>
<td>≥ 4</td>
<td>0</td>
<td>1/2</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1/2</td>
</tr>
</tbody>
</table>

Remember: Continuous phase transitions are usually characterized by asymptotic power laws (=algebraic decay). The corresponding exponents are expected to be universal, i.e., they coincide in all models with phase transitions belonging to the same class.

Critical exponents within mean field theory:

This is part of our tutorial.

To compute the critical exponent $\tilde{\beta}$ of the Ising model within mean field theory, let us Taylor-expand the parametric representation of the red curve:

\[
\langle s \rangle = \tanh z = z - \frac{z^3}{3} + O(z^5) \quad (4.51)
\]

\[
\frac{T}{T_c} = \frac{\tanh z}{z} = 1 - \frac{z^2}{3} + O(z^4) \quad (4.52)
\]

Solving the second equation for $z$ and inserting the resultant that the first one we obtain

\[
\langle s \rangle = \sqrt{\frac{3}{T_c}} (T_c - T)^{1/2} + \ldots . \quad (4.53)
\]

Hence the critical exponent (within mean field theory) is given by $\tilde{\beta} = 1/2$.

Ising critical exponents – General picture:

Studying the Ising model in various dimensions, the following picture emerges:

- **In one dimension**, the Ising model is always in the paramagnetic phase, i.e., a phase transition is absent. Therefore, we cannot define critical exponents. However, in some situations we may interpret the behavior of the one-dimensional Ising model as having a phase transition at zero temperature $T_c = 0$, corresponding to $\beta_c = \infty$.

Remark: To understand the physical reason why the Ising model in one dimension does not exhibit a phase transition, let us recall the physical mechanism behind the transition. Spontaneous symmetry breaking means that the system magnetizes itself in a particular direction without any external field. It does so because such a state is entropically favorable. This requires that the spontaneously magnetized state is stable against fluctuations.

For example, let us assume that the system is magnetized in the positive direction. Because of the interaction with the heat bath, the spins will nevertheless fluctuate, forming little droplets of spins in opposite direction. If such a droplet was able to grow it could destroy
the magnetized state. Therefore, a robust mechanism is needed, which eliminates minority islands of down spins. In the Ising model, this mechanism is due to an effective surface tension of domain walls. In fact, looking at the energy functional, we immediately see that neighboring spins oriented in opposite direction cost energy. In other words, domain walls between differently oriented domains are energetically punished. Consequently, the system tries to reduce the total length of domain walls, which effectively introduces some kind of surface tension.

Because of the surface tension, the droplets acquire a roundish form. In addition, small droplets tend to shrink. It is this mechanism which stabilizes the spontaneously magnetized state. Obviously, this mechanism can work only in space dimensions larger or equal than 2, the simple reason being that in one dimension domain walls are just points which cannot shrink. This explains why the Ising transition does not exist in one-dimensional systems.

According to a famous argument by Landau, this applies to all equilibrium critical phenomena, i.e., equilibrium phase transitions are impossible in one-dimensional systems.

- **In two dimensions**, the Ising model exhibits a phase transition and the corresponding critical exponents are given by simple rational values. In the theory of critical phenomena, rational values are an exception rather than the rule. The deep reason behind these rational values is the existence of the powerful symmetry in the two-dimensional case, the so-called conformal symmetry.

- **In three dimensions**, the Ising model still exhibits a nontrivial phase transition. An exact solution of this case is still unknown. The critical exponents can only be determined numerically and by means of field theoretic methods, suggesting that they are given by irrational values. If the reader could explain the values of these exponents this would guarantee her/him a professorship.

- **In four dimensions and above**, the interaction becomes so strongly mixed that the mean field approximation becomes valid. Phenomenologically, there exists usually a well-defined dimension, called the upper critical dimension $d_c$, at which a given model crosses over to mean field behavior. For the Ising model, the upper critical dimension is 4. For $d \geq d_c$, the mean field exponents are exact. Precisely at the critical dimension $d = d_c$, the mean field power laws are superseded by logarithmic corrections.

### 4.1.5. Numerical simulation of the Ising model

Recall that a stochastic Markov process is defined by its configuration space $\Omega^{\text{sys}}$ and a set of rates $w_{s \rightarrow s'}$. Contrarily, equilibrium models are solely defined in terms of their configuration space $\Omega^{\text{sys}}$ and their energy functional $E_s$. As a great advantage of equilibrium models, the stationary distribution $P_s = \frac{1}{Z} e^{-\beta E_s}$ is (almost) for free. However, the definition of equilibrium models does not provide any information about the relaxation modes. In other words, the energies $E_s$ tell us how the stationary state looks like but they don’t tell us how the stationary state is actually reached.

On the other hand, if we want to simulate the model on a computer, meaning that the computer just mimics the Markovian random dynamics, we have to know the transition rates. This means that we have to invent rates in such a way that the system relaxes into a stationary state which is just exactly given by the Boltzmann weights. It turns out that this procedure is not unique.
Figure 4.5: Typical screenshots of the two-dimensional Ising model in the equilibrium state. Left: Ordered phase below the critical point. Middle: Critical state at the critical point. Right: Supercritical disordered state.

How should the rates be designed? In principle they can be chosen freely, provided that they meet the following two criteria:

- The network of transitions has to be ergodic, that is, every configuration can be reached from any other configuration by a sequence of dynamic transitions.

- In the stationary state the dynamics should by the condition of detailed balance.

As discussed in section 3.2 on page 56 the condition of detailed balance states that the probability currents between pairs of configurations vanishes in both directions:

$$J_{s \rightarrow s'} = J_{s' \rightarrow s} \iff P_s w_{s \rightarrow s'} = P_{s'} w_{s' \rightarrow s} \quad \forall s, s' \in \Omega^{sys}. \quad (4.54)$$

Knowing that $P_s = \frac{1}{Z} e^{-\beta E_s}$ this means that the rates have to be chosen in such a way that obey the condition

$$\frac{w_{s \rightarrow s'}}{w_{s' \rightarrow s}} = \frac{P_{s'}}{P_s} = \exp(\beta(E_s - E_{s'})). \quad (4.55)$$

Thus, if we choose any rate, the condition of detailed balance determines the corresponding rate in opposite direction. In other words, if we choose the matrix elements in the upper triangle of the Liouvillian, detailed balance gives us the lower triangle, thereby halving the degrees of freedom.

This leaves us with enormous number of $|\Omega^{sys}|^2 - |\Omega^{sys}|)/2$ undetermined the rates of which sufficiently many have to be nonzero in order to make sure that the result-

Figure 4.6: Illustration of the meaning of detailed balance. The figure shows a system with three states A,B,C while the blue arrows symbolize transitions at rate 1. The left panel shows a situation where the probability currents cancel one another in the stationary state, obeying detailed balance. Clearly, the stationary state given by $p_A = p_B = p_C = 1/3$. The system shown on the right hand side possesses the same stationary state. However, there is a non-vanishing cyclic probability current, meaning that the system is still out of equilibrium.
ing transition network is ergodic. This gives rise to a virtually infinite zoo of legitimate dynamical procedures which all relax into the same equilibrium state. In practice, however, one introduces further conditions in order to reach one of the following goals:

- We may either seek for a dynamics, which mimics what is happening in nature as faithfully as possible. The most important examples are local spin flip algorithms.
- Alternatively, we may be looking for a dynamics, which is particularly efficient when implemented on a computer. as will be discussed below, this includes the so-called cluster algorithms.

Local spin flip dynamics

The simplest class of dynamical procedures for the Ising model involve only single spin flips. As a starting point, note that the energy functional authorizing model can be written as

\[ E_s = -\sum_i H_i s_i, \quad H_i = h + J \sum_{j \in <i>} s_j. \]  

(4.56)

Considering a single spin as a subsystem, its probability to be oriented in positive direction is given by the Boltzmann weight

\[ p_i = p_i(\uparrow) = \frac{e^{\beta H_i}}{e^{\beta H_i} + e^{-\beta H_i}}. \]  

(4.57)

**Heat bath dynamics:**

This suggests to introduce the following dynamics. For each update a site \(i\) is randomly selected and the probability \(p_i\) is computed according to the formula given above. Then the selected spin is oriented according to this probability, i.e., we generate a uniformly distributed random number \(z\) between zero and one and set the new value at the selected site to

\[ s_i^{\text{new}} := \text{sign}(p_i - z). \]  

(4.58)

Obviously, this dynamic a procedure doesn’t care about the previous orientation of the updated spin. As can be verified easily, this procedure amounts to introduce the following transition rates for local spin flips:

\[ w(\downarrow \to \uparrow) \propto e^{\beta H_i} \]  

\[ w(\uparrow \to \downarrow) \propto e^{-\beta H_i} \]  

(4.59)

(4.60)

The corresponding a ratio of the rates is

\[ \frac{w(\downarrow \to \uparrow)}{w(\uparrow \to \downarrow)} = e^{2\beta H_i}. \]  

(4.61)

In order to prove that these dynamical rules obey detailed balance. Let us compare the energy of the configurations before and after a spin flip. For example, if we consider
the transition $\downarrow \rightarrow \uparrow$, the resulting energy change is given by

$$\Delta E = E_{\text{new}} - E_{\text{old}} = E_{\uparrow} - E_{\downarrow} = -2H_i.$$  \hspace{1cm} (4.62)

Detailed balance is established if and only if

$$\frac{w(\downarrow \rightarrow \uparrow)}{w(\uparrow \rightarrow \downarrow)} = \frac{P_{\text{new}}}{P_{\text{old}}} = \frac{e^{-\beta E_{\text{new}}}}{e^{-\beta E_{\text{old}}}} = e^{-\beta \Delta E} = e^{+2\beta H_i},$$ \hspace{1cm} (4.63)

reproducing Eq. (4.61).

**Glauber dynamics:**

*Glauber dynamics* differs from heat bath dynamics in so far as the local spins are not oriented but flipped with a certain probability. More specifically, the following steps are carried out:

- A site $i$ is randomly selected.
- The probabilities $p_i$ are computed by Eq. (4.57).
- If $s_i = -1$ the spin is flipped with probability $p_i$. Otherwise, if $s_i = +1$, the spin is flipped with probability $1 - p_i$.

As can be verified easily, this amounts to the rates

$$w(\downarrow \rightarrow \uparrow) \propto e^{\beta H_i},$$ \hspace{1cm} (4.64)

$$w(\uparrow \rightarrow \downarrow) \propto e^{-\beta H_i},$$ \hspace{1cm} (4.65)

hence Glauber dynamics and heat bath dynamics are statistically equivalent.

**Metropolis dynamics:**

The Metropolis algorithm differs from Glauber dynamics in that one of the moves is carried out with certainty and not with a finite probability. The Metropolis algorithm consists of the following steps.

- Choose a site $i$ randomly.
- Compute the energy gain $\Delta E = E_{\text{new}} - E_{\text{old}} = 2s_iH_i$ that would result from carrying out a spin flip at site $i$ ($s_i$ is the orientation before the update) and set
- If $\Delta E < 0$, i.e., if the system goes into an energetically more favorable state, the spin at site $i$ is flipped with certainty (probability 1).
  Otherwise, if $\Delta E \geq 0$ the spin is flipped only with probability $p$.

This means that one accepts the spin flip with the probability

$$p := \min(1, e^{-\beta \Delta E}).$$ \hspace{1cm} (4.66)

and dismisses it otherwise.
4.2 Ising model: Transfer matrix and quantum limit

In the Metropolis algorithm the rates for local spin flips and no longer independent, instead they depend on the actual contract duration of the immediate environment, and coded in the sign of the local field \( H_i \). If \( H_i \geq 0 \) we have

\[
w(\downarrow \rightarrow \uparrow) \propto 1 \quad \text{and} \quad w(\uparrow \rightarrow \downarrow) \propto e^{-2\beta H_i}, \tag{4.67}
\]

while for \( H_i < 0 \) the rates are given by

\[
w(\downarrow \rightarrow \uparrow) \propto e^{2\beta H_i} \quad \text{and} \quad w(\uparrow \rightarrow \downarrow) \propto 1, \tag{4.68}
\]

In both cases, the ratio of the rates is again compatible with Eq. (4.61), establishing detailed balance.

Summary: Local spin flip dynamics for the Ising model
Choose a random site \( i \) and let

\[
p_i = \frac{e^{\beta H_i}}{e^{\beta H_i} + e^{-\beta H_i}}.
\]

Furthermore, let \( z \in [0, 1] \) be a random number. Then perform an update as follows:

- **Standard heat bath dynamics:**
  \[
s_{i}^{\text{new}} := \text{sign}(p_i - z)
\]

- **Glauber dynamics:**
  \[
s_{i}^{\text{new}} := \begin{cases} 
+\text{sign}(p_i - z) & \text{if } s_{i}^{\text{old}} = +1 \\
-\text{sign}(1 - p_i - z) & \text{if } s_{i}^{\text{old}} = -1
\end{cases}
\]

- **Metropolis dynamics:**
  \[
s_{i}^{\text{new}} := \begin{cases} 
+\text{sign}(p_i^+ - z) & \text{if } s_{i}^{\text{old}} = +1 \\
-\text{sign}(p_i^- - z) & \text{if } s_{i}^{\text{old}} = -1
\end{cases}
\]

where \( p_i^\pm = \min(1, e^{\pm 2\beta H_i}) \).

All the dynamical procedures listed above relax into the same equilibrium state of the Ising model. In the equilibrium stationary state the rates obey detailed balance.

4.2. Ising model: Transfer matrix and quantum limit

4.2.1. One-dimensional Ising model

**Transfer matrix**

We consider a one-dimensional classical Ising model with \( M \) Ising spins \( s_m = \pm 1 \), with the index \( m \) running from 0 to \( M - 1 \). The model is defined by the energy function

\[
E_s = -J \sum_{m=0}^{M-1} s_m s_{m+1} - h \sum_{m=0}^{M-1} s_m, \tag{4.71}
\]

\[\text{[Footnote]}\]

\[\text{[Footnote]}\] We use here \( M \) instead of \( N \) because in the 2D case we intend to study an \( M \times N \)-dimensional lattice by similar methods.
where $J$ is the coupling constant between neighboring spins and $h$ is the external magnetic field. If not specified otherwise, we use periodic boundary conditions, i.e. index operations are carried out modulo $M$, meaning that $s_M \equiv s_0$.

The model is assumed to be in contact with an external heat bath with inverse temperature $\beta = 1/T$. Maximizing the total entropy of system and environment (= minimizing the free energy), the system relaxes into a stationary equilibrium state described by the partition function
\[
Z(\beta, h) = \sum_s e^{-\beta E_s},
\]
where $s = \{s_0, \ldots, s_{M-1}\}$ denotes the configuration of all spins. In what follows it is convenient to rewrite the energy function in the form
\[
E_s = -JM_{\text{drop}} + \frac{J}{2} \sum_{m=0}^{M-1} (s_m - s_{m+1})^2 - \frac{h}{2} \sum_{m=0}^{M-1} (s_m + s_{m+1}).
\]

Because of the quadratic term in the first some we get an additional constant contribution $-JM$. Since the energy is only defined up to a constant and since such contributions cancel in the partition sum by normalization, we can simply drop this constant. This allows us to rewrite the partition sum by
\[
Z(\beta, h) = \sum_{s} T(s_m, s_{m+1})
\]
where
\[
T(s_m, s_{m+1}) = \exp \left[ -\frac{\beta J}{2} (s_m - s_{m+1})^2 + \frac{\beta h}{2} (s_m + s_{m+1}) \right]
\]
accounts for the energy contributions related to adjacent sites. Since $s_m = \pm 1$ and $s_{m+1} = \pm 1$ this expression takes one of the following four possible values:
\[
T(+1, +1) = e^{\beta h} \\
T(+1, -1) = e^{-2\beta J} \\
T(-1, +1) = e^{-2\beta J} \\
T(-1, -1) = e^{-\beta h}
\]

Let us now define the canonical basis vectors of the single-site configuration space, namely $|+\rangle = \left( \begin{array}{c} 1 \\ 0 \end{array} \right)$ and $|\rangle = \left( \begin{array}{c} 0 \\ 1 \end{array} \right)$. With these vectors we can express the four contributions listed above in the form
\[
T(s_m, s_{m+1}) = \langle s_m | T | s_{m+1} \rangle
\]
where
\[
T = \left( \begin{array}{cc} e^{\beta h} & e^{-2\beta J} \\ e^{-2\beta J} & e^{-\beta h} \end{array} \right)
\]
is the so-called transfer matrix which, in the present case of the one-dimensional Ising model, is just 2×2-matrix. The name “transfer matrix” comes from the fact that for a
given probability distribution of the orientation of $s_n$, the application of the transfer matrix renders the probability distribution of the spin at the next site $n + 1$. In other words, the transfer matrix is the operator for translation on the Ising lattice by one unit. With the transfer matrix the partition sum can be written as

$$Z(\beta) = \sum_{s_0 = \pm 1} \langle s_0 | T | s_1 \rangle \sum_{s_1 = \pm 1} \langle s_1 | T | s_2 \rangle \sum_{s_2 = \pm 1} \langle s_2 | T | ... T | s_{M-1} \rangle \sum_{s_{M-1} = \pm 1} \langle s_{M-1} | T | s_0 \rangle$$

which, as can be seen, reduces to the simple expression

$$Z(\beta) = \text{Tr}[T^M].$$

Correlation functions for vanishing external field

The transfer matrix contains all the information that is needed to compute arbitrary $n$-point correlation functions. For simplicity let us restrict ourselves to the case of vanishing external field $h = 0$ (the general case works similarly, but the formulas are more involved). Solving the eigenvalue problem $T|\psi\rangle = \lambda|\psi\rangle$ one finds two eigenvalues and eigenvectors:

$$\lambda_{1,2} = 1 \pm e^{-2\beta J}, \quad |\psi_{\pm}\rangle = \frac{1}{\sqrt{2}} \left( \begin{array}{c} 1 \\ \pm 1 \end{array} \right)$$

We are now interested in computing the two-point spin-spin correlation function

$$\langle s_i s_j \rangle = \frac{1}{Z} \sum_s e^{-\beta E_s} s_i s_j = \frac{1}{Z} \text{Tr}[T^i \sigma^z T^j - i \sigma^z T^{M-1}].$$

Rotating the transfer matrices cyclically this turns into

$$\langle s_i s_j \rangle = \frac{1}{Z} \text{Tr}[T^{M-j+i} \sigma^z T^{j-i} \sigma^z].$$

Inserting identities $1 = |\psi_+\rangle\langle \psi_+ | + |\psi_-\rangle\langle \psi_- |$ and using the fact that

$$\langle \psi_+ | \sigma^z | \psi_+ \rangle = \langle \psi_- | \sigma^z | \psi_- \rangle = 0, \quad \langle \psi_+ | \sigma^z | \psi_- \rangle = \langle \psi_- | \sigma^z | \psi_+ \rangle = 1$$

it is straightforward to compute the correlator:

$$\langle s_i s_j \rangle = \frac{\lambda_+^{M-j+i} \lambda_-^{j-i} + \lambda_-^{M-j+i} \lambda_+^{j-i}}{\lambda_+^M + \lambda_-^M}.$$ 

In this sense we have solved the one-dimensional Ising model exactly by means of a transfer matrix method.

In the thermodynamic limit $M \to \infty$ the correlations function reduces to

$$\langle s_i s_j \rangle = \left( \tanh(\beta J) \right)^{|j-i|}.$$
Comparing this result with an asymptotic exponential decay $\langle s_i s_j \rangle \sim e^{-|j-i|/\xi}$, the solution of the equation $\tanh J = \exp\left(-1/\xi\right)$ gives the correlation length

$$\xi = \frac{1}{\ln(\tanh \beta J)} = \frac{1}{\ln(\coth \beta J)}$$

(4.86)

This expression is finite for all $\beta$ and diverges in the limit $\beta \to \infty$. This means that the Ising model in one dimension has no phase transition at finite temperature, but it behaves in a way as if the phase transition was located at $T = 0$. Of course, a ‘transition’ at $T = 0$ is not a transition in the usual sense since $T < 0$ is inaccessible so that it does not really separate two different phases.

**Relation to a quantum system**

In the following we would like to relate the one-dimensional classical Ising model to a zero-dimensional quantum system. The idea is that the Hamiltonian of the quantum system can be considered as the generator of the transfer matrix, meaning that the spatial dimension of the original model is reinterpreted as some kind of ‘time’ in the quantum system. In other words, instead of moving in space by means of the transfer matrix we would like to move in time by means of a suitable Hamiltonian. This Hamiltonian should be related to the transfer matrix by

$$T = \exp(-\tau H),$$

(4.87)

where $\tau$ depends on $J$ and $h$. Generally this is impossible, but there is a special limit in which such a relation can be established to leading order. Namely, if we consider the limit of large coupling constant $J \to \infty$ and if we tune $h$ in such a way that

$$e^{-2\beta J} = \tau \quad \text{and} \quad e^{\beta h} = 1 + \eta \tau$$

(4.88)

for some fixed $\eta > 0$ (which means that $\tau \to 0$ and $h \to 0$), then the transfer matrix can be written as

$$T = \left( \begin{array}{cc} 1 + \eta \tau & \tau \\ \tau & 1 - \eta \tau \end{array} \right) = 1 + \tau (\sigma^x + \eta \sigma^z) \approx \exp(-\tau H)$$

(4.89)

where

$$H = -\sigma^x - \eta \sigma^z$$

(4.90)

is the quantum Hamiltonian of the system. As this Hamiltonian involves two Pauli matrices acting on a single site, this quantum Hamiltonian can be considered as zero-dimensional.

**Remark:** The strong-coupling limit can be explained as follows. If we take $J \to \infty$, the neighboring spins are strongly coupled, that is, they are likely to be identical in their orientation. Therefore, the corresponding probability distributions of neighboring spins will differ only slightly, meaning that the transfer matrix is close to the identity: $T \approx 1 + \tau H$. 

© HAYE HINRICHS — PHYSICS OF COMPLEX SYSTEMS — CC BY-SA 4.0
Wick rotation

Note that in the previous example, the transfer matrix is given by $T = \exp(\tau H)$, which differs from a true quantum-mechanical problem, where we would deal with a unitary time evolution operator $U = \exp\left(\frac{i}{\hbar} H t\right)$. Nevertheless, we consider $H$ as a quantum Hamiltonian and treat it as such.

Both expressions differ by a so-called Wick rotation. A Wick rotation is basically an analytic continuation in the complex plane which allows us to map a problem of equilibrium statistical mechanics in $d$ dimension to the time evolution of a $d$-1-dimensional quantum problem.

More specifically, in equilibrium statistical mechanics, the expectation value of a quantity $A$ in a system with energies $E_s$ which is in contact with a heat bath reads

$$\langle A \rangle_\beta = \frac{1}{Z} \sum_s A_s e^{-\beta E_s}.$$  

We can easily construct an analogous quantum-mechanical problem. To this end consider a Hamiltonian $H$ on a finite-dimensional Hilbert space with eigenenergies $E_n$ and eigenstates $|n\rangle$. Furthermore, let us define a coherent superposition of all eigenstates $|\psi\rangle = \sum_n |n\rangle$ and an arbitrary state $|a\rangle = \sum_n a_n |n\rangle$. Suppose that $|\psi\rangle$ evolves in time by $|\psi(t)\rangle = e^{-\frac{i}{\hbar} H t} |\psi\rangle$. Then we get the expression

$$\langle a |\psi(t)\rangle = \sum_n a_n e^{-\frac{i}{\hbar} E_n t},$$

which formally looks (up to normalization) like Eq. (4.91), the essential difference being that the exponent is imaginary in quantum mechanics. Setting $t = -i \hbar \beta$ we can therefore we can interpret statistical mechanics as a quantum-mechanical problem evolving in imaginary time.

Mathematically, the key point of the Wick theorem is that the exponential function is holomorphic (analytic and differentiable) on $\mathbb{C}$, obeying the Cauchy-Riemann differential equations. The same applies of course to a finite sum of exponential functions. If a holomorphic function is given on the real line, we can use the Cauchy-Riemann differential equation to compute it everywhere, in particular on the imaginary line. Vice versa, if the function is given on the imaginary line, we can compute it on the real line. Thus, by solving a quantum problem in real time, the solution of the corresponding problem in statistical mechanics (with imaginary time) is obtained by analytic continuation.
4.2.2. Two-dimensional Ising model

We now consider the two-dimensional classical Ising model. As we have seen above, this case is much more interesting since the 2D-Ising model exhibits a phase transition at finite temperature.

The 2D-Ising is defined on a square lattice with $M \times N$ sites with classical Ising spins $s_{m,n} = \pm 1$, where $m \in \{0, \ldots, M - 1\}$ and $n \in \{0, \ldots, N - 1\}$. The Ising spins are coupled with the coupling constant $J_\perp$ in horizontal and with $J_\parallel$ in vertical direction. In addition, we have again an external field $h$:

$$E_s = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left( -J_\perp s_{m,n}s_{m,n+1} - J_\parallel s_{m,n}s_{m+1,n} - h s_{m,n} \right).$$ \hspace{1cm} (4.93)

Again we assume periodic boundary conditions in both directions, setting $s_{M,n} \equiv s_{0,n}$ and $s_{m,N} \equiv s_{m,0}$. Note that the resulting topology of the model is that of a torus. Moreover, note that for $J_\perp = J_\parallel$ and $N = M$, the model is completely symmetric in horizontal and vertical direction.

Kramers–Wannier duality

Again the partition sum of the Ising model is given by $Z(\beta) = \sum_s e^{-\beta E_s}$. As in the mean field approximation, it is expected that the second-order phase transition can only be observed if the external field $h$ vanishes. Thus the phase transition is controlled by effectively two parameters, namely,

$$K_\perp := \beta J_\perp \quad \text{and} \quad K_\parallel := \beta J_\parallel.$$ \hspace{1cm} (4.94)

Therefore, the 2D Ising model defined above does not have a single critical point, instead we expect it to have a whole line of critical points, as sketched schematically in the adjacent figure.

For the 2D Ising model on a square lattice, it is in fact possible to compute the critical line exactly. This amazing discovery goes back to Kramers and Wannier, who discovered a nontrivial duality in the Ising model. We will not go into great detail here, but it is interesting how they obtained this result. It turns out that there are two different ways to rewrite the partition sum of the Ising model, known as low- and high-temperature expansion. The low-temperature expansion reads

$$Z(K_\perp, K_\parallel) = 2^{2K_\perp + K_\parallel} \sum_{\text{islands}} (e^{-2K_\perp})^{n_\perp} (e^{-2K_\parallel})^{n_\parallel}.$$ \hspace{1cm} (4.95)

The sum runs over all oppositely oriented islands in the current configuration.
land is surrounded by a domain wall of oppositely oriented pairs of neighboring spins. The numbers \( n_\perp \) and \( n_\parallel \) count these pairs in horizontal and vertical direction, respectively. Although this expression is exact, it is particularly useful in the limit \( \beta \to \infty \). Here the Ising model is preferentially magnetized either in positive or negative direction, explaining the factor 2 in front.

Likewise it is possible to rewrite the partition sum as an expansion which is particularly useful in the limit of high temperatures. Marking the corresponding parameters by a tilde, this expansion reads

\[
Z(\tilde{K}_\perp, \tilde{K}_\parallel) = (2 \cosh \tilde{K}_\perp \cosh \tilde{K}_\parallel)^{MN} \sum_{\text{contours}} (\tanh \tilde{K}_\perp)^{\ell_\perp} (\tanh \tilde{K}_\parallel)^{\ell_\parallel}. \tag{4.96}
\]

Here the sum runs over all contours on the lattice with certain constraints which in the end ensure that it has formally the same structure as the sum over the islands in Eq. (4.95). Since the factor in front of the sum is physically irrelevant, the formal similarity of both expansions allows us to identify

\[
e^{-2K_\perp} = \tanh \tilde{K}_\parallel, \quad e^{-2K_\parallel} = \tanh \tilde{K}_\perp. \tag{4.97}
\]

This means that a two-dimensional Ising model with the parameters \((K_\perp, K_\parallel)\) can be mapped exactly to a two-dimensional Ising model with the parameters \((\tilde{K}_\parallel, \tilde{K}_\perp)\), thereby relating low- and high-temperature behavior. Note that the duality relation exchanges \( \perp \) and \( \parallel \). We can rewrite the equations given above by

\[
\sinh 2\tilde{K}_\perp = 2 \sinh \tilde{K}_\perp \cosh \tilde{K}_\parallel = 2 \tanh \tilde{K}_\perp \cosh^2 \tilde{K}_\parallel = \frac{2 \tanh \tilde{K}_\perp}{1 - \tanh^2 \tilde{K}_\perp} = \frac{2}{1 - e^{-4K_\parallel}} = \frac{2}{e^{2K_\parallel} - e^{-2K_\parallel}} = \frac{1}{\sinh 2K_\parallel},
\]

arriving at the relations

\[
\sinh 2K_\parallel \sinh 2\tilde{K}_\perp = 1, \quad \sinh 2\tilde{K}_\parallel \sinh 2K_\parallel = 1. \tag{4.99}
\]

This duality transformation tells us that the two-dimensional Ising model with \((K_\perp, K_\parallel)\) and \((\tilde{K}_\perp, \tilde{K}_\parallel)\) exhibit (up to the transformation) the same physics. Consequently, if the model exhibits a phase transition at \((K_\perp, K_\parallel)\), it will also exhibit a phase transition at \((\tilde{K}_\perp, \tilde{K}_\parallel)\). However, as we expect the Ising model to exhibit only a single phase transition and not several ones, we can conclude that the system will undergo a transition if \(K_\perp = \tilde{K}_\perp\) and \(K_\parallel = \tilde{K}_\parallel\). Therefore, the transition takes place along a critical line given by

\[
\sinh 2K_\parallel \crl \sinh 2\tilde{K}_\perp = 1. \tag{4.100}
\]

In particular, if the model is isotropic (meaning that \(K_\perp = K_\parallel = K\) we get \(\sinh K \crl = 1\), hence \(K \crl \approx 0.440687\), or equivalently:

\[
T_c = 2.26919 J. 
\]
Transfer matrix

Following the same line of arguments as in Eq. (4.73), we can write the partition sum of the 2D Ising model in terms of a transfer matrix acting in one of the two (here the vertical) direction. We first rewrite the energy as

\[ E_s = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left[ -J_\perp s_{m,n}s_{m,n+1} - J_\parallel s_{m,n}s_{m+1,n} - h s_{m,n} \right] \]  

(4.101)

as

\[ E_s = -\int_\text{drop} MN + \sum_{n=0}^{M-1} \sum_{m=0}^{N-1} \left[ -\frac{J_\perp}{2} (s_{m,n}s_{m,n+1} + s_{m+1,n}s_{m,n+1}) \right. \\
\left. + \frac{J_\parallel}{2} (s_{m,n} - s_{m+1,n})^2 - \frac{h}{2} (s_{m,n} + s_{m+1,n}) \right]. \]

(4.102)

Again we would like to express the partition sum as a product of the form

\[ Z(K_\perp, K_\parallel) = \sum_s \prod_{m=0}^{M-1} T(s_m, s_{m+1}) \]  

(4.103)

where \( s_m = \{s_{m,0}, \ldots, s_{m,N-1}\} \) is the configuration of the \( m^{\text{th}} \) row and

\[ T(s_m, s_{m+1}) = \exp \left[ \sum_{n=0}^{N-1} \left( \frac{K_\perp}{2} (s_{m,n}s_{m,n+1} + s_{m+1,n}s_{m,n+1}) \right. \right. \\
\left. \left. \left. - \frac{K_\parallel}{2} (s_{m,n} - s_{m+1,n})^2 \right) + \frac{\beta h}{2} (s_{m,n} + s_{m+1,n}) \right) \right] \]

(4.104)

accounts for the contributions in horizontal and vertical direction and for the interaction with the external magnetic field. Introducing a \( 2^N \)-dimensional vector space with the canonical configuration vectors

\[ |\vec{s}\rangle = |s_0\rangle \otimes |s_1\rangle \otimes \cdots \otimes |s_{N-1}\rangle, \quad |+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \]

(4.105)

and rewriting

\[ T(\vec{s}_m, \vec{s}_{m+1}) = \langle \vec{s}_{m+1}|T|\vec{s}_m\rangle \]

(4.106)

where \( T \) is the \( 2^N \times 2^N \) transfer matrix, we find that

\[ T = \prod_{n=0}^{N-1} \exp \left( \frac{K_\perp}{2} \sigma_n^x \sigma_{n+1}^x + \frac{\beta h}{2} \sigma_n^z \right) \]

(4.107)

Here \( \sigma_n^x \) (and similarly \( \sigma_n^z \)) denotes the Pauli matrix acting at site \( n \), i.e.

\[ \sigma_n^x = 1 \otimes 1 \otimes \cdots \otimes \sigma_n^x \otimes \cdots \otimes 1. \]

(4.108)
We now consider the strongly anisotropic limit $K_\perp \to 0, K_\parallel \to \infty$, defining
\[
e^{-2K_\parallel} = \lambda \tau, \quad K_\perp = \tau, \quad \beta h = \eta \tau. \quad (4.109)
\]
In this limit we can rewrite the transfer matrix to lowest order as
\[
T = \prod_{n=0}^{N-1} \left[ 1 + \tau \left( \sigma_n^z \sigma_{n+1}^z + \lambda \sigma_n^x + \eta \sigma_n^z \right) \right] \quad (4.110)
\]
which can be interpreted as being generated via $T = e^{-\tau H}$ by the "quantum" Hamiltonian
\[
H = -\sum_{n=0}^{N-1} \left( \sigma_n^z \sigma_{n+1}^z + \lambda \sigma_n^x + \eta \sigma_n^z \right). \quad (4.111)
\]
Because of Eq. (4.100) the critical point in the strongly anisotropic limit for vanishing external field is $K_\perp^{\text{crit}} = \lambda \tau$, which corresponds to taking $\lambda_c = 1$ and $\eta = 0$.

In the literature it is customary to cyclically replace the Pauli matrix by $\sigma^x \to \sigma^y \to \sigma^z \to \sigma^x$, which amounts to a similarity transformation of the Hamiltonian. Moreover the Hamiltonian is usually studied for vanishing external magnetic field $\eta = 0$. Thus, the final form of the Hamiltonian of the Ising quantum chain with $N$ sites reads
\[
H = -\sum_{n=0}^{N-1} \left( \sigma_n^z \sigma_{n+1}^z + \lambda \sigma_n^x \right) \quad (4.112)
\]
To summarize, $H$ is a $2^N \otimes 2^N$ matrix acting on a Hilbert space $\mathbb{C}^{2^N}$, where
\[
\sigma_n^x = 1 \otimes \cdots \otimes 1 \otimes \sigma^x \otimes 1 \otimes \cdots \otimes 1 \quad \text{site } n \quad (4.113)
\]
denotes the Pauli matrix $\sigma^x$ acting at site $n$, assuming periodic boundary conditions modulo $\sigma_N^x \equiv \sigma_0^x$. The Ising quantum chain exhibits a phase transition at the critical value $\lambda_c = 1$.

**Remark:** On the level of the quantum chain, it is customary to replace the letter $\lambda$ by the letter $h$ and to denote it as the *transverse field*. This *transverse field* must not be confused with the *external field* $h$ or $\eta = \beta h / \tau$. The *external field* is a homogeneous magnetic field acting on all Ising spins of the classical model in 2D, while the *transverse field* is not an external field but rather the coupling constant in *transverse* direction along the 1D quantum chain. Unfortunately, both quantities are often denoted by $h$ in the literature.

## 4.3. Exact solution of the quantum $XY$ chain

The Ising quantum chain is a special case of a broader class of spin models, called the *quantum XY chain*, which can be diagonalized exactly. In this Section we outline the sequence of steps that have to be carried out in order to diagonalize the Hamiltonian.\(^4\)

The quantum XY chain with $N$ sites and periodic boundary conditions is defined by

\(^4\)The original version of this Section was written by Pascal Fries.
the Hamiltonian
\[ H = -\sum_{n=0}^{N-1} \left[ \frac{1 + \gamma}{2} \sigma^x_n \sigma^x_{n+1} + \frac{1 - \gamma}{2} \sigma^y_n \sigma^y_{n+1} + h \sigma^z_n \right], \]  
(4.114)

where the parameters \( \gamma \) controls the anisotropy while \( h \) is the strength of the transverse field (this is not the magnetic field, see remarks on the previous page). As before, the Hamiltonian \( H \) acts on the Hilbert space \( (\mathbb{C}^2)^{\otimes N} \cong \mathbb{C}^{2N} \). As can be seen, the special case of the Ising model is included by setting \( \gamma = 1 \).

Following the seminal works by Onsager [?] (1942) and Lieb, Schultz and Mattis [?] (1961), the Ising quantum chain Hamiltonian can be shown to describe a theory of discrete free fermions. This means that an appropriate sequence of transformations allows us to diagonalize the Hamiltonian, writing it in the form
\[ H = E_0 + \sum_k \omega_k \eta_k^+ \eta_k \]  
(4.115)

with certain excitation energies \( \omega_k \) and anticommuting fermionic operators \( \eta_k \) and \( \eta_k^+ \) so that \( \eta_k^+ \eta_k \) can be interpreted as a number operator with eigenvalues \( \pm 1 \). In the following we describe this sequence of transformations in detail.

**Jordan-Wigner transformation**

The Hamiltonian, as given in Eq. (4.114), is translationally invariant along the chain, hence it seems reasonable to look at it in Fourier space. The Fourier transforms of the operators \( \sigma^x/y/z \), however, seem to follow a rather complicated algebra. For this reason we first rewrite Eq. (4.114) in terms of another set of operators, effectively mapping the problem to free fermionic fields on a lattice, whose algebra is much easier to understand. To this end, we introduce the spin flip operators
\[ \sigma_n^\pm = \frac{1}{2} (\sigma_n^x \pm i \sigma_n^y) \]  
(4.116)

with
\[ \sigma_n^x = \sigma_n^+ + \sigma_n^-, \quad \sigma_n^y = -i(\sigma_n^+ - \sigma_n^-), \quad \sigma_n^z = \sigma_n^+ \sigma_n^- - \sigma_n^- \sigma_n^+. \]  
(4.117)

In terms of the spin flip operators, the Hamiltonian takes the form
\[ H = -\sum_{n=0}^{N-1} \left[ \gamma (\sigma_n^+ \sigma_{n+1}^+ + \sigma_n^- \sigma_{n+1}^-) + \sigma_n^+ \sigma_{n+1}^- + \sigma_n^- \sigma_{n+1}^+ + h (\sigma_n^+ \sigma_n^- - \sigma_n^- \sigma_n^+) \right] \]  
(4.118)
or somewhat shorter

\[ H = -\sum_{n=0}^{N-1} \left( \gamma \sigma^+_n \sigma^+_n + \sigma^+_n \sigma^-_{n+1} + \text{h.c.} \right) + h \left( \sigma^+_n \sigma^-_n - \sigma^-_n \sigma^+_n \right), \]  

(4.119)

where ‘+ h.c.’ stands for “plus Hermitian conjugate” or “plus the same expression daggered”.

The operators \( \sigma^+ \) and \( \sigma^- \) are conjugate to each other and act like creation and annihilation operators. This means that on the same site they behave as fermions with the anticommutation relations

\[ \{ \sigma^+_n, \sigma^-_m \} = 0, \quad \{ \sigma^+_n, \sigma^-_n \} = 1. \]  

(4.120)

On different sites, however, they rather behave like bosons with commutation relations

\[ [\sigma^+_n, \sigma^+_m] = [\sigma^-_n, \sigma^-_m] = [\sigma^-_n, \sigma^+_m] = 0 \quad \forall \ n \neq m. \]  

(4.121)

This means that the operators \( \sigma^\pm_m \) behave partly as fermions and partly as bosons. In order to work with truly fermionic fields, we perform a so-called Jordan-Wigner transformation by defining new operators

\[ \psi_n = \prod_{i=0}^{n-1} (\sigma^-_i), \quad \psi^+_n = \prod_{i=0}^{n-1} (\sigma^+_i) \]  

(4.122)

Because of \( \{ \sigma^\pm, \sigma^\pm \} = 0 \) these operators exhibit full fermionic commutation relations:

\[ \{ \psi_n, \psi_m \} = 0, \quad \{ \psi^+_n, \psi_m \} = \delta_{nm}. \]  

(4.123)

**Remark**: It is quite easy to understand how this works. To each operator \( \sigma^\pm \) acting on site \( j \) we prepend a product of \( \sigma^- \) to the left of the actual position, as illustrated in Fig. Fig. 4.7. If the indices are equal, these “\( \sigma^- \)-strings” in the brackets just cancel due to \( (\sigma^-)^2 = 1 \). However, if the indices are different, then one of the two operators has to be commuted with one of the \( \sigma^- \)-factors. This produces a minus sign which is needed for the anticommutation relations of operators acting on different sites.

Now we want to rewrite the Hamiltonian [4.119] in terms of these new operators. To this end we note that the \( \sigma^- \)-strings cancel between products of operators acting on the same site and on nearest neighbors, i.e.

\[ \sigma^+_n \sigma^+_n = \psi^+_n \psi^+_n, \quad \sigma^+_n \sigma^-_{n+1} = \psi^+_n \psi_{n+1}, \quad \sigma^+_n \sigma^-_n = \psi^+_n \psi_n \quad (n = 0, 1, \ldots, N - 2) \]

However, for products of operators acting on the last and the first site of the chain, which arise due to periodic boundary conditions, one obtains a \( \sigma^- \)-string extending over the entire chain, e.g.

\[ \psi^+_m \psi^+_0 = (-\sigma^-_0) \cdots (-\sigma^-_{N-2}) \sigma^-_{N-1} \sigma^+_0 = -P \sigma^-_{N-1} \sigma^+_0 \]  

(4.124)

where \( P = \prod_{n=0}^{N-1} \sigma^-_n \) is the Jordan-Wigner string with the properties \( P^2 = P = P^* \).
Therefore we have
\begin{align*}
\sigma_{N-1}^+ \sigma_0^- &= -P \psi_{N-1}^+ \psi_0^+, \\
\sigma_{N-1}^- \sigma_0^+ &= +P \psi_{N-1}^- \psi_0^+,
\end{align*}
so that the Hamiltonian (4.119) can be written as
\begin{align*}
H &= - \sum_{n=0}^{N-2} \left( \gamma \psi_n^+ \psi_{n+1}^+ + \psi_n^+ \psi_n + h.c. \right) - h \sum_{n=0}^{N-1} \left( \psi_n^+ \psi_n - \psi_n \psi_n^+ \right) \\
&\quad + P \left( \gamma \psi_{N-1}^+ \psi_0^+ + \psi_{N-1}^+ \psi_0 + h.c. \right) \\
&\quad + P \left( \gamma \psi_{N-1}^+ \psi_0^+ + \psi_{N-1}^+ \psi_0 + h.c. \right) \\
&\quad + P \left( \gamma \psi_{N-1}^+ \psi_0^+ + \psi_{N-1}^+ \psi_0 + h.c. \right)
\end{align*}
(4.126)
At first glance, the appearance of the operator $P$ seems to complicate the solution. On the other hand, $P$ is by construction diagonal in the spin configuration basis with
\begin{align*}
P |\uparrow\uparrow\downarrow\ldots\rangle &= \pm |\uparrow\uparrow\downarrow\ldots\rangle.
\end{align*}
(4.127)
It is easy to see that the eigenvalue is $+1$ if the number of down spins is even and $-1$ otherwise. In other words, $P$ measures the parity of the number of down spins. Since all terms in the original Hamiltonian (4.119) are quadratic in $\sigma_j^\pm$, they can only flip in pairs, meaning that the parity is conserved under time evolution:
\begin{align*}
[H, P] &= 0.
\end{align*}
(4.128)
Therefore, $H$ and $P$ can be diagonalized simultaneously. This allows us to define two projectors
\begin{align*}
\Pi_\pm &= \frac{1}{2} (1 \pm P)
\end{align*}
(4.129)
on the $\pm 1$-eigenspaces of $P$ and to express the Hamiltonian as
\begin{align*}
H &= \Pi_+ H_+ \Pi_+ + \Pi_- H_- \Pi_- \\
&= H_+ \Pi_+ + H_- \Pi_-
\end{align*}
(4.130)
where $H_\pm$ is the Hamiltonian in the respective sector, with the operator $P$ can be replaced by its eigenvalue:
\begin{align*}
H_\pm &= - \sum_{n=0}^{N-2} \left( \gamma \psi_n^+ \psi_{n+1}^+ + \psi_n^+ \psi_n + h.c. \right) - h \sum_{n=0}^{N-1} \left( \psi_n^+ \psi_n - \psi_n \psi_n^+ \right) \\
&\quad \pm \left( \gamma \psi_{N-1}^+ \psi_0^+ + \psi_{N-1}^+ \psi_0 + h.c. \right). \\
&\quad \pm \left( \gamma \psi_{N-1}^+ \psi_0^+ + \psi_{N-1}^+ \psi_0 + h.c. \right).
\end{align*}
(4.131)
Note that the two operators $H_\pm$ are identical up to their boundary condition: While $H_-$ has standard periodic boundary conditions, $H_+$ has some kind of “twisted” boundary conditions. In the following we have to solve these two cases individually.

Now we have expressed $H$ in terms of quadratic expressions in the fields. This allows us to write the Hamiltonian in the respective sector as a quadratic form
\begin{align*}
H_\pm &= - \sum_{n,m=0}^{N-1} \left[ A_{nm}^\pm \left( \psi_n^+ \psi_m - \psi_n \psi_m^+ \right) + B_{nm}^\pm \left( \psi_n^+ \psi_m - \psi_n \psi_m^+ \right) \right],
\end{align*}
(4.132)
where $A^\pm$ and $B^\pm$ are matrices of the form
\[
A^\pm = \frac{1}{2} \begin{bmatrix} 2h & 1 & \ldots & \pm 1 \\ 1 & 2h & \ldots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ \pm 1 & 1 & \ldots & 2h \end{bmatrix}, \quad B^\pm = \frac{\gamma}{2} \begin{bmatrix} 0 & 1 & \ldots & \pm 1 \\ -1 & 0 & 1 & \ldots \\ \vdots & \vdots & \ddots & \vdots \\ \pm 1 & -1 & 0 \end{bmatrix}. \tag{4.133}
\]

Defining the field vector $[\psi \quad \psi^\dagger]$ by
\[
[\psi \quad \psi^\dagger]^\dagger = [\psi^\dagger \quad \psi] = [\psi_0^\dagger, \ldots, \psi_{N-1}^\dagger, \psi_0, \ldots, \psi_{N-1}] \tag{4.134}
\]
the Hamiltonian (4.132) can be written in the compact form
\[
H_\pm = -[\psi \quad \psi^\dagger]^\dagger \begin{bmatrix} A^\pm & B^\pm \\ -B^\dagger & -A^\dagger \end{bmatrix} [\psi \quad \psi^\dagger]. \tag{4.135}
\]

**Relation between discrete Fourier transformations and $z$-transformations**

For a given set of $N$ complex numbers $c_0, \ldots, c_{N-1}$ the *discrete Fourier transform* (DFT) and its inverse are usually defined by
\[
\tilde{c}_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} e^{2\pi i nk/N} c_n, \quad c_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{-2\pi i kn/N} \tilde{c}_k \tag{4.136}
\]
where $k = 0, \ldots, N-1$ enumerates the Fourier modes. The DFT is a unitary linear map from position to momentum space.

On a lattice it is convenient to relate the DFT with a *$z$-transformation*. This type of transformation associates with each data set $c_0, c_1, \ldots, c_{N-1}$ a function $c(z)$ by
\[
c(z) := \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} z^n c_n. \tag{4.137}
\]

Obviously, the $z$-transformation is related to the DFT by $\tilde{c}_n = c(e^{2\pi i n/N})$, where the arguments $e^{2\pi i n/N}$ are the $N$th root of unity on the unit circle in the complex plane. This relation can be formalized conveniently by introducing the $N \times N$ translation operator
\[
Z = \begin{bmatrix} 0 & 1 & & & \frac{1}{2} \\ 1 & 0 & & & \cdot \\ & \ddots & \ddots & \ddots \\ & & \cdot & \cdot & \cdot \\ & & & 1 & 0 \end{bmatrix} \tag{4.138}
\]
which simply shifts all components of a vector cyclically. It is easy to see that this matrix
is unitary and returns to the identity after \( N \) iterations, i.e.
\[
ZZ^\dagger = Z^\dagger Z = 1, \quad Z^N = 1.
\] (4.139)

Consequently, the eigenvalues \( z \) of the eigenvalue problem
\[
Z|z\rangle = z|z\rangle
\] (4.140)
are root of unity (\( z^N = 1 \)). More specifically, the solution of the eigenvalue problem is given in term of plane waves
\[
z = e^{2\pi ik/N} \quad \text{and} \quad \langle n|z \rangle = \frac{1}{\sqrt{N}} z^{-n},
\] (4.141)
where \( k = 0, \ldots, N - 1 \). Since \( Z \) is unitary the eigensystem is orthonormal
\[
\langle z|z' \rangle = \frac{1}{N} \sum_{n=0}^{N-1} (z^*)^n = \frac{1}{N} \sum_{n=0}^{N-1} e^{2\pi i (k-k')n/N} = \delta_{k,k'} =: \delta_{z,z'}
\] (4.142)
so that
\[
1 = \sum_{z^N=1} |z\rangle\langle z| \quad \text{and} \quad Z = \sum_{z^N=1} z|z\rangle\langle z|.
\] (4.143)
Moreover, the \( z \)-transformation (4.137) can be expressed as \( c(z) = \langle z|c \rangle \).

**Taking twisted boundary conditions into account**

In the case of the Ising model, the Fourier transformation outlined above works only for \( H_- \), where we have periodic boundary conditions. For \( H_+ \), one needs a similar transformation for twisted boundary conditions. Fortunately, this can be accounted for by introducing two different translation operators
\[
Z_\pm = \begin{bmatrix}
0 & \mp1 \\
1 & 0 \\
\vdots & \vdots \\
1 & 0
\end{bmatrix}
\] (4.144)
which is again unitary and fulfills the condition \( Z_\pm = \mp1 \). The eigenvalues are given by
\[
z_\pm = e^{\pi ik_/N} \quad \text{where} \quad \begin{cases} k_+ = 1, 3, 5, \ldots, 2N - 1 \\ k_- = 0, 2, 4, \ldots, 2N - 2. \end{cases}
\] (4.145)
and the system of eigenvectors is orthonormal in both cases. For this reason it turns out that both \( H_+ \) and \( H_- \) can be diagonalized with the same \( z \)-formalism, the only difference being that the sums are carried out over \( z^n = -1 \) or \( z^n = +1 \), respectively.
4.3 Exact solution of the quantum XY chain

**Fourier transformation**

In order to Fourier-transform the Hamiltonian, we first note that the transformed fields

\[
\psi(z) \equiv \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} z^n \psi_n \quad \text{and} \quad \psi^\dagger(z) \equiv [\psi(z)]^\dagger = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} z^{-n} \psi_n^\dagger
\]  

(4.146)

obey the same anticommutation relations as before, i.e.

\[
\{\psi(z), \psi(z')\} = \{\psi^\dagger(z), \psi^\dagger(z')\} = 0,
\]

\[
\{\psi(z), \psi^\dagger(z')\} = \frac{1}{N} \sum_{n,m=0}^{N-1} z^n (z')^{-m} \{\psi_n, \psi_m^\dagger\} = \frac{1}{N} \sum_{n=0}^{N-1} (z/z')^n = \delta_{zz'},
\]

(4.147)

where \(\delta_{zz'}\) equals 1 if \(z\) and \(z'\) correspond to the same \(k\) and zero otherwise.

Secondly we note that the matrices \(A^\pm\) and \(B^\pm\) in Eq. (4.133) can be expressed as

\[
A^\pm = h \hat{1} + \frac{1}{2} (Z_\pm + Z_\pm^\dagger), \quad B^\pm = -\frac{\gamma}{2} (Z_\pm - Z_\pm^\dagger),
\]

(4.148)

meaning that they are diagonal in the \(z\)-basis

\[
\langle z|A|z'\rangle = \delta_{zz'} (h + \text{Re}(z)), \quad \langle z|B|z'\rangle = \delta_{zz'} (-i\gamma \text{Im}(z)),
\]

(4.149)

where we dropped the \(\pm\) subscript. This allows us to write down the Hamiltonian in Fourier space by inserting identities of the form \(\hat{1} = \sum_{n=1}^{N} |z\rangle \langle z|\):

\[
H_\pm = -\left[ \begin{array}{cc} A^\pm & B^\pm \\ -B^\pm & -A^\pm \end{array} \right] \left[ \begin{array}{c} \psi \\ \psi^\dagger \end{array} \right] = -\sum_{z^n = 1} z^n \sum_{z^m = 1} \left[ \begin{array}{cc} \psi^\dagger \langle z| & |z\rangle \langle z^*| \\ & \end{array} \right] \left[ \begin{array}{c} A^\pm & B^\pm \\ -B^\pm & -A^\pm \end{array} \right] \left[ \begin{array}{c} |z'\rangle \langle z'| \\ |z'^*\rangle \langle z^{*'}| \end{array} \right] \left[ \begin{array}{c} \psi^\dagger \langle z^*| & |z^*\rangle \langle z'| \\ & \end{array} \right] \right]
\]

\[
= -\sum_{z^n = 1} z^n \left[ \begin{array}{cc} \psi(z) & \langle h + \text{Re}(z) \rangle \\ \langle i\gamma \text{Im}(z) \rangle & \langle -i\gamma \text{Im}(z) \rangle \end{array} \right] \left[ \begin{array}{c} |z\rangle \\ |z^*\rangle \end{array} \right] = -\sum_{k=1}^{N} \left[ \begin{array}{cc} \psi_k & \langle h + \text{Re}(z) \rangle \\ \langle i\gamma \text{Im}(z) \rangle & \langle -i\gamma \text{Im}(z) \rangle \end{array} \right] \left[ \begin{array}{c} |\psi_k\rangle \\ |\psi_k^*\rangle \end{array} \right],
\]

(4.150)

**Remark:** This result can easily be translated back to the ‘traditional’ form of a Fourier transformation

\[
H_\pm = \sum_k \left[ \begin{array}{cc} h + \cos \frac{2\pi k}{N} & \langle i\gamma \sin \frac{2\pi k}{N} \rangle \\ \langle -i\gamma \sin \frac{2\pi k}{N} \rangle & \langle h + \text{Re}(z) \rangle \end{array} \right] \left[ \begin{array}{c} \psi_k \langle h + \text{Re}(z) \rangle \\ \langle i\gamma \text{Im}(z) \rangle \end{array} \right],
\]

(4.151)

where the wave number index \(k\) runs over the ranges defined in (4.145).

**Bogoljubov transformation**

The Fourier transformation carried out above mixes the components of the fields \(\psi_n\) and similarly the components of the fields \(\psi_n^\dagger\), preserving the commutation relations. However, for all terms with \(z \neq z^*\), the resulting quadratic form is not yet diagonal. To obtain a fully diagonal form, another transformation is necessary, namely, a so-called Bogoljubov transformation.

Unlike the Fourier transformation, which mixes components of \(\psi\) and components of \(\psi^\dagger\) among themselves, the Bogoljubov transformation mixes the components of \(\psi\) with compo-
nents of $\psi^\dagger$ in such a way that the fermionic anticommutation relations are preserved. More specifically, we define new operators $\eta(z)$ and $\eta^\dagger(z)$ by

$$\eta(z) := u(z)\psi(z) + v(z^*)\psi^\dagger(z^*), \quad \eta^\dagger(z) := u^*(z)\psi^\dagger(z) + v^*(z^*)\psi(z^*)$$  \hspace{1cm} (4.152)

with

$$u(z) = \cos \phi(z), \quad v(z^*) = -i \sin \phi(z^*) = i \sin \phi(z),$$  \hspace{1cm} (4.153)

where the ‘angle’ $\phi(z)$ is given by

$$\tan(2\phi(z)) = \frac{\gamma \text{Im}(z)}{h + \text{Re}(z)}, \quad [h + \text{Re}(z)] \cos(2\phi(z)) \leq 0.$$  \hspace{1cm} (4.154)

With this definition the Bogoljubov transformation can be written in the compact form

$$\begin{bmatrix} \eta(z) \\ \eta^\dagger(z^*) \end{bmatrix} = \begin{bmatrix} \cos \phi(z) & -i \sin \phi(z) \\ -i \sin \phi(z) & \cos \phi(z) \end{bmatrix} \begin{bmatrix} \psi(z) \\ \psi^\dagger(z^*) \end{bmatrix} =: U \begin{bmatrix} \psi(z) \\ \psi^\dagger(z^*) \end{bmatrix}$$  \hspace{1cm} (4.155)

Note that for $z \in \mathbb{R}$ the transformation $U$ reduces to the identity. It is a straight-forward exercise to show that $\eta, \eta^\dagger$ are fermionic operators obeying the anticommutation relations

$$\{\eta(z), \eta(z')\} = \{\eta^\dagger(z), \eta^\dagger(z')\} = 0, \quad \{\eta(z), \eta^\dagger(z')\} = \delta_{z,z'}.$$  \hspace{1cm} (4.156)

The Bogoljubov transformation specified above was chosen in such a way that it diagonalizes the Hamiltonian \[4.150\].

$$H_\pm = -\sum_{z^N=\pm 1} \begin{bmatrix} \psi(z) \\ \psi^\dagger(z^*) \end{bmatrix}^\dagger U^{-1} U \begin{bmatrix} h + \text{Re}(z) & -i \gamma \text{Im}(z) \\ i \gamma \text{Im}(z) & -(h + \text{Re}(z)) \end{bmatrix} U^{-1} \begin{bmatrix} \psi(z) \\ \psi^\dagger(z^*) \end{bmatrix}$$

\hspace{1cm} (4.157)

$$= +\sum_{z^N=\pm 1} \begin{bmatrix} \eta(z) \\ \eta^\dagger(z^*) \end{bmatrix}^\dagger \begin{bmatrix} \Lambda(z) \\ -\Lambda(z) \end{bmatrix} \begin{bmatrix} \eta(z) \\ \eta^\dagger(z^*) \end{bmatrix},$$

where

$$\Lambda(z) = \sqrt{[h + \text{Re}(z)]^2 + [\gamma \text{Im}(z)]^2}.$$  \hspace{1cm} (4.158)

This may be rewritten as

$$H_\pm = \sum_{z^N=\pm 1} \Lambda(z) \left[ \eta^\dagger(z)\eta(z) - \eta(z^*)\eta^\dagger(z^*) \right]$$

\hspace{1cm} (4.159)

$$= E^{(0)}_\pm + 2 \sum_{z^N=\pm 1} \Lambda(z) \eta^\dagger(z)\eta(z)$$

where

$$E^{(0)}_\pm = -\sum_{z^N=\pm 1} \Lambda(z)$$  \hspace{1cm} (4.160)

is the ground state energy in the respective sector.

### 4.4. Conformal invariance

to be written
4.5 Continuum limit of the Ising model

So far we have introduced and quantitatively understood the Ising model on a $d$-dimensional lattice. Is it possible to devise a theory for the Ising model in terms of continuous degrees of freedom?

By “continuous” we mean that any discrete element of the Ising model is replaced by a continuous option. More specifically, we would like to replace the lattice positions $i$ by a continuous vector $r \in \mathbb{R}^d$. Moreover, we would like to replace the discrete spins $s_i = \pm 1$ by a continuous field $\phi(r) \in \mathbb{R}$.

Before introducing continuous coordinates let us consider the discretization of the spin itself. Replacing the binary spin variable $s_i = \pm 1$ by a continuous degree of freedom $\phi$ we have to make sure that the continuous variable $\phi$ favors two $Z_2$-symmetric values. This is most easily achieved if we introduce a symmetric double-well potential

$$ V(\phi) = \phi^2 + \lambda(\phi^2 - 1)^2 = \lambda \phi^4 + (1 - 2\lambda)\phi^2 + 1 $$

(4.161)

Next, let us write down the “action” for this continuous field on the lattice. To this end, let us denote by $r_i$ the position of the lattice sites. Then, the energy functional reads

$$ E[\phi] = \sum_i \left[ -2\kappa \sum_u \phi(r_i + u)\phi(r_i) + \phi^2(r_i) + \lambda(\phi^2(r_i) - 1)^2 \right], $$

(4.162)

where the second sum runs over all displacements pointing to the nearest neighbors of $r_i$.

The continuum limit of this expression is straightforward. We know that on a square lattice the Laplacian has the approximate discrete representation

$$ \Delta \phi(r) \approx \left( \sum_u \phi(r + u) \right) - 2d \phi(r), $$

(4.163)

where $d$ is the dimension of the lattice. This allows us to express the energy functional as an integral

$$ E[\phi] = \int d^d r \mathcal{L}[\phi](r) $$

(4.164)

with the Lagrange density

$$ \mathcal{L}[\phi] = -2\kappa (\nabla \phi)(\nabla \phi) + (1 - 4\kappa d)\phi^2 + \lambda(\phi^2 - 1)^2 $$

(4.165)

Assuming that the field vanishes asymptotically at infinity, we can partially integrate this expression, obtaining the Lagrange density

$$ \mathcal{L}[\phi] = +2\kappa (\nabla \phi)(\nabla \phi) + (1 - 4\kappa d)\phi^2 + \lambda(\phi^2 - 1)^2 $$

(4.166)

Usually one introduces new constants, rewriting the Lagrangian as

$$ \mathcal{L}[\phi] = \frac{1}{2} (\nabla \phi)(\nabla \phi) - \frac{m}{2} \phi^2 + \frac{g}{4!} \phi^4. $$

(4.167)

The partition function is then given as a functional integral over the Boltzmann factor of the energy functional:

$$ Z = \int D\phi \exp \left[ -\beta \left( \frac{1}{2} (\nabla \phi)(\nabla \phi) - \frac{m}{2} \phi^2 + \frac{g}{4!} \phi^4 \right) \right] $$

(4.168)
Here, the first integral has to be read as the “sum over all possible configurations of the continuous field $\phi$.” This partition sum is known as $\phi^4$ field theory, the simplest nontrivial field theory with so-called loop corrections.
5. Nonequilibrium phase transitions

In this Chapter, we are primarily interested in phase transitions in systems far from equilibrium. To this end we consider stochastic processes that violate detailed balance so strongly that concepts of equilibrium statistical physics can no longer be applied, even in an approximate sense. We are interested in the question whether stochastic processes far from equilibrium can exhibit new phenomena that cannot be observed under equilibrium conditions. This applies in particular to phase transitions far from equilibrium.

5.1. Directed percolation

The probably most important class of non-equilibrium processes, which display a non-trivial phase transition from a fluctuating phase into an absorbing state, is Directed Percolation (DP). As will be discussed below, the DP class comprises a large variety of models that share certain basic properties.

5.1.1. Directed bond percolation on a lattice

To start with let us first consider a specific model called directed bond percolation which is often used as a simple model for water percolating through a porous medium. The model is defined on a tilted square lattice whose sites represent the pores of the medium. The pores are connected by small channels (bonds) which are open with probability $p$ and closed otherwise. As shown in Fig. 5.1 water injected into one of the pores will percolate along open channels, giving rise to a percolation cluster of wetted pores whose average size will depend on $p$.

There are two fundamentally different versions of percolation models. In isotropic percolation the flow is undirected, i.e., the spreading agent (water) can flow in any direction through open

Figure 5.1.: Isotropic versus directed bond percolation. The figure shows two identical realizations of open and closed bonds on a finite part of a tilted square lattice. A spreading agent (red) is injected at the central site (blue circle). In the case of isotropic percolation (left) the agent percolates through open bonds in any direction. Contrarily, in the case of directed percolation, the agent is restricted to percolate along a preferred direction, as indicated by the arrow.
Figure 5.2.: Directed bond percolation. The process shown here starts with a single active seed at the origin. It then evolves through a sequence of configurations along horizontal lines (called states) which can be labeled by a time-like index \( t \). An important quantity to study would be the total number of active sites \( N(t) \) at time \( t \).

bonds (left panel of Fig. 5.1). A comprehensive introduction to isotropic percolation is given in the textbook by Stauffer [7]. In the present lecture, however, we are primarily interested in the case of directed percolation. Here the clusters are directed, i.e., the water is restricted to flow along a preferred direction in space, namely downwards, as indicated by the arrow in Fig. 5.1. In the context of porous media this preferred direction may be interpreted as a gravitational driving force. Using the language of electronic circuits DP may be realized as a random diode network (cf. Sect. 5.2.4 on page 118).

The strict order of cause and effect in DP allows one to interpret the preferred direction as a temporal coordinate. For example, in directed bond percolation, we may enumerate horizontal rows of sites by an integer time index \( t \) (see Fig. 5.2). Instead of viewing it as a static model of directed connectivity, we shall now on interpret DP as a dynamical process which evolves in time. Denoting wetted sites as active and dry sites as inactive the process starts with a certain initial configuration of active sites that can be chosen freely. For example, in Fig. 5.2 the process starts with a single active seed at the origin. As \( t \) increases the process evolves stochastically through a sequence of configurations of active sites at time \( t \). An important quantity, which characterizes these intermediate states, is the total number of active sites \( N(t) \), as illustrated in Fig. 5.2.

Regarding directed percolation as a reaction-diffusion process the local transition rules may be interpreted as follows. Each active site represents a particle \( A \). If the two subsequent bonds are both closed, the particle will have disappeared at the next time step by a death process \( A \rightarrow \emptyset \) (see Fig. 5.3a). If only one of the bonds is open, the particle diffuses stochastically to the left or to the right, as shown in Fig. 5.3b-5.3c. Finally, if the two bonds are open the particle creates an offspring \( A \rightarrow 2A \) (Fig. 5.3d). However, it is important to note that each site in directed bond percolation can be either active or inactive. In the particle language this means that each site can be occupied by at most one particle. Consequently, if two particles happen to reach the same site, they merge irreversibly forming a single one by coalescence \( 2A \rightarrow A \), as illustrated in Fig. 5.3e. Summarizing these reactions, directed bond percolation can be interpreted as a reaction-diffusion process which effectively follows the reaction scheme

\[
\begin{align*}
A \rightarrow \emptyset & \quad \text{death process} \\
A \rightarrow 2A & \quad \text{offspring production} \\
2A \rightarrow A & \quad \text{coagulation}
\end{align*}
\]

combined with single-particle diffusion.

The dynamical interpretation in terms of particles is of course the natural language for any algorithmic implementation of DP on a computer. As the configuration at time \( t \) depends exclusively on the previous configuration at time \( t-1 \) it is not necessary to store the entire cluster.
in the memory, instead it suffices to keep track of the actual configuration of active sites at a
given time and to update this configuration in parallel sweeps according to certain probabilistic
rules. In the case of directed bond percolation, one obtains a stochastic cellular automaton with
certain update rules in which each active site of the previous configuration activates its nearest
neighbors of the actual configuration with probability $p$. In fact, as shown in Fig. 5.4 a simple
non-optimized C-code for directed bond percolation takes less than a page, and the core of the
update rules takes only a few lines.

## 5.1.2. Absorbing states and critical behavior

As only active sites at time $t$ can activate sites at time $t + 1$, the configuration without active sites
plays a special role. Obviously, such a state can be reached by the dynamics but it cannot be
left. In the literature such states are referred to as absorbing. Absorbing states can be thought of
as a trap: Once the system reaches the absorbing state it becomes trapped and will stay there
forever. As we will see below, a key feature of directed percolation is the presence of a single
absorbing state, usually represented by the empty lattice.

The mere existence of an absorbing state demonstrates that DP is a dynamical process far
from thermal equilibrium. As explained in the Introduction, equilibrium statistical mechanics
deals with stationary equilibrium ensembles that can be generated by a dynamics obeying de-
tailed balance, meaning that probability currents between pairs of sites cancel each other. As
the absorbing state can only be reached but not be left, there is always a non-zero current of
probability into the absorbing state that violates detailed balance. Consequently the temporal
evolution before reaching the absorbing state cannot be described in terms of thermodynamic
ensembles, proving that DP is a non-equilibrium process.

The enormous theoretical interest in DP – more than 800 articles refer to this class of models –
is related to the fact that DP displays a non-equilibrium phase transition from a fluctuating phase
into the absorbing state controlled by the percolation probability $p$. The existence of such a tran-
sition is quite plausible since offspring production and particle death compete with each other.
As will be discussed below, phase transitions into absorbing states can be characterized by cer-
tain universal properties which are independent of specific details of the microscopic dynamics.
In fact, the term ‘directed percolation’ does not stand for a particular model, rather it denotes a
whole universality class of models which display the same type of critical behavior at the phase
transition. The situation is similar as in equilibrium statistical mechanics, where for example the
Ising universality class comprises a large variety of different models. In fact, DP is probably as
fundamental in non-equilibrium statistical physics as the Ising model in equilibrium statistical
mechanics.

In DP the phase transition takes place at a certain well-defined critical percolation probability
$p_c$. As illustrated in Fig. 5.5 the behavior on both sides of $p_c$ is very different. In the subcrit-
critical regime $p < p_c$ any cluster generated from a single seed has a finite life time and thus a
finite mass. Contrarily, in the supercritical regime $p > p_c$ there is a finite probability that the
generated cluster extends to infinity, spreading roughly within a cone whose opening angle de-
pends on $p - p_c$. Finally, at criticality finite clusters of all sizes are generated. These clusters are

![Figure 5.3: Interpretation of the dynamical rules of directed bond percolation as a reaction-diffusion process: a) death process, b)-c) diffusion, d) offspring production, and e) coagulation.](image)
Nonequilibrium phase transitions

---

Figure 5.4. Simple non-optimized program in C for directed bond percolation. The numerical results are shown below in Fig. 6.
sparse and remind of self-similar fractal objects. As we will see, a hallmark of such a scale-free behavior is a power-law behavior of various quantities.

The precise value of the percolation threshold $p_c$ is non-universal (i.e., it depends on the specific model) and can only be determined numerically. For example, in the case of directed bond percolation in 1+1 dimensions, the best estimate is $p_c = 0.644700185(5)$. Unlike isotropic bond percolation in two dimensions, where the critical value is exactly given by $p_{ciso} = 1/2$, an analytical expression for the critical threshold of DP in finite dimensions is not yet known. This seems to be related to the fact that DP is a non-integrable process. It is in fact amazing that so far, in spite of its simplicity and the enormous effort by many scientists, DP resisted all attempts to be solved exactly, even in 1+1 dimensions.

In order to describe the phase transition quantitatively an appropriate order parameter is needed. For simulations starting from a single active seed a suitable order parameter is the average number of particles $\langle N(t) \rangle$ at time $t$, where $\langle \ldots \rangle$ denotes the average over many independent realizations of randomness (called runs in the numerical jargon). For example, the program shown in Fig. 5.4 averages this quantity over 10,000 runs and stores the result in a file that can be viewed by a graphical tool such as xmgrace. As shown in Fig. 5.6, there are three different cases:

- For $p < p_c$ the average number of active sites first increases and then decreases rapidly. As demonstrated in the inset, this decrease is in fact exponential. Obviously, the typical crossover time, where exponential decay starts, depends on the distance from criticality $p_c - p$.

- At criticality the average number of active sites increases according to a power law $\langle N(t) \rangle \sim t^\theta$. A standard regression of the data gives the exponent $\theta \simeq 0.302$, shown in the figure as a thin dashed line. As can be seen, there are deviations for low $t$, i.e., the process approaches a power-law only asymptotically.

- In the supercritical regime $p > p_c$ the slow increase of $\langle N(t) \rangle$ crosses over to a fast linear increase with time. Again the crossover time depends on the distance from criticality $p - p_c$.

The properties of $\langle N(t) \rangle$ above and below criticality can be used to determine the critical threshold numerically. This iterative procedure works as follows: Starting with a moderate simulation time it is easy to specify a lower and an upper bound for $p_c$ by hand, e.g., $0.64 < p_c < 0.65$ in
the case of directed bond percolation. This interval is then divided into two equal parts and the process is simulated in between, e.g., at $p = 0.645$. In order to find out whether this value is sub- or supercritical one has to check deviations for large $t$ from a straight line in a double logarithmic plot. If the curve veers down (up) the procedure is iterated using the upper (lower) interval. In detecting the sign of curvature the human eye is quite reliable but it is also possible to recognize it automatically. If there is no obvious deviation from a straight line the simulation time and the number of runs has to be increased appropriately.

**Warning:** Determining the numerical error for the estimate of a critical exponent never use the statistical $\chi^2$-error of a standard regression! For example, for the data produced by the minimal program discussed above, a linear regression by *xmgrace* would give the result $\theta = 0.3017(2)$. However, the estimate in the literature $\theta = 0.313686(8)$ lies clearly outside these error margins, meaning that the actual error must be much higher. A reliable estimate of the error can be obtained by comparing the slopes over the last decade of simulations performed at the upper and the lower bound of $p_c$.

In principle the accuracy of this method is limited only by the available CPU time. We note, however, that this standard method assumes a clean asymptotic power law scaling, which for DP is indeed the case. However, in some cases the power law may be superposed by slowly varying deviations, e.g. logarithmic corrections, so that the plotted data at criticality is actually not straight but slightly curved. With the method outlined above one is then tempted to ‘compensate’ this curvature by tuning the control parameter, leading to unknown systematic errors. Recently this happened, for example, in the case of the diffusive pair contact process, as will be described in Sect. 5.3.4.

5.1.3. The Domany-Kinzel cellular automaton

An important model for DP, which includes directed bond percolation as a special case, is the celebrated Domany-Kinzel model [2, 7]. The Domany-Kinzel model is a stochastic cellular automaton defined on a diagonal square lattice which evolves by parallel updates according to certain conditional transition probabilities $P[s_i(t+1)|s_{i-1}(t), s_{i+1}(t)]$, where $s_i(t) \in \{0, 1\}$ denotes the occupancy of site $i$ at time $t$. These probabilities depend on two parameters $p_1, p_2$ and
are defined by

\[ \begin{align*}
P[1|0,0] &= 0, \\
P[1|0,1] &= P[1|1,0] = p_1, \\
P[1|1,1] &= p_2, \\
\end{align*} \]

with \( P[0|\cdot,\cdot] = 1 - P[1|\cdot,\cdot] \). On a computer the Domany-Kinzel model can be implemented as follows. To determine the state \( s_i(t) \) of site \( i \) at time \( t \) we generate for each site a random number \( z \in (0,1) \) from a flat distribution and set

\[ s_i(t+1) = \begin{cases} 
1 & \text{if } s_{i-1}(t) \neq s_{i+1}(t) \text{ and } z_i(t) < p_1, \\
1 & \text{if } s_{i-1}(t) = s_{i+1}(t) = 1 \text{ and } z_i(t) < p_2, \\
0 & \text{otherwise}.
\end{cases} \]  

This means that a site is activated with probability \( p_2 \) if the two nearest neighbors at the previous time step were both active while it is activated with probability \( p_1 \) if only one of them was active. Thus the model depends on two percolation probabilities \( p_1 \) and \( p_2 \), giving rise to the two-dimensional phase diagram shown in Fig. 5.7. The active and the inactive phase are now separated by a line of phase transitions. This line includes several special cases. For example, the previously discussed case of directed bond percolation corresponds to the choice \( p_1 = p \) and \( p_2 = p(2-p) \). Another special case is directed site percolation, corresponding to the choice \( p_1 = p_2 = p \). In this case all bonds are open but sites are permeable with probability \( p \) and blocked otherwise. Finally, the special case \( p_2 = 0 \) is a stochastic generalization of the rule ‘W18’ of Wolfram’s classification scheme of cellular automata [7] (Yes, it’s him, the Mathematica® Wolfram!). Numerical estimates for the corresponding critical parameters are listed in Table 5.1.

There is strong numerical evidence that the critical behavior along the whole phase transition line (except for its upper terminal point) is that of DP, meaning that the transitions always exhibit the same type of long-range correlations. The short-range correlations, however, are non-universal and may change when moving along the phase transition line. They may even change the visual appearance of the clusters, as illustrated in Fig. 5.8 where four typical snapshots of critical clusters are compared. Although the large-scale structure of the clusters in the first three cases is roughly the same, the microscopic texture seems to become bolder as we move up along the phase transition line. As shown in Ref. [7], this visual impression can be traced back to an increase of the mean size of active islands.
Approaching the upper terminal point the mean size of active islands diverges and the cluster becomes compact. For this reason this special point is usually referred to as compact directed percolation. However, this nomenclature may be misleading for the following reasons. The exceptional behavior at this point is due to an additional symmetry between active and inactive sites along the upper border of the phase diagram at $p_2 = 1$. Here the DK model has two symmetric absorbing states, namely, the empty and the fully occupied lattice. For this reason the transition does no longer belong to the universality class of directed percolation, instead it becomes equivalent to the (1+1)-dimensional voter model \cite{7, 8}, or the Glauber-Ising model at zero temperature. Since the dynamic rules are invariant under the replacement $p_1 \leftrightarrow 1 - p_1$, the corresponding transition point is located at $p_1 = 1/2$.

The Domany-Kinzel model can be generalized easily to higher spatial dimensions. For example, Fig. 5.9 shows a possible cluster in 2+1-dimensional directed bond percolation. Generally, in the $d+1$-dimensional Domany-Kinzel model the activation probability of site $i$ at time $t+1$ depends on the number $n_i(t) = \sum_{j \in \langle i \rangle} s_j(t)$ of active nearest neighbors at time $t$, i.e. the conditional probabilities

$\begin{align*}
P[1|0] &= 0, \\
P[1|n] &= p_n, \quad (1 \leq n \leq 2d)
\end{align*}$

are controlled by $2d$ parameters $p_1, \ldots, p_{2d}$. The special case of directed bond percolation corresponds to the choice $p_n = 1 - (1 - p)^n$ while for equal parameters $p_n = p$ one obtains directed site percolation in $d+1$ dimensions.

### 5.1.4. The contact process

Another important model for directed percolation, which is popular in mathematical communities, is the contact process. The contact process was originally introduced by Harris \cite{7} as a model for epidemic spreading (see Sect. 5.4). It is defined on a $d$-dimensional square lattice whose sites can be either active ($s_i(t) = 1$) or inactive ($s_i(t) = 0$). In contrast to the Domany-Kinzel model, which is a stochastic cellular automaton with parallel updates, the contact process evolves by asynchronous updates, i.e., the three elementary processes (offspring produc-
5.1 Directed percolation

![Figure 5.9: Lattice geometry of directed bond percolation in 2+1 dimensions. The red lines represent a possible cluster generated at the origin.](image)

...tion, on-site removal and diffusive moves) occur spontaneously at certain rates. Although the microscopic dynamics differs significantly from the Domany-Kinzel model, the contact process displays the same type of critical behavior at the phase transition. In fact, both models belong to the universality class of DP.

In the ‘epidemic’ language the contact process effectively corresponds to the following reaction-diffusion scheme:

\[
\begin{align*}
A &\rightarrow 2A: \text{ infection of neighbors at rate } \lambda \\
A &\rightarrow \emptyset: \text{ spontaneous self-recovery at rate 1} \\
2A &\rightarrow A: \text{ infected individuals cannot be infected twice} \\
\emptyset A &\leftrightarrow A\emptyset: \text{ nearest-neighbor diffusion}
\end{align*}
\]

On a computer the \(d+1\)-dimensional contact process can be implemented as follows. For each attempted update a site \(i\) is selected at random. Depending on its state \(s_i(t)\) and the number of active neighbors \(n_i(t) = \sum_{j \in \langle i \rangle} s_j(t)\) a new value \(s_i(t+dt) = 0, 1\) is assigned according to certain transition rates \(w[s_i(t) \rightarrow s_i(t+dt), n_i(t)]\). In the standard contact process these rates are defined by

\[
\begin{align*}
w[0 \rightarrow 1, n] &= \lambda n/2d, \\
w[1 \rightarrow 0, n] &= 1.
\end{align*}
\]

Here the parameter \(\lambda\) plays the same role as the percolation probability in directed bond percolation. Its critical value depends on the dimension \(d\). For example, in 1+1 dimensions the best-known estimate is \(\lambda_c \approx 3.29785(8)\) [?].

As demonstrated in Ref. [?] the evolution of the contact process can be described in terms of a master equation whose Liouville operator \(\mathcal{L}\) can be constructed explicitly on a finite lattice. Diagonalizing this operator numerically one obtains a spectrum of relaxational modes with at least one zero mode which represents the absorbing state. In the limit of large lattices the critical threshold \(\lambda_c\) is usually the point from where on the first gap in the spectrum of \(\mathcal{L}\) vanishes.

5.1.5. The critical exponents of directed percolation

In equilibrium statistical physics, continuous phase transitions as the one in the Ising model can be described in terms of a phenomenological scaling theory. For example, the spontaneous magnetization \(M\) in the ordered phase vanishes as \(|M| \sim (T_c - T)^\beta\) as the critical point is approached. Here \(\beta\) is a universal critical exponent, i.e. its value is independent of the specific realization of the model. Similarly, the correlation length \(\xi\) diverges in as \(\xi \sim |T - T_c|^{-\nu}\) for \(T \rightarrow T_c\) with another universal exponent \(\nu\). The critical point itself is characterized by the
absence of a macroscopic length scale so that the system is invariant under suitable scaling transformations (see below).

In directed percolation and other non-equilibrium phase transitions into absorbing states the situation is very similar. However, as non-equilibrium system involves time which is different from space in character, there are now two different correlation lengths, namely, a spatial correlation length $\xi_\perp$ and a temporal correlation length $\xi_\parallel$ with two different associated exponents $\nu_\perp$ and $\nu_\parallel$. Their ratio $z = \nu_\parallel / \nu_\perp$ is called dynamical exponent as it relates spatial and temporal scales at criticality.

What is the analogon of the magnetization in DP? As shown above, in absorbing phase transitions the choice of the order parameter depends on the initial configuration. If homogeneous initial conditions are used, the appropriate order parameter is the density of active sites at time $t$

$$\rho(t) = \lim_{L \to \infty} \frac{1}{L} \sum_i s_i(t) .$$

Here the density is defined as a spatial average in the limit of large system sizes $L \to \infty$. Alternatively, for a finite system with periodic boundary conditions we may express the density as

$$\rho(t) = \langle s_i(t) \rangle ,$$

where $\langle \ldots \rangle$ denotes the ensemble average over many realizations of randomness. Because of translational invariance the index $i$ is arbitrary. Finally, if the process starts with a single seed, possible order parameters are the average mass of the cluster

$$N(t) = \langle \sum_i s_i(t) \rangle$$

and the survival probability

$$P_{\text{sur}}(t) = \langle 1 - \prod_i (1 - s_i(t)) \rangle .$$

These quantities allow us to define the four standard exponents

$$\rho(\infty) \sim (p - p_c)^\beta ,$$

$$P_{\text{sur}}(\infty) \sim (p - p_c)^{\beta'} ,$$

$$\xi_\perp \sim |p - p_c|^{-\nu_\perp} ,$$

$$\xi_\parallel \sim |p - p_c|^{-\nu_\parallel} .$$

The necessity of two different exponents $\beta$ and $\beta'$ can be explained in the framework of a field-theoretic treatment, where these exponents are associated with particle creation and annihilation operators, respectively. In DP, however, a special symmetry, called rapidity reversal symmetry, ensures that $\beta = \beta'$. This symmetry can be proven most easily in the case of directed bond percolation, where the density $\rho(t)$ starting from a fully occupied lattice and the survival probability $P_{\text{sur}}(t)$ for clusters grown from a seed are exactly equal for all $t$. Hence both quantities scale identically and the two corresponding exponents have to be equal. This is the reason why DP is characterized by only three instead of four critical exponents.

Directed percolation is characterized by three exponents $\beta, \nu_\perp, \nu_\parallel$.

The cluster mass $N(t) \sim t^\theta$ scales algebraically as well. The associated exponent, however, is not independent, instead it can be expressed in terms of the so-called generalized hyperscaling
5.1 Directed percolation

relation [?]

\[ \theta = \frac{d v_\perp - \beta - \beta'}{v_\parallel}. \] (5.15)

In order to determine the critical point of a given model by numerical methods, \( N(t) \) turned out to be one of the most sensitive quantities.

5.1.6. Mean field approximation

Homogeneous mean field approximation

In a mean field approximation one usually replaces local interactions by some effective global interactions (the ‘mean field’) in order to get a suitable approximation. There are in fact many possible mean field approximations, differing in their degree of simplification. For DP the simplest mean field scheme would be to suppress any kind of spatial information and to characterize a configuration solely by the density of active sites at a given time, denoted as \( \rho(t) \). This is what chemists do: They consider only the concentration of active sites and write down an equation of motions:

\[ \frac{d}{dt} \rho(t) = \lambda \rho(t) (1 - \rho(t)) - \rho(t). \] (5.16)

Here the first term on the r.h.s. describes infection: it is proportional to the density of individuals carrying the infection times the density of non-yet-infected individuals, multiplied by the infection rate \( \lambda \). The second term describes spontaneous recovery at rate 1.

Obviously, this equation has two stationary solutions, where the l.h.s. vanishes, namely, \( \rho = 0 \), representing the empty lattice (absorbing state), and \( \rho = (\lambda - 1)/\lambda \). The good news is that this simple mean field equation does reproduce a continuous phase transition. It takes place at the critical value \( \lambda_c = 1 \), where the density is found to scale as

\[ \rho \sim (\lambda - 1)^\beta + \mathcal{O}((\lambda - 1)^2), \] (5.17)

hence the critical exponent \( \beta = 1 \). The bad news is that this is quite far from what we observe in 1D, namely, \( \lambda_c \approx 3.298 \) and \( \beta = 0.276 \).

In the literature it is common to drop the \( \lambda \)-dependence in the quadratic term by the replacement \( \rho \rightarrow \rho/\lambda \), which yields the same results. Moreover, it is useful to define the distance from criticality, denoted here as \( a \):

\[ a = \lambda - \lambda_c. \] (5.18)

The resulting mean field equation reads:

\[ \frac{d}{dt} \rho(t) = a \rho - \rho^2 \] (5.19)

with the stationary solutions \( \rho = 0 \) and \( \rho = a \).

Non-homogeneous mean field approximation

The homogeneous mean-field equations can be extended easily to include an explicit dependence on the position. Since the infection effectively performs a random-walk in space, it is natural to add a term for diffusion. The resulting equation reads

\[ \frac{d}{dt} \rho(x, t) = a \rho - \rho^2 + D \nabla^2 \rho, \] (5.20)
where $D$ is the diffusion constant. As will be shown in an exercise, this allows us to extract the critical exponent $\nu_{\perp}$ within the mean field approximation. However, the equation has a shortcoming in so far as perfectly homogeneous state (it density which does not depend on $x$) remains homogeneous for all times. This is certainly not realistic since DP always generates non-homogeneous dates by itself. We will come back to this point when discussing so-called Langevin equation (see Sect. 5.2.1 on page 113).

### 5.1.7. Phenomenological scale invariance

Despite its simplicity an exact solution of DP is not yet known. In the attempt to characterize the phase transition of DP we therefore rely on numerical simulations, approximations and phenomenological approaches. One such phenomenological approach is the postulate of scale invariance. Scale invariance means that

- the emerging physics at the critical point is scale-free, just as in a fractal.
- apart from the imposed cutoffs (the lattice spacing and the system size) no finite length or time scale can be singled out at the critical point.
- the behavior in the vicinity of the critical point is characterized by only one length scale $\xi_{\perp}$ and one length scale $\xi_{\parallel}$.

Starting point of a phenomenological scaling theory for absorbing phase transitions is the assumption that the macroscopic properties of the system close to the critical point are invariant under scale transformations of the form

\[
\begin{align*}
\text{Scaling of distance from criticality:} & \quad a \to b^{-1}a \\
\text{Scaling of distance-like quantities:} & \quad x \to b^{\nu_{\perp}}x \\
\text{Scaling of time-like quantities:} & \quad t \to b^{\nu_{\parallel}}t \\
\text{Scaling of the density of active sites:} & \quad \rho \to b^{-\beta} \rho \\
\text{Scaling of the survival probability:} & \quad P_s \to b^{-\beta'} P_s
\end{align*}
\]

where $b > 0$ is some scale factor and $a \sim p - p_c$ denotes the distance from criticality.

The scale factor $b$ can be thought of as a zoom operation. For $b > 1$ we zoom into the system, meaning that length and time scales are growing and that we are moving closer to words criticality. At the same time, more and more fractal details of the clusters become visible, leading to an effective decrees of the density of active sites. Contrarily, for $b < 1$ we zoom out, seeing decreasing length and time scales and increasing coarse-grained density of active sites. This process of scale changes is at the core of renormalization.

Scaling invariance strongly restricts the form of functions. For example, at the critical point in an infinite system, $\rho$ is a function $f(t)$ depending on time. If we now apply the scale transformation on both sides we get:

\[
\rho = f(t) \quad \Rightarrow \quad b^{-\beta} \rho = f(b^{\nu_{\parallel}} t).
\]

Since this relation is valid for all scale factors $b > 0$, we have the freedom to choose a special one. Therefore let us choose $b$ in such a way that the argument of the function takes on the value 1, i.e., we choose $b = t^{-1/v_1}$. Then we have

\[
t^{\beta/\nu_{\parallel}} \rho = f(1) \quad \Rightarrow \quad \rho \propto t^{-\beta/\nu_{\parallel}}.
\]

This implies that a function depending on only one argument is forced to behave like power law with a well-defined exponent determined by the scale transformation. In this relation, $f(1)$
5.1 Directed percolation

Figure 5.10.: Data collapse of the survival probability measured in a finite DP process with $L$ sites (see text). As can be seen, the collapse is very accurate. The shape of the collapsed curves resembles the scaling function $F$.

plays the role of the proportionality factor. In the literature the power law derived above is usually written in the form

$$\rho(t) \sim t^{-\delta}, \quad \delta = \frac{\beta}{v_\parallel},$$

(5.24)

where the tilde $\sim$ as the meaning of being asymptotically proportional, that is, in the limit of large times.

Similarly, starting from an initial seed, the survival probability $P_s(t)$ can be shown to decay

$$P_s(t) \sim t^{-\delta'}, \quad \delta' = \frac{\beta'}{v_\parallel}$$

(5.25)

with $\delta = \delta'$ in the case of DP.

A similar reasoning holds in situations where more than one parameter involved. For example, let us consider the decay of the particle density at criticality ($\Delta = 0$) in a finite system of size $L$. Here we expect the density $\rho$ to be function of both the actual time $t$ and their given system size $L$, i.e., $\rho = f(t, L)$. Noting that the system size behaves like a length, the application of the scale transformation listed above results into

$$\rho = f(t, L) \Rightarrow b^{-\beta} \rho = f(b^{\parallel}t, b^{\perp}L).$$

(5.26)

Choosing again $b = t^{-1/v_\parallel}$ this turns into

$$t^{\beta/v_\parallel} \rho = f(1, t^{-v_\perp/v_\parallel}L).$$

(5.27)

Rewriting the function appearing on the right hand side by $F(x) = f(1, x^{-v_\perp/v_\parallel})$ we arrive at the usual form that can be found in the literature, namely

$$\rho(t, L) \sim t^{-\delta} F(t/L^z),$$

(5.28)

where $z = v_\parallel/v_\perp$ is the so-called dynamical exponent. So we can conclude that a function like $\rho(t, L)$, which depends on two parameters, is a reduced by scale invariance to a leading power law multiplied by a function of only one parameter like $F(t/L^z)$. More generally, a function depending on $N$ parameters will be reduced to a power law times a function depending on $N - 1$ parameters. This function is referred to as the scaling function of the corresponding scaling law. In this context let us note the following:

- The scaling function on the right hand side (like $F(t/L^z)$ in the preceding example) de-
pends on parameters which by construction are invariant under scale transformations. For this reason these parameters are also called *scale-invariant ratios*. For example, $t/L^z$ is such a scale-invariant ratio.

- It turns out that scaling functions by themselves are universal, too. That is, in all models belonging to the DP universality class we obtain *exactly* the same type of scaling function which can be extracted numerically provided that we invest sufficient computational power. This is highly nontrivial since functions provide much more information than the critical exponents alone.

The universality of scaling functions is exploited in the numerical technique involving *data collapses*. In the example given above, where $\rho \sim t^{-\delta} F(t/L^z)$, we would run several simulations at different system sizes, typically arranged in powers of 2 (e.g. $L=128, 256, 512,...$). Plotting $\rho t^\delta$ versus the scale-invariant argument $t/L^z$, our graph shows the function $F$ in all cases, that is, the graphs for different system sizes have to collapse and what we see is exactly the shape of the scaling function. The same applies to a measurement of the survival probability in finite-size simulations, see Fig. 5.10. In practice data collapses are often used to determine the critical point by searching for the best possible collapse of the curves.

5.1.8. Universality

As outlined in the introduction, the working hypothesis in the field of continuous non-equilibrium phase transitions is the notion of *universality*. This concept expresses the expectation that the critical behavior of second-order phase transitions can be associated with a *finite* set of possible *universality classes*, each corresponding to a certain type of underlying field theory.

The phenomenon of Universality can be understood to some extent in the framework of field-theoretic renormalization group techniques. As we shall see in more detail below, the field-theoretic action involves certain relevant operators whose form is usually determined by the symmetry properties of the process, while other details of the microscopic dynamics lead to contributions which are irrelevant in the field-theoretic sense. This explains why various different models may belong to the same universality class.

**DP conjecture**

In particular the DP class – so to say the “Ising”-class of non-equilibrium statistical physics – is extremely robust with respect to the microscopic dynamic rules. The large variety and robustness of DP models led Janssen and Grassberger to the conjecture that a model should belong to the DP universality class if the following conditions hold [7, 7]:

1. The model displays a continuous phase transition from a *fluctuating* active phase into a *unique* absorbing state.

2. The transition is characterized by a *positive one-component* order parameter.

3. The dynamic rules involve only *short-range* processes.

4. The system has no unconventional attributes such as additional symmetries or quenched randomness.

Although this conjecture has not yet been proven rigorously, it is strongly supported by numerical evidence. In fact, DP seems to be even more general and has been identified even in systems that violate some of the four conditions.

The universality classes can be characterized in terms of their critical exponents and scaling
functions. Hence in order to identify a certain universality class, a precise estimation of the critical exponents is an important numerical task. In the case of DP, the numerical estimates suggest that the critical exponents are given by irrational numbers rather than simple rational values. In addition, the scaling functions, which have been neglected in the literature for long time, provide a wealth of useful information, as shown e.g. in a review by Lübeck [?].

5.2. Field-theoretic approach to Directed Percolation

5.2.1. Langevin equation

The upper critical dimension

As can be seen in Table 5.2, the values of the critical exponents depend on the spatial dimension of the system. For example, the exponent $\beta \approx 0.2765$ in 1D and grows with the dimension until it saturates at a constant value $\beta = 1$, as shown in the adjacent figure. The crossover to a simple constant value takes place at $d = 4$, and the same phenomenology can be observed for all other critical exponents as well. This dimension, from where on the critical exponents are constant, is called the upper critical dimension and is denoted by $d_c$.

What happens above the upper critical dimension? Generally, a high-dimensional space offers more freedom than a low-dimensional one. For the clusters of directed percolation this means that the spreading agent has more possibilities to spread, and therefore it is less likely that the infection returns to the same point from where it came. For this reason correlations are less significant in high dimensions. In fact, the simple values of the critical exponents observed for $d > d_c$ are just the mean field values – an approximation where correlations are ignored. Of course, this alone does not yet explain why the mean field behavior sets in sharply at $d = d_c = 4$.

To understand the specific value of $d_c$, we have to analyze the so-called Langevin equation of DP.

Langevin equation

Let us go back to the extended mean field equation in Eq. (5.20) for an inhomogeneous coarse-grained density $\rho(x,t)$ which includes a diffusive term with a Laplacian. Adding a factor $\lambda$ in front of the nonlinear term, this equation reads

$$\frac{d}{dt} \rho(x,t) = a \rho - \rho^2 + D \nabla^2 \rho.$$
If this equation starts with a homogeneous (i.e. x-independent) initial state, then \( \rho(x, t) \) will stay strictly homogeneous for all times. This is of course not true in the full model, where we have a stochastic process which evolves into an inhomogeneous fluctuating state independent of the initial state. What this extended mean field equation is missing is therefore an element of stochasticity. This can be included by adding another term with a ‘noise’, denoted as \( \xi(x, t) \), giving the so-called \textit{Langevin equation} for directed percolation:

\[
\partial_t \rho(x, t) = a \rho(x, t) - \lambda \rho^2(x, t) + D \nabla^2 \rho(x, t) + \xi(x, t) .
\] (5.29)

This Langevin equation can be derived more rigorously from the master equation of the contact process \cite{7}. It is a \textit{stochastic partial differential equation} with \textit{multiplicative noise}, i.e. the noise amplitude depends on the local density of active sites \( \rho(x, t) \). This dependence ensures that the noise vanishes as soon as there are no active sites, resembling the behavior of the original DP process.

More specifically, \( \xi(x, t) \) is a density-dependent white Gaussian noise field characterized by the correlations

\[
\langle \xi(x, t) \rangle = 0 ,
\] (5.30)

\[
\langle \xi(x, t) \xi(x', t') \rangle = \Gamma \rho(x, t) \delta^d(x - x') \delta(t - t') .
\] (5.31)

Since the amplitude of \( \xi(x, t) \) is proportional to \( \sqrt{\rho(x,t)} \), it is easy to see that the absorbing state \( \rho(x, t) = 0 \) does not fluctuate. The square-root behavior is related to the fact that the noise describes density fluctuations on a coarse-grained scale, which can be viewed as the sum of individual noise contributions generated by each particle averaged over some mesoscopic box size. According to the central limit theorem, if the number of particles in this box is sufficiently high, \( \xi(x, t) \) approaches a Gaussian distribution with an amplitude proportional to the square root of the number of active sites in the box.

\textbf{Scaling analysis of the Langevin equation}

Applying the scaling transformation \cite{5} to Eq. (5.29) with constant dilatation factor \( b \) we get

\[
b^{-\beta - \nu_\parallel} \partial_t \rho(x, t) = b^{-\beta - 1} a \rho(x, t) - b^{-2\beta} \lambda \rho^2(x, t) + b^{-\beta - 2\nu_\perp} D \nabla^2 \rho(x, t) + b^{\frac{1}{2}(-\beta - \nu_\perp - \nu_\parallel)} \xi(x, t) .
\] (5.32)

The last scaling factor marked in red color originates from the scaling behavior of the noise correlator in Eq. (5.30)

\[
\langle \xi(x, t) \xi(x', t') \rangle = \Gamma \rho(x, t) \delta^d(x - x') \delta(t - t')
\] (5.33)

\[
\rightarrow b^{-\beta - \nu_\perp - \nu_\parallel} \Gamma \rho(x, t) \delta^d(x - x') \delta(t - t') ,
\]

and since the correlator is quadratic in the noise, we have to take the square root in Eq. (5.32). Let us now impose scale invariance, i.e., we ask for values of the critical exponents in which the Langevin equation is invariant under scale transformations. To this end we first bring all scaling factors to the right side:

\[
\partial_t \rho(x, t) = b^{\nu_\parallel - 1} a \rho(x, t) - b^{\nu_\parallel - \beta} \lambda \rho^2(x, t) + b^{\nu_\perp - 2\nu_\parallel} D \nabla^2 \rho(x, t) + b^{\frac{1}{2}(\nu_\perp - \beta + \nu_\parallel)} \xi(x, t) .
\]

Obviously, the Langevin equation is a scale-invariant if all scaling factors in front of the terms drop out. From the first three terms on the right inside we can immediately read off the critical exponents

\[
\rho^{\text{MF}} = 1, \quad \nu_{\perp}^{\text{MF}} = 1/2, \quad \nu_{\parallel}^{\text{MF}} = 1.
\] (5.34)
This exponents are called the mean field (MF) critical exponents of directed percolation. Note that in contrast to the coefficients \(\lambda, D, \Gamma\), the coefficient \(a\) of the linear term is co-moving under scale transformations. If we want this coefficient to be invariant we have to set it to zero. This is just the mean field critical point \(a_c = 0\) were the phase transition takes place.

This consideration does not yet involve the scaling behavior of the noise in the last term. Interestingly, this is the only term where the dimension of the system enters explicitly (coming from the \(\delta^d\)-function in the noise correlator). Here we observe that \(\frac{1}{2}(+\beta - d v_\perp + v_\parallel)\) is positive for \(d < d_c = 4\) (which means that the term grows and the renormalization) and negative for \(d > d_c = 4\) (which means that the term scales to zero under renormalization). At this point we understand why the mean field behavior sets in at a particular upper critical dimension \(d_c\), and that the value of the upper critical dimension can be extracted by a simple dimensional analysis of the corresponding Langevin equation.

Note that the Langevin equation may also include higher order terms such as \(\rho^3(x, t), \nabla^4 \rho(x, t)\), or higher-order contributions of the noise, but a simple dimensional analysis along the same lines reveals that they all are irrelevant under renormalization. This explains why we observe the phenomenon of universality: The universal features are encoded in the lowest-order terms of the equation of motion, while higher-order terms, which scale to zero under renormalization, encode the non-universal microscopic details.

### 5.2.2. The integral equation for DP

Introducing the Fourier-transformed field \(\phi(k, \omega)\) by

\[
\phi(k, \omega) = \int_{-\infty}^{\infty} Dxt e^{i(k \cdot x - \omega t)} \rho(x, t), \quad \rho(x, t) = \int_{-\infty}^{\infty} Dk \omega e^{-i(k \cdot x - \omega t)} \phi(k, \omega)
\]

with the shortcut notation

\[
Dxt := d^d x dt, \quad Dk \omega = \frac{1}{(2\pi)^{d+1}} d^d k \, d\omega
\]

the DP Langevin equation can be written in the form

\[
- i \omega \phi(k, \omega) = a \phi(k, \omega) - \frac{1}{2} \int Dk' \omega' \phi(k', \omega') \phi(k - k', \omega' - \omega) - Dk^2 \phi(k, \omega) - \eta(k, \omega)
\]

Here \(\eta(k, \omega)\) denotes the Fourier-transformed noise

\[
\eta(k, \omega) = \int_{-\infty}^{\infty} Dxt e^{i(k \cdot x - \omega t)} \eta(x, t)
\]

with the correlations

\[
\langle \eta(k, \omega) \eta(k', \omega') \rangle = \Gamma \phi(k + k', \omega + \omega').
\]

Without the nonlinear term and without the noise Eq. (5.37) reduces to the so-called free equation

\[
- i \omega \phi(k, \omega) = a \phi(k, \omega) - Dk^2 \phi(k, \omega) - \eta(k, \omega).
\]

Defining the bare Greens function

\[
G_0(k, \omega) = \frac{1}{Dk^2 - a - i\omega'}
\]
which is also known as the free propagator, this equation is formally solved by the free solution
\[
\phi_0(k, \omega) = G_0(k, \omega) \eta(k, \omega).
\] (5.42)

Since \( \eta(k, \omega) = (Dk^2 - a - i\omega) \phi_0(k, \omega) \) we can eliminate the noise in Eq. (5.37), obtaining an integral equation of the form
\[
\phi(k, \omega) = \phi_0(k, \omega) - \frac{\lambda}{2} G_0(k, \omega) \int Dk' \omega' \phi(k', \omega') \phi(k - k', \omega - \omega') - \frac{1}{2} \lambda \phi^2 - \eta.
\] (5.43)

This is the Schwinger-Dyson equation. Apart from the assumptions made at the outset of theory, namely, the formulation terms of coarse-grained densities, this equation chemical considered as exact.

5.2.3. Field-theoretic formulation of DP

Assume that we have an equilibrium system in which those states and realize that satisfy the DP Langevin equation. The partition sum of such a system is the integral over all realizations of the field \( \phi(x, t) \) and all possible realizations of the noise \( \eta(x, t) \) (weighted by the probability to find this particular realization of the noise) times a \( \delta \)-function with the Langevin equation as its argument:
\[
Z \sim \int D[\eta] P[\eta] \int D[\phi] \delta\left( \partial_t \phi - D \nabla^2 \phi - a \phi + \frac{1}{2} \lambda \phi^2 - \eta \right).
\] (5.44)

Here \( \int D[...] \) denotes functional integration over all realizations while \( P[\eta] \) is the probability distribution to find the noise \( \eta[x, t] \). Since the noise is assumed to be white and uncorrelated, its distribution will be Gaussian:
\[
P[\eta] = \frac{1}{g[\phi]} \exp\left( - \int Dxt \frac{\eta^2(x, t)}{2 \Gamma \phi(x, t)} \right),
\] (5.45)

where \( g[\phi] \) plays the role of a (\( \phi \)-dependent) normalization constant. In order to handle the \( \delta \)-function in Eq. (5.44), we use a simple trick, namely, we employ the well-known representation of the \( \delta \)-function
\[
\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk \ e^{ikx}
\] (5.46)

reformulated as a functional
\[
\delta[f] \propto \int D[\tilde{\phi}] \exp\left( i \int d^d x \ dt \tilde{\phi}(x, t) f(x, t) \right)
\] (5.47)

This functional can be interpreted as follows: \( \delta[f] \) is infinite if the function \( f(x, t) \) vanishes identically everywhere, otherwise \( \delta[f] = 0 \). Notice that we have introduced a new field \( \tilde{\phi}(x, t) \) which plays the role of the wavenumber \( k \) in Eq. (5.46). Inserting this
expression into the partition sum, we obtain

$$Z \sim \int D[\eta] P[\eta] \int D[\phi] D[\hat{\phi}] \exp \left[ i \int d^4x dt \left( \hat{\phi} (\partial_t - D \nabla^2 - a) \phi - \frac{1}{2} \lambda \hat{\phi} \phi^2 - \phi \eta \right) \right].$$  

(5.48)

Now a mysterious operation performed, of which I have only a very limited understanding, namely, the so-called **Wick rotation**. A Wick rotation is basically something like analytic continuation in the complex plane, which in the present case amounts to replacing the imaginary unit $i$ in the exponential by a minus sign. Out the Wick rotation and pulling all contributions independent of the noise in front, the partition sum can be rewritten as

$$Z \sim \int D[\phi] D[\hat{\phi}] \exp \left[ - \int d^4x dt \left( \hat{\phi} (\partial_t - D \nabla^2 - a) \phi - \frac{1}{2} \lambda \hat{\phi} \phi^2 \right) \right] + \mathcal{N}[\phi, \hat{\phi}]$$  

(5.49)

where

$$\mathcal{N}[\phi, \hat{\phi}] = \int D[\eta] P[\eta] \exp \left( \int d^4x dt \hat{\phi} \eta \right)$$  

(5.50)

captures the noise contributions. Inserting the noise probability distribution Eq. (5.45) we get an expression which can be integrated by quadratic completion:

$$\mathcal{N}[\phi, \hat{\phi}] = g[\phi] \int D[\eta] \exp \left( \int d^4x dt \left( \phi \eta - \frac{\eta^2}{2 \Gamma \phi} \right) \right) = g[\phi] \exp \left( \frac{\Gamma}{2} \int d^4x dt \phi \hat{\phi}^2 \right),$$  

(5.51)

just in the same way as we compute a Gaussian integral by quadratic completion:

$$\int_{-\infty}^{+\infty} dx \frac{1}{\sqrt{2\pi a}} \exp \left( bx - \frac{x^2}{2a} \right) = \exp \left( \frac{1}{2} ab^2 \right).$$  

(5.52)

The resulting partition function reads

$$Z \sim \int D[\phi] D[\hat{\phi}] \exp \left[ - \int d^4x dt \left( \hat{\phi} [\partial_t - D \nabla^2 - a] \phi + \frac{\lambda}{2} \hat{\phi} \phi^2 - \Gamma \hat{\phi}^2 \phi \right) \right]$$  

(5.53)

Here we recognize the inverse bare propagated the square brackets. Note that we have successfully eliminated the noise. This comes at the price to introduce an additional fluctuating field, namely, $\hat{\phi}$. A non-equilibrium statistical physics this field is termed the **response field**. Roughly speaking the response field can be thought of as a creation operator, which generates an infection, while the ordinary field $\phi$ can be understood as a and elation operator measuring the intensity of the infection at a given time and position. Two-point correlation functions therefore involve the product of $\phi$ and $\hat{\phi}$.

In Eq. (5.53) we have two nonlinear terms which are cubic in the fields. They are weighted by two different coefficients $\lambda$ and $\Gamma$. In what follows it is convenient to assume that both coefficients coincide, i.e., $\lambda = \Gamma$. This can be achieved by rescaling the fields $\phi$ and $\hat{\phi}$ by reciprocal factors

$$\phi \mapsto \sqrt{\frac{\lambda}{\Gamma}} \phi, \quad \hat{\phi} \mapsto \sqrt{\frac{\Gamma}{\lambda}} \hat{\phi}$$  

(5.54)
and replacing $\sqrt{\lambda \Gamma}$ by $\Gamma$. The resulting partition sum reads

$$Z \sim \int D[\phi] \int D[\tilde{\phi}] \exp[-\mathcal{L}_0 - \mathcal{L}_{\text{int}}],$$

(5.55)

where

$$\mathcal{L}_0[\phi, \tilde{\phi}] = \int d^d x \, dt \, \tilde{\phi}(x, t) \left( \partial_t - D \nabla^2 - a \right) \phi(x, t)$$

(5.56)

is the free Lagrangian while

$$\mathcal{L}_{\text{int}}[\phi, \tilde{\phi}] = \int d^d x \, dt \, \frac{\Gamma}{2} \phi(x, t) \tilde{\phi}(x, t) \left( \phi(x, t) - \tilde{\phi}(x, t) \right)$$

(5.57)

is the interaction part. Interestingly, this field theory, called Reggeon field theory, made its first appearance in the 60s in the context of high-energy physics, where it was suggested as an effective interaction between bosons in order to explain the phenomenology of strong interaction. However, with the discovery of QCD, Reggeon field theory was dismissed and forgotten. Later, in the mid-eighties, this field theory was rediscovered by Janssen and Grassberger, who realized that it is suitable to describe the phase transition in directed percolation.

Note that the Lagrangians given above are invariant under the replacement

$$\phi(x, t) \mapsto \tilde{\phi}(x, -t), \quad \tilde{\phi}(x, t) \mapsto -\phi(x, -t).$$

(5.58)

This remarkable symmetry, which is known as the rapidity reversal symmetry of directed percolation, implies that the two fields $\phi$ and $\tilde{\phi}$ have to scale identically. Consequently, the exponents $\beta$ and $\beta'$ introduced above coincide. This symmetry can be checked explicitly on the lattice in directed bond percolation process (see exercise).

### 5.2.4. Multifractal properties of currents on directed percolation clusters

So far we have seen that the critical behavior of DP and other absorbing phase transitions can be described in terms of scaling laws that involve three independent critical exponents $\eta, \nu_\perp,$ and $\nu_\parallel$. This type of scaling is usually referred to as simple scaling, as opposed to multiscaling, where a whole spectrum of exponents exists. For example, in DP at criticality starting with a homogeneous initial state any integral power $\rho^n$ of the order parameter $\rho$ scales in the same way, i.e.

$$\rho^n(t) \sim t^{-\delta}, \quad n = 1, 2, 3, \ldots.$$  

(5.59)

Let us now consider an electric current running on a directed percolation cluster according to Kirchhoff’s laws, interpreting the cluster as a random resistor-diode network. By introducing such a current the theory is extended by an additional physical concept. In fact, even though the DP cluster itself is known to be characterized by simple scaling laws, a current running on it turns out to be distributed in a multifractal manner. This phenomenon was first discovered in the case of isotropic percolation [?, ?] and then confirmed for DP [?, ?].

As shown in Fig. 5.11, in directed bond percolation at criticality an electric current $I$
running from one point to the other is characterized by a non-trivial distribution of currents. The multifractal structure of this current distribution can be probed by studying the moments
\[ M_\ell := \sum_b \left( \frac{I_b}{I} \right)^\ell \quad \ell = 0, 1, 2, \ldots, \quad (5.60) \]
where the sum runs over all bonds \( b \) that transport a non-vanishing current \( I_b > 0 \).
For example, \( M_0 \) is just the number of conducting bonds while \( M_1 \) is essentially the total resistance between the two points. \( M_2 \) is the second cumulant of the resistance fluctuations and can be considered as a measure of the noise in a given realization. Finally, \( M_\infty \) is the number of so-called red bonds that carry the full current \( I \). The quantity \( M_\ell \) is found to scale as a power law
\[ M_\ell (t) \sim t^{\psi_\ell/\nu}. \quad (5.61) \]
In the case of simple scaling, the exponents \( \psi_\ell \) would depend linearly on \( \ell \). In the present case, however, a non-linear dependence is found both by field-theoretic as well as numerical methods (see Ref. [?]). This proves that electric currents running on DP clusters have multifractal properties.

Again it should be emphasized that multifractality is not a property of DP itself, rather it emerges as a new feature whenever an additional process, here the transport of electric currents, is confined to live on the critical clusters of DP.

### 5.2.5. Characterizing non-equilibrium transition by Yang-Lee zeroes in the complex plane

In equilibrium statistical mechanics a large variety of continuous phase transitions has been analyzed by studying the distribution of so-called Yang-Lee zeros [?, ?, ?]. To determine these zeros the partition sum of a (finite) equilibrium system is expressed as a polynomial of the control parameter, which is usually a function of temperature. E.g., for the Ising model the zeros of this polynomial lie on a circle in the complex plane and heckle the real line from both sides in the vicinity of the phase transition as the system size increases. This explains why the analytic behavior in finite system crosses over to a non-analytic behavior at the transition point in the thermodynamic limit.
Recently, it has been shown that the concept of Yang and Lee can also be applied to non-equilibrium systems [?], including DP [?]. To this end one has to consider the order parameter in a finite percolation tree as a function of the percolation probability $p$ in the complex plane. This can be done by studying the survival probability $P(t)$ (see Eq. (5.10)), which is defined as the probability that a cluster generated in a single site at time $t = 0$ survives at least up to time $t$. In fact, the partition sum of an equilibrium system and the survival probability of DP are similar in many respects. They both are positive in the physically accessible regime and can be expressed as polynomials in finite systems. As the system size tends to infinity, both functions exhibit a non-analytic behavior at the phase transition as the Yang-Lee zeros in the complex plane approach the real line.

In directed bond percolation the survival probability is given by the sum over the weights of all possible configurations of bonds, where each conducting bond contributes to the weight with a factor $p$, while each non-conducting bond contributes with a factor $1 - p$. As shown in Ref [?], the polynomial for the survival probability can be expressed as a sum over all cluster configurations $c$ reaching the horizontal row at time $t$. The polynomial is of the form

\[ P(t) = \sum_c p^n(1 - p)^m, \]  

(5.62)

where $n$ denotes the number of bonds while $m$ is the number of bonds belonging to its cluster’s hull. Summing up all weights in Eq. (5.62), one obtains a polynomial of degree $t^2 + t$. For example, the first few polynomials are given by

\[ P(0) = 1 \]  

(5.63)

\[ P(1) = 2p - p^2 \]

\[ P(2) = 4p^2 - 2p^3 - 4p^4 + 4p^5 - p^6 \]

\[ P(3) = 8p^3 - 4p^4 - 10p^5 - 3p^6 + 18p^7 + 5p^8 - 30p^9 + 24p^{10} - 8p^{11} + p^{12} \]

\[ P(4) = 16p^4 - 8p^5 - 24p^6 - 8p^7 + 6p^8 + 84p^9 - 29p^{10} - 62p^{11} - 120p^{12} \]

\[ + 244p^{13} + 75p^{14} - 470p^{15} + 495p^{16} - 268p^{17} + 83p^{18} - 14p^{19} + p^{20} \]

As $t$ increases, the number of cluster configurations grows rapidly, leading to complicated polynomials with very large coefficients. The distribution of zeros in the complex plane for $t = 15$ is shown in Fig. 5.12. As can be seen, the distribution reminds of a fractal, perhaps being a signature of the non-integrable nature of DP. As expected, the zeros approach the phase transition point from above and below. Their distance to the transition point is found to scale as $t^{-1/\nu}$ in agreement with basic scaling arguments.

5.3. Other classes of absorbing phase transitions

So far we discussed directed percolation as the most important class of non-equilibrium phase transitions into absorbing states. Because of the robustness of DP it is interesting to search for other universality classes. The ultimate goal would be to set up a table of possible non-trivial universality classes from active phases into absorbing states.

Although various exceptions from DP have been identified, the number of firmly
established universality classes is still small. A recent summary of the status quo can be found in Refs. [?, ?]. In these lecture notes, however, we will only address the most important classes with local interactions.

### 5.3.1. Parity-conserving particle processes

The parity-conserving (PC) universality class comprises phase transitions that occur in reaction-diffusion processes of the form

\[ A \rightarrow (n+1)A \]

\[ 2A \rightarrow \emptyset \]  \hspace{1cm} (5.64)

combined with single-particle diffusion, where the number of offspring \( n \) is assumed to be even. As an essential feature, these processes conserve the number of particles modulo 2. A particularly simple model in this class with \( n = 2 \) was proposed by Zhong and ben-Avraham [?]. The estimated critical exponents

\[ \eta = \eta' = 0.92(2), \quad \nu_\parallel = 3.22(6), \quad \nu_\perp = 1.83(3) \]  \hspace{1cm} (5.66)

differ significantly from those of DP, establishing PC transitions as an independent universality class. The actual values of \( \delta \) and \( \theta \) depend on the initial condition. If the process starts with a single particle, it will never stop because of parity conservation, hence \( \delta = 0 \), i.e. the usual relation \( \delta = \eta / \nu_\parallel \) does no longer hold. However, if it starts with two particles, the roles of \( \delta \) and \( \theta \) are exchanged, i.e. \( \theta = 0 \). The theoretical reasons for this exchange are not yet fully understood.

The relaxational properties in the subcritical phase differ significantly from the standard DP behavior. While the particle density in DP models decays exponentially as
Nonequilibrium phase transitions

$\rho(t) \sim e^{-t/\xi}$, in PC models it decays algebraically since the decay is governed by the annihilation process $2A \rightarrow \emptyset$.

A systematic field theory for PC models can be found in Refs. [?, ?], confirming the existence of the annihilation fixed point in the inactive phase. However, the field-theoretic treatment at criticality is extremely difficult as there are two critical dimensions: $d_c = 2$, above which mean-field theory applies, and $d'_c \approx 4/3$, where for $d > d'_c$ ($d < d'_c$) the branching process is relevant (irrelevant) at the annihilation fixed point. Therefore, the physically interesting spatial dimension $d = 1$ cannot be accessed by a controlled $\epsilon$-expansion down from upper critical dimension $d = 2$.

5.3.2. The voter universality class

Order-disorder transition in models with a $Z_2$-symmetry which are driven by interfacial noise belong to the so-called voter universality class [?]. As will be explained below, the voter class and the parity conserving class are identical in one spatial dimension but different in higher dimensions.

To understand the physical mechanism that generates the phase transition in the voter model, let us first discuss the difference between interfacial and bulk noise. Consider for example the Glauber-Ising model in two spatial dimensions at $T = 0$. This model has two $Z_2$-symmetric absorbing states, namely, the two fully ordered states. Starting with a random initial configuration one observes a coarsening process forming ordered domains whose size grows as $\sqrt{t}$. In the Ising model at $T = 0$ domain growth is curvature-driven, leading to an effective surface tension of the domain walls. In fact, as shown in Fig. 5.13 the domain walls produced by the Glauber-Ising model appear to be smooth and indeed the density of domain walls is found to decay as $1/\sqrt{t}$.

Increasing temperature occasional spin flips occur, leading to the formation of small minority islands inside the existing domains. For small temperature the influence of surface tension is strong enough to eliminate these minority islands, stabilizing the ordered phase. However, increasing $T$ above a certain critical threshold $T_c$ this mechanism breaks down, leading to the well-known order-disorder phase transition in the Ising model. Thus, from the perspective of a dynamical process, the Ising transition results from a competition between surface tension of domain walls and bulk noise.

Let us now compare the Glauber-Ising model with the classical voter model in two
spatial dimensions. The classical voter model \[?] is a caricatural process in which sites (voters) on a square lattice adopt the opinion of a randomly-chosen neighbor. As the Ising model, the voter model has two symmetric absorbing states. Moreover, an initially disordered state coarsens. However, as shown in the right panel of Fig. 5.13 already the visual appearance is very different. In fact, in the voter model the domain sizes are found to be distributed over the whole range between 1 and \(\sqrt{t}\). Moreover, in contrast to the Glauber-Ising model, the density of domain walls decays only logarithmically as \(1/\ln t\). This marginality of the voter model is usually attributed to the exceptional character of its analytic properties \[?, ?, ?\] and may be interpreted physically as the absence of surface tension.

In the voter model even very small thermal bulk noise would immediately lead to a disordered state. However, adding interfacial noise one observes a non-trivial continuous phase transition at a finite value of the noise amplitude. Unlike bulk noise, which flips spins everywhere inside the ordered domains, interfacial noise restricts spin flips to sites in the vicinity of domain walls.

Recently Al Hammal et al. \[?\] introduced a Langevin equation describing voter transitions. It is given by

\[
\frac{\partial}{\partial t} \rho = (a\rho - b\rho^3)(1 - \rho^2) + D \nabla^2 \rho + \sigma \sqrt{1 - \rho^2} \xi, \tag{5.67}
\]

where \(\xi\) is a Gaussian noise with constant amplitude. For \(b > 0\) this equation is found to exhibit separate Ising and DP transitions, while for \(b \leq 0\) a genuine voter transition is observed. With these new results the voter universality class is now on a much firmer basis than before.

In one spatial dimension, kinks between domain walls can be interpreted as particles. Here interfacial noise between two domains amounts to generating pairs of additional domain walls nearby. This process, by its very definition, conserves parity and can be interpreted as offspring production \(A \rightarrow 3A, 5A, \ldots\) while pairwise coalescence of domain walls corresponds to particle annihilation \(2A \rightarrow \emptyset\). For this reason the voter class and the parity-conserving class coincide in one spatial dimension. However, their behavior in higher dimensions, in particular the corresponding field theories, are expected to be different. Loosely speaking, the parity-conserving class deals with the dynamics of zero-dimensional objects (particles), while in the voter class the objects of interest are \(d\)-1-dimensional hypermanifolds (domain walls).

5.3.3. Absorbing phase transitions with a conserved field

According to the conjecture by Janssen and Grassberger (cf. Sect. 5.1.8), non-DP behavior is expected if the dynamics is constrained by additional conservation laws. For example, as shown in the previous subsections, parity conservation or a \(Z_2\)-symmetry may lead to different universality classes. Let us now consider phase transitions in particle processes in which the total number of particles is conserved. According to an idea by Rossi et al. \[?\] this leads to a different universality class of phase transitions which is characterized by an effective coupling of the process to a non-diffusive conserved field. Models in this class have infinitely many absorbing states and are related to certain
models of self-organized criticality (for a recent review see Ref. [?]).

As an example let us consider the conserved threshold transfer process (CTTP). In this model each lattice site can be vacant or occupied by either one or two particles. Empty and single occupied sites are considered as inactive while double occupied sites are regarded as active. According to the dynamical rules each active site attempts to move the two particles randomly to neighboring sites, provided that these target sites are inactive. By definition of these rules, the total number of particles is conserved. Clearly, it is the background of solitary particles that serves as a conserved field to which the dynamics of active sites is coupled.

In \( d \geq 2 \) spatial dimensions this model shows the same critical behavior as the Manna sand pile model [?]. The corresponding critical exponents in \( d = 2 \) dimensions were estimated by [?]

\[
\eta = 0.639(9), \quad \eta' = 0.624(29), \quad \nu_\perp = 0.799(14), \quad \nu_\parallel = 1.225(29).
\] (5.68)

Obviously, this set of exponents differs from those of all other classes discussed above. In one spatial dimension the situation is more complicated because of a split of the CTTP and Manna universality classes, as described in detail in Ref. [?]

5.3.4. The diffusive pair contact process

Among the known transitions into absorbing states, the transition occurring in the so-called contact process with diffusion (PCPD) is probably the most puzzling one (see Ref. [?] for a recent review). The PCPD is a reaction-diffusion process of particles which react spontaneously whenever two of them come into contact. In its simplest version the PCPD involves two competing reactions, namely

\[
\text{fission:} \quad 2A \rightarrow 3A, \\
\text{annihilation:} \quad 2A \rightarrow \emptyset.
\]

In addition individual particles are allowed to diffuse. Moreover, there is an additional mechanism such that the particle density cannot diverge. In models with at most one particle per site this mechanism is incorporated automatically.

The PCPD displays a non-equilibrium phase transition caused by the competition of fission and annihilation. In the active phase, the fission process dominates, maintaining a fluctuating steady-state, while in the subcritical phase the annihilation process dominates so that the density of particles decreases continuously until the system reaches one of the absorbing states. The PCPD has actually two absorbing states, namely, the empty lattice and a homogeneous state with a single diffusing particle.

The pair-contact process with diffusion was already suggested in 1982 by Grassberger [?], who expected a critical behavior “distinctly different” from DP. Eight years ago the problem was rediscovered by Howard and Täuber [?], who proposed a bosonic field-theory for the one-dimensional PCPD. In this theory the particle density is unrestricted and thus diverges in the active phase. The first quantitative study of a restricted PCPD by Carlon et al. [?] using DMRG techniques led to controversial results.
5.3 Other classes of absorbing phase transitions

Figure 5.14.: High-performance simulation of the PCPD model introduced by Kockelkoren and Chaté. The plot shows the density of active sites multiplied by the expected power law. As can be seen, the lines are slightly curved. Kockelkoren and Chaté simulated the process up to about $10^7$ Monte Carlo updates (dotted line), identifying the blue line in the middle as critical curve. Extending these simulations by one decade one recognizes that this curve is actually subcritical and that the true critical threshold has to be slightly higher. Obviously a slow drift towards DP (slope indicated by dashed line) cannot be excluded.

and released a still ongoing debate concerning the asymptotic critical behavior of the 1+1-dimensional PCPD at the transition. Currently the main viewpoints are that the PCPD:

- represents a new universality class with well-defined critical exponents [?],
- represents two different universality classes depending on the diffusion rate [?, ?] and/or the number of space dimensions [?],
- can be interpreted as a cyclically coupled DP and annihilation process [?],
- is a marginally perturbed DP process with continuously varying exponents [?],
- may have exponents depending continuously on the diffusion constant [?],
- may cross over to DP after very long time [?, ?], and
- is perhaps related to the problem of non-equilibrium wetting in 1+1 dimensions [?].

Personally I am in favor of the conjecture that the PCPD in 1+1 dimensions belongs to the DP class. This DP behavior, however, is masked by extremely slow (probably logarithmic) corrections. Searching the critical point by fitting straight lines in a double logarithmic plot may therefore lead to systematic errors in the estimate of the critical threshold since the true critical line is not straight but slightly curved. This in turn leads to even larger systematic errors for the critical exponents. However, as the computational effort is increased, these estimates seem to drift towards DP exponents.

The problem of systematic errors and drifting exponents can be observed for example in the work by Kockelkoren and Chaté, who tried to establish the PCPD as a new universality class as part of a general classification scheme [?]. Introducing a particu-
larly efficient model they observed clean power laws in the decay of the density over several decades, leading to the estimates

\[ \delta = \frac{\eta}{\nu_\parallel} = 0.200(5), \quad z = \frac{\nu_\perp}{\nu_\parallel} = 1.70(5), \quad \eta = 0.37(2). \]  

(5.69)

However, increasing the numerical effort by a decade in time, it turns out that their critical point \( p_c = 0.795410(5) \), including its error margin, lies entirely in the inactive phase (see Fig. 5.14). In the attempt to obtain an apparent power-law behavior, it seems that the authors systematically underestimated the critical point.

Presently it is still not yet clear whether the PCPD belongs to the DP universality class or not. Apparently computational methods have reached their limit and more sophisticated techniques are needed to settle this question.

5.4. Epidemic spreading with long-range interactions

Directed percolation is often used as a caricatural process for epidemic spreading. Suppose that infected and healthy individuals are sitting in a train, as shown in Fig. 5.15. On the one hand, infected people infect their nearest neighbors with a certain probability per unit time. On the other hand, infected individuals may recover spontaneously. Depending on the rates for infection and recovery, this toy model for epidemic spreading just resembles a simple DP process.

Although DP is too simplistic to describe epidemic spreading in reality, there are some important analogies. Certainly, epidemic spreading in Nature is a non-equilibrium process with a transition-like behavior at some threshold of the infection rate. For example, as an increasing number of people refuses vaccinations, the question arises at which percentage of unprotected individuals certain diseases, that became almost extinct, will again percolate through the society.

Epidemic spreading in Nature is of course a much more complex phenomenon. For example, it takes place in a very disordered environment and involves short- and long-range interactions. Moreover, individuals protect themselves by sophisticated immunization strategies. Certainly, physicist will never be able to predict epidemic spreading in Nature quantitatively. However, it is possible to extend DP towards a more realistic description of epidemic spreading and to study how they influence the behavior at the transition. Some of these extensions will be discussed in the following.
5.4 Epidemic spreading with long-range interactions

5.4.1. Immunization and mutations

As a first step towards a more realistic description of epidemic spreading we may include the effect of immunization. For example, we may declare all sites that were active at least once in the past as immune. One then introduces two different infection probabilities, namely, a probability for first infection \( p_0 \), and a second (usually smaller) probability \( p \) for the reinfection of immune sites. The case of perfect immunization (vanishing reinfection probability) is known as general epidemic process \([\text{?}]\) which can be regarded as a dynamical procedure to grow isotropic percolation clusters.

Introducing a finite reinfection probability one obtains the phase diagram shown in Fig. 5.16. It comprises a curved phase transition line with the same critical behavior as in the generalized epidemic process which separates phases of finite and annular growth. Moreover, there is a horizontal transition line above which compact cluster growth is observed. The critical properties along this line are partly dictated by the DP behavior inside immune regions, combined with non-universal properties for the growth of the clusters at its boundaries \([\text{?}]\). Both transition lines meet in a point with an interesting multicritical behavior. Extending this model by possible mutations of the spreading agent, the memory of immunization is lost. As a result one observes a controlled crossover back to DP \([\text{?}]\).

5.4.2. Long-range infections

Realistic diseases spread by different transport mechanisms, including direct contact between local individuals, transport by carriers such as mosquitos, and long-range transport e.g. by air planes. Usually it is very difficult to predict how these transport mechanism contribute to epidemic spreading. As an interesting empirical approach, Brockmann and Geisel traced the spatio-temporal trajectories of individual dollar notes within the United States \([\text{?},\text{?}]\). In agreement with previous conjectures \([\text{?}]\) they found out that the transport distances are distributed algebraically with some empirical exponent. Moreover, the time intervals at which the dollar notes were registered were found
to obey a power law as well.

Motivated by such empirical studies it is near at hand to generalize DP such that the spreading distances \( r \) are distributed as a power law

\[
P(r) \sim r^{-d-\sigma}, \quad (\sigma > 0)
\]

where \( \sigma \) is a control exponent. In the literature such algebraically distributed long-range displacements are known as \textit{Lévy flights} \cite{levy} and have been studied extensively e.g. in the context of anomalous diffusion \cite{anomalous}. In the present context of epidemic spreading it turns out that such long-range flights do not destroy the transition, instead they change the critical behavior provided that \( \sigma \) is sufficiently small. More specifically, it was observed both numerically and in mean field approximations that the critical exponents change \textit{continuously} with \( \sigma \) \cite{disorder,levy,field}. As a major breakthrough, Janssen et al introduced a renormalizable field theory for epidemic spreading transitions with spatial Lévy flights \cite{field}, computing the critical exponents to one-loop order. Because of an additional scaling relation only two of the three exponents were found to be independent. These results were confirmed numerically by Monte Carlo simulations \cite{monte}. 

### 5.4.3. Incubation times

As a second generalization one can introduce a similar long-range mechanism in \textit{temporal} direction. Such ‘temporal’ Lévy flights may be interpreted as incubation times \( \Delta t \) between catching and passing on the infection. As in the first case, these incubation times are assumed to be algebraically distributed as

\[
P(\Delta t) \sim \Delta t^{-1-\kappa}, \quad (\kappa > 0) \quad (5.71)
\]

where \( \kappa \) is a control exponent. However, unlike spatial Lévy flights, which take place equally distributed in all directions, such temporal Lévy flights have to be directed forward in time. Again it was possible to compute the exponents by a field-theoretic renormalization group calculation \cite{field}.

Recently, we studied the mixed case of epidemic spreading by spatial Lévy flights \textit{combined} with algebraically distributed incubation times \cite{mixed}. In this case the corresponding field theory was found to render two additional scaling relations, namely,

\[
2\eta - (\sigma - d)v_\perp - v_\parallel = 0, \quad (5.72)
\]

\[
2\eta - d v_\perp + (\kappa - 1)v_\parallel = 0. \quad (5.73)
\]

Hence only one of the three exponents is independent. In particular, the dynamical exponent \( z \) locks onto the ratio \( \sigma / \kappa \). A systematic numerical and field-theoretic study leads to a generic phase diagram in terms of the control exponents \( \sigma \) and \( \kappa \) which is shown in Fig. 5.17. It includes three types of mean-field phases, a DP phase, two phases corresponding to purely spatial or purely temporal Lévy flights, and a novel fluctuation-dominated phase describing the mixed case in which the critical exponents have been computed by a field-theoretic renormalization group calculation to one-loop order.
5.5 Surface growth and non-equilibrium wetting

Another interesting direction of non-equilibrium physics is the study of wetting far from equilibrium. Wetting phenomena occur in a large variety of experiments, where a planar substrate is exposed to a gas phase. Usually the term ‘wetting’ refers to a situation where a bulk phase in contact with a substrate coexists with a layer of a different phase which is preferentially attracted to the surface of the substrate. By changing physical parameters such as temperature and chemical potential, the system may undergo a wetting transition from a non-wet phase, where the thickness of the layer stays finite, to a wet phase, where the layer becomes macroscopic.

In many experimental situations it is reasonable to assume that a wetting stationary layer is in thermal equilibrium. In fact, methods of equilibrium statistical mechanics turned out to be very successful in a large variety of theoretical and experimental studies [?]. Therefore, the question arises whether models for wetting far from equilibrium may exhibit new physical phenomena which cannot be observed under equilibrium conditions.

Non-equilibrium wetting is usually modeled as a Kardar-Parisi-Zhang (KPZ) growth process [?] growing on top of a hard substrate. Theoretically such a system can be described by a KPZ equation in a potential [?, ?]

\[
\frac{\partial h(x,t)}{\partial t} = \sigma \nabla^2 h(x,t) - \frac{\partial V(h(x,t))}{\partial h(x,t)} + \lambda (\nabla h(x,t))^2 + \xi(x,t),
\]

(5.74)

where \(\xi(x,t)\) is a Gaussian noise. It is important to note that the nonlinear term \(\lambda (\nabla h(x,t))^2\) in this equation is a relevant perturbation of the underlying field theory, i.e., even if \(\lambda\) is very small, it will be amplified under renormalization group transformations, driving the system away from thermal equilibrium. In fact, as can be shown by constructing a closed loop, it is this term that breaks detailed balance.
Some time ago we introduced a simple solid-on-solid (SOS) model for non-equilibrium wetting in 1+1 dimensions \cite{?, ?, ?}. The model is controlled by an adsorption rate $q$, a desorption rate $p$, and optionally by a special deposition rate $q_0$ on sites at height zero (desorption at the edges takes place at rate 1, see Fig. 5.18). Setting $q_0 = q$ and varying the growth rate the model exhibits a continuous wetting transition at a certain critical growth rate $q_c(p)$. This wetting transition is related to the unpinning process of an interface from the substrate. Moreover, for $q_0 \neq q$ the model can emulate a short-range interaction between the interface and the substrate \cite{?, ?}. It was found that sufficiently strong attractive interaction modifies the nature of the wetting transition and makes it first order. In addition, it has been demonstrated that there exists an extended region in the phase diagram, where the pinned and the moving phases coexist in the sense that the transition time from the pinned to the moving phase grows exponentially with the system size so that the two phases become stable in the thermodynamic limit. This type of phase coexistence is in fact a new phenomenon that occurs only far away from equilibrium and should be experimentally observable.
6. Systems out of equilibrium

So far we have studied complex systems in thermal equilibrium. Thermal equilibrium means that (a) the system is stationary, and (b) there are no physical currents flowing through the system, as, for example, heat, particles, or electrical currents. Contrarily, whenever such currents are present, the system is said to be out of equilibrium. In this chapter we are going to study such situations in more detail.

6.1. Dynamics of subsystems

Recalling basic facts about Markov processes

Before starting, let us briefly summarize the formalism of stochastic Markov processes developed in Chapters 1 and 2. At any time $t$ the system is in a certain configuration $c \in \Omega$, where $\Omega$ denotes the configuration space. The system evolves in time by spontaneous jumps from one configuration to the other with certain transition rates $w_{c \rightarrow c'}$. These transition rates are not necessarily symmetric, i.e., generally we have $w_{c \rightarrow c'} \neq w_{c' \rightarrow c}$.

The actual trajectory of the system cannot be predicted. What can be predicted is the probability $P_c(t)$ to find the system at time $t$ in the configuration $c$. Introducing the probability currents

$$J_{c \rightarrow c'}(t) = P_c(t)w_{c \rightarrow c'}$$

these probabilities evolve deterministically according to the master equation

$$\frac{\partial}{\partial t}P_c(t) = \sum_{c' \neq c} J_{c' \rightarrow c}(t) - \sum_{c' \neq c} J_{c \rightarrow c'}(t),$$

which is the linear partial differential equation. Listing these probabilities in a canonical order as components of a vector $|P_t\rangle$, we can rewrite the master equation compactly as

$$\partial_t |P_t\rangle = -\mathcal{L}|P_t\rangle.$$ (6.3)

Here $\mathcal{L}$ is the Liouville operator with the matrix elements

$$\langle c' | \mathcal{L} | c \rangle = -w_{c \rightarrow c'} + \delta_{c,c'} \sum_{c''} w_{c \rightarrow c''},$$ (6.4)

where the diagonal elements $\langle c | \mathcal{L} | c \rangle = \sum_{c''} w_{c \rightarrow c''}$ are the so-called escape rates describing how much probability flows away from $c$ to elsewhere.
The formal solution of the master equation reads
\[ |P_t\rangle = \exp(-\mathcal{L}t)|P_0\rangle. \] (6.5)

This solution looks simple, but the evaluation of the matrix exponential function requires to diagonalize the Liouvillian, which is generally a non-trivial task.

Introducing the bra vector \( \langle 1 | = (1,1,1,\ldots,1) \) the normalization condition (probability conservation) requires that \( \langle 1 | P_t \rangle = 1 \) holds for all times, implying \( \langle 1 | \mathcal{L} = 0 \). This means that the column sum of the Liouvillian matrix vanishes. Consequently, the column sum of a finite-time evolution matrix \( T = e^{-\mathcal{L}t} \) is equal to 1.

A Markov process is called stationary if the probability distribution \( P_c(t) \) does not depend on time. In the vector space formalism, this means that \( |P_{\text{stat}}\rangle \) is a right eigenvector to the eigenvalue zero:
\[ \partial_t |P_{\text{stat}}\rangle = -\mathcal{L} |P_{\text{stat}}\rangle = 0. \] (6.6)

A system is called isolated if it does not interact by any means with the outside world. Arguing that the quantum-mechanical evolution of an isolated system would be unitary and thus invariant under time reversal, we concluded that the corresponding transition rates have to be symmetric, i.e.,
\[ w_{c\rightarrow c'} = w_{c'\rightarrow c}, \quad \text{(isolated system)} \] (6.7)
meaning that the matrix \( \mathcal{L} \) of an isolated system is symmetric. This implies that left and right eigenvectors have the same components, hence \( \mathcal{L}^T |1\rangle = \mathcal{L} |1\rangle = 0 \). Thus, if the stationary state is unique, it is given by
\[ |P_{\text{stat}}\rangle = \frac{1}{|\Omega|} |1\rangle. \] (6.8)

This result, known as Gibbs postulate or equal-a-priori postulate, tells us that an isolated system relaxes into a stationary equilibrium state in which all configurations are equally probable. In this situation the entropy of the system takes on its maximal value
\[ S = \ln |\Omega|. \]

As we have discussed in detail, the Second Law of thermodynamics states in addition that the average entropy of an isolated system on its way into this equilibrium state can only increase, i.e., \( \Delta S \geq 0 \). One of the purposes of this chapter is to prove and generalize this inequality.

**Systems embedded into the environment**

As already discussed in Chapter 3, a system is usually embedded into a larger system, called environment. This surrounding system could be very large, and we may think in an over-simplified scenario of the whole universe. Most importantly, this surrounding system itself is assumed to be isolated. We want to describe the whole environment in the same way as the embedded system itself, i.e., in terms of configurations and tran-
6.1 Dynamics of subsystems

A subsystem is defined by a projection \( \pi \) which maps each configuration \( c \in \Omega_{\text{tot}} \) of the total system (left) onto a particular configuration \( s \in \Omega_{\text{sys}} \) of the subsystem (right), dividing the configuration space of the total system into sectors.

**Figure 6.1**: A subsystem is defined by a projection \( \pi \) which maps each configuration \( c \in \Omega_{\text{tot}} \) of the total system (left) onto a particular configuration \( s \in \Omega_{\text{sys}} \) of the subsystem (right), dividing the configuration space of the total system into sectors.

In statistical physics the process of projecting a larger system onto a smaller one is often referred to as *coarse-graining* the configuration space, as illustrated in Fig. 6.1.

As discussed previously, this divides the configurational space of the total system into sectors denoted as

\[
\Omega_{\text{sys}} = \{ c \in \Omega_{\text{tot}} | \pi(c) = s \} =: \pi^{-1}(s).
\]

(6.10)

If the total system evolves only *within* such a sector, the configuration of the laboratory system will not change. However, it is important to realize that a sector \( \pi^{-1}(s) \) may not be fully connected within itself, even though the total system is ergodic. The sector \( \pi^{-1}(s) \) may rather decompose into several internally connected subsectors. Moving directly from one subsector to another within \( \pi^{-1}(s) \) is then impossible, it may only be possible to reach them by a detour through different sectors, which requires to temporarily change the configuration of the laboratory system. As we will see below, this is basically what happens in any type of cyclically moving engine.

**Remember**:

- **‘sys’**: *System*: The laboratory system, measurable, configurations \( s \in \Omega_{\text{sys}} \)
- **‘env’**: *Environment*: Everything else, non-measurable.
- **‘tot’**: *Total system*: System and environment together, configurations \( c \in \Omega_{\text{tot}} \)

For every configuration \( c \) of the ‘universe’ our laboratory system will be in a well-defined configuration denoted as \( s \in \Omega_{\text{sys}} \). In order to express this relationship, we have introduced a projection

\[
\pi : \Omega_{\text{tot}} \to \Omega_{\text{sys}} : c \mapsto s = \pi(c).
\]

(6.9)
Coarse-grained master equation

Under projective coarse-graining the probability distribution can be coarse-grained simply by adding all probabilities in the respective sector:

$$P_s(t) = \sum_{c \in \Omega^\text{tot}} P_c(t) = \sum_{c(\pi)} P_c(t). \quad (6.11)$$

Here we introduced the compact notation $\sum_{c(\pi)}$ of a sum running over all configurations of the total system with $\pi(c) = s$. Likewise, the probability currents between different configurations can be coarse-grained as well:

$$J_{s \rightarrow s'}(t) = \sum_{c(\pi)} \sum_{c'(\pi')} J_{c \rightarrow c'}(t) = \sum_{c(\pi)} \sum_{c'(\pi')} w_{c \rightarrow c'}. \quad (6.12)$$

If the total system (consisting of laboratory system and environment) evolve according to the master equation

$$\frac{d}{dt} P_c(t) = \sum_{c' \in \Omega^\text{tot}} \left( J_{c' \rightarrow c}(t) - J_{c \rightarrow c'}(t) \right). \quad (6.13)$$

one can easily show that the laboratory system will evolve according to the coarse-grained master equation

$$\frac{d}{dt} P_s(t) = \sum_{s' \in \Omega^\text{sys}} \left( J_{s' \rightarrow s}(t) - J_{s \rightarrow s'}(t) \right). \quad (6.14)$$

Note that although the rates of the (isolated) total system are symmetric and time-independent, the effective rates

$$w_{s \rightarrow s'}(t) = \frac{J_{s \rightarrow s'}(t)}{P_s(t)} = \frac{\sum_{c(\pi)} P_c(t) \sum_{c'(\pi')} w_{c \rightarrow c'}}{\sum_{c(\pi)} P_c(t)}. \quad (6.15)$$

are generally neither symmetric nor constant.

Detailed balance

At this point let us briefly recall the concept of detailed balance (see page 56). This condition holds whenever the system is in thermal equilibrium with the surrounding environment and if the environment itself is in equilibrium. If a total system is in thermal equilibrium, we have seen that all configurations $c \in \Omega^\text{tot}$ are equally probable, implying that the coarse-grained probabilities of the laboratory system are given by

$$p^\text{stat}_s = \sum_{c(\pi)} p^\text{stat}_c = \sum_{c(\pi)} \frac{1}{\Omega^\text{tot}} = \frac{|\Omega^\text{tot}|}{|\Omega^\text{tot}|}. \quad (6.16)$$

This means that the stationary probability of a certain configuration $s$ of the laboratory system is proportional to the number of configurations in the corresponding sector.
Likewise, the stationary probability currents in the laboratory system are given by

\[ j_{s \rightarrow s'}^{\text{stat}} = \frac{1}{\Omega_{\text{tot}}} \sum_{c(s)} \sum_{c'(s')} w_{c \rightarrow c'} . \tag{6.17} \]

Knowing that the transition rates of the total system have to be constant and symmetric, this equation implies that the probability currents have to be symmetric as well, leading directly to the condition of detailed balance:

\[ j_s = j_{s'} . \tag{6.18} \]

or, equivalently, in textbook form:

\[ p_s w_s \rightarrow s' = p_{s'} w_{s'} \rightarrow s \quad \forall s \neq s' . \tag{6.19} \]

**Alternative definition of detailed balance:** Not so known even among experts is an alternative definition of detailed balance which does not rely on the knowledge of the stationary state. Starting point is the notion of stochastic paths or stochastic trajectories which are nothing but sequences of configurations through which the system evolves. Let us now consider a cyclically closed stochastic path

\[ c_0 \rightarrow c_1 \rightarrow c_2 \rightarrow \ldots \rightarrow c_N \rightarrow c_0 \tag{6.20} \]

which goes back to its starting point. Then it is easy to show that the system obeys detailed balance if and only if the following identity holds for any closed stochastic trajectory:

\[ \prod_{i=1}^{N} \frac{w_{c_i \rightarrow c_{i+1}}}{w_{c_{i+1} \rightarrow c_i}} = 1 . \tag{6.21} \]

This criterion tells us that if the product of the transition rates along any closed cycle of transitions in forward direction is exactly equal to the product of rates in backward direction, we can conclude that the stationary state will obey a detailed balance. To prove this, we first note that the trivial identity

\[ \prod_{i=1}^{N} p_{c_i \rightarrow c_{i+1}} = 1 \tag{6.22} \]

holds for any closed cycle of transition, i.e., the stationary probabilities simply drop out along a closed cycle of transitions. Therefore, we have

\[ 1 = \prod_{i=1}^{N} j_{c_i \rightarrow c_{i+1}}^{\text{stat}} = \prod_{i=1}^{N} p_{c_i \rightarrow c_{i+1}} w_{c_{i+1} \rightarrow c_i} = \prod_{i=1}^{N} p_{c_i \rightarrow c_{i+1}} \prod_{i=1}^{N} w_{c_{i+1} \rightarrow c_i} = \prod_{i=1}^{N} \frac{w_{c_i \rightarrow c_{i+1}}}{w_{c_{i+1} \rightarrow c_i}} , \tag{6.23} \]

proving the property of detailed balance.

**Remember:** Definition of detailed balance:
A stationary state is said to obey detailed balance if all probability currents cancel pairwise.
As an alternative criterion, the product of the transition rates along a closed cycle have to be
equal in forward and backward direction.

Nonequilibrium

If the transition rates $w_{s \rightarrow s'}(t)$ do not obey the condition of detailed balance, the laboratory system is said to be out of equilibrium. Since we have derived detailed balance from the assumption that the total system has already reached its equally-distributed microcanonical stationary state, it conversely follows that a laboratory system out of equilibrium requires the total system to be nonstationary. As such, the total system has not yet reached the state of maximal entropy. Therefore, as a hallmark of nonequilibrium, we expect the entropy of the total system (and therewith the entropy of the environment) to increase. Such an increase of entropy in the environment is interpreted as an entropy production caused by the nonequilibrium dynamics of the laboratory system.

Before discussing entropy production in more detail, let us consider a simple example of nonequilibrium systems. First, we note that a system with only two configurations always obeys detailed balance by definition (exercise). Therefore, nonequilibrium requires at least three configurations.

Such a toy system with three configurations $s_1, s_2, s_3$ is shown on the right hand side. As usual, the configurations are symbolized by colored bullets while the transition rates are represented by arrows, the length indicating their magnitude. In this example the rates are cyclically biased in one direction, leading to a clockwise oriented probability current. If the jump rates in the respective direction are equal, as indicated by the arrows in the figure, this system will relax towards a stationary state $P_{s_1}^{\text{stat}} = P_{s_2}^{\text{stat}} = P_{s_3}^{\text{stat}} = \frac{1}{3}$. Nevertheless the system keeps on jumping preferably in clockwise direction, even though the probability distribution is stationary. In the literature this is known as a nonequilibrium steady state (NESS).

Remember: Stationarity vs. equilibrium: Equilibrium systems are always stationary, but conversely stationary systems are not always in equilibrium. The condition of equilibrium is stronger because it requires in addition that the probability currents between all pairs of configurations vanish. This happens if and only if the rates satisfy the condition of detailed balance.

If the system keeps on jumping preferably in clockwise direction, how does the corresponding dynamics in the environment look like? First of all, we expect each of the three configurations to correspond to a sector of configurations in $\Omega^e \subseteq \Omega^m$. Does this mean that in case of the clock model there are only three such sectors?

In order to answer this question on an intuitive basis, recall that the transition rates often isolated system have to be symmetric. Therefore, asymmetric rates in the laboratory system are only possible between sectors of different sizes. In other words, the system jumps preferentially from $s$ to $s'$ if the target sector is larger, i.e., $|\Omega^e_s| > |\Omega^e_{s'}|$. This is an example of a so-called entropic force which drags the system into the direction of increasing entropy. Obviously, this inequality cannot be satisfied with only three sectors, rather the sector size has to increase whenever the system jumps in clockwise direction.
Imagine that a fly is disturbing you while you are doing the exercises. Then you can get rid of the fly by opening the window. If there are no other flies outside, it is clear that the fly inside will eventually get out of the window and it will be unlikely to return. In some sense it seems as if there was a ‘force’ dragging the fly out of the window. Of course, this is not a mechanical force in the usual sense, it is rather a purely probabilistic effect: the fly itself performs an undirected random walk, but as there is so much more space outside (a large entropy), it is highly unlikely that the fly, once outside, will return. This apparent drag is a simple example of a so-called *entropic force*.

If the clock model completes one cycle $s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_1$ it ends up in the same configuration from where it started, but, from the perspective of the total system, in a different sector of the environment. In other words, although the system returns to its original starting point, the environment has changed significantly. This is exactly what all kinds of heat engines do: They run in a closed cycle, but they interact at the same time with the outside world, consuming energy and producing entropy. In the following let us study this phenomenon of entropy production in more detail.

### 6.2. Entropy production

As the environment seems to keep track of the winding number, it is meaningful to unfold the closed cycle in our three-state clock model into a linear chain of transitions, as sketched in Fig. 6.2. In this linear chain each bullet is associated with a well-defined sector, which increases in size as we go to the right. Nevertheless keep in mind that we still have to identify every third bullet, indicated by the same color.

How does the environment look like? The simple answer is that we don’t know. We made the assumption that it can be described as a stochastic Markov process as well, but neither the structure of the configuration space nor the network of transition rates $w_{c \rightarrow c'}$ is actually accessible to our observation. Then, if we don’t know anything about the physics of the environment, how can we quantify the entropy production?
Before addressing this question, let us again study the three-state clock model in an idealized situation which is shown schematically in Fig. 6.3. Here we simply assume that the number of states of the total system doubles whenever we jump in clockwise direction. For simplicity, we also assume that jumps within the same sector (i.e., jumps between the green bullets within a column) are forbidden, while neighboring sectors are assumed to be fully connected by microscopic transitions, as indicated by thin black lines in the figure. By definition, the corresponding rates have to be symmetric in both directions. Moreover, as the strongest assumption, let us assume that all non-vanishing transition rates in the total system are identical:

\[ w_{s \rightarrow s'} = \begin{cases} 1 & \text{if } w_{s \rightarrow s'} \neq 0 \\ 0 & \text{if } w_{s \rightarrow s'} = 0. \end{cases} \quad (6.24) \]

With these assumptions, it is clear that whenever the laboratory system hops to the right or to the left, each of the corresponding configurations in the total system is by definition equally probable. In other words, each sector equilibrates instantaneously. Therefore, the actual entropy of the respective sector is just the logarithm of the configurations the corresponding sector.

**Summary:** Idealized assumptions made about the embedding of the clock model:

- No transitions within the sectors
- Fully connected network of transitions between neighboring sectors
- All non-zero rates in the total system equal to 1 in both directions.

More specifically, let us assume that the laboratory system is in one particular configuration, let’s say in the third from left (the blue bullet marked by ’s’). Then, according to Fig. 6.3 the total system is in one of four possible configurations with equal probability, meaning that

\[ P_{c \rightarrow s}(t) = \frac{1}{4} P_{s}(t). \quad (6.25) \]

Suppose that the system wants to jump to the right, performing a transition s \( \rightarrow \) s’. In the total system, there are eight possibilities to realize this transition. If all the corre-
sponding transition rates $w_{c \to c'}$ are equal to 1, we have

$$J_{s \to s'}(t) = \sum_{c(s)} \sum_{c'(s')} P_c(t) P_{c'}(t) w_{c \to c'} = 8 P_s(t)$$  \hspace{1cm} (6.26)$$

so that $w_{s \to s'} = J_{s \to s'}(t)/P_s(t) = 8$. On the other hand, the current in opposite direction is given by

$$J_{s' \to s}(t) = \sum_{c'(s')} \sum_{c(s)} P_{c'}(t) P_c(t) w_{c' \to c} = 4 P_{s'}(t),$$  \hspace{1cm} (6.27)$$

hence $w_{s' \to s} = J_{s' \to s}(t)/P_{s'}(t) = 4$. Therefore, the rate of jumping to the right is twice as large as the rate of jumping to the left, simply because the corresponding configuration space of the total system doubles its size. Generally, the two equations given above can be written in the form

$$J_{s \to s'}(t) = \frac{|\Omega_{s'}^{tot}|}{P_s(t)} \Rightarrow w_{s \to s'} = \frac{|\Omega_{s'}^{tot}|}{|\Omega_s^{tot}|}.$$  \hspace{1cm} (6.28)$$

hence

$$w_{s \to s'} = \frac{|\Omega_{s'}^{tot}|}{|\Omega_s^{tot}|},$$  \hspace{1cm} (6.29)$$

Since all configurations in the corresponding sector are equally probable, the entropy of the total system is simply given by $H_{s}^{tot} = \ln |\Omega_{s}^{tot}|$. Therefore, we can conclude that the increase of the entropy in the total system during the jump of the laboratory system from $s$ to $s'$ is given by

$$\Delta H_{s \to s'}^{tot} = \ln \frac{|\Omega_{s'}^{tot}|}{|\Omega_{s}^{tot}|}.\hspace{1cm} (6.30)$$

Since we are conditioned on a jump from $s$ to $s'$, the laboratory system does not have any entropy. For this reason the change in to total entropy equals the change of the entropy in the environment. Using Eq. (6.29) this can be written as

$$\Delta H_{s \to s'}^{env} = \ln \frac{w_{s \to s'}(t)}{w_{s' \to s}(t)}.$$  \hspace{1cm} (6.31)$$

In the literature this formula is known as the Schnakenberg formula [2]. However, we have to keep in mind that these formulas with derived under very special assumptions, namely, a fully connected transition network with equal transition rates between the respective sectors. In a general setting the entropy production can be higher. Robert Ziener, a former master student, investigated this circumstance in his master thesis and published the results at arxiv.org/pdf/1505.07719.pdf.

There is also a very interesting story how entropy production was discovered in physical chemistry. If you are interested, have a look at Appendix E.
Figure 6.4: Schematic drawing how the entropy of an isolated system evolves in time. As shown in the middle, each transition contributes with a little $\delta$-peak, which may be positive or negative. Accumulating all peaks in a long time limit, one obtains a quasi-continuous distribution, as shown in the right panel. The second law of thermodynamics states that the first moment of this distribution is non-negative.

6.3. Fluctuation theorem

The discovery of fluctuation theorems is a very recent development. As we will see, the standard version of the fluctuation includes the second law of thermodynamics, but being an equation rather than an inequality it is much stronger.

The second law states that the entropy of an isolated system increases on average. However, this does not mean that the entropy increases monotonically, it actually fluctuates in positive and negative direction, but in such a way that the average is positive. This is shown schematically in Fig. 6.4. Accumulating infinitely many fluctuations in the histogram, one obtains a quasi-continuous distribution, as shown in the right panel of the figure. The second law of thermodynamics tells us that the average (the first moment) is non-negative, but it does not tell us anything about the shape of the curve. As we will see below, fluctuation theorems restrict the form of the curve by an equation rather than an inequality.

**Formal definition: Strong, weak and integral fluctuation theorems**

Consider a real-valued random variable $X$ and the corresponding probability density $P(X)$. A so-called strong fluctuation theorem is a relation of the following form:

$$P(X) = e^{X}P(-X).$$

(6.32)

This equation tells us that the left part of the distribution (for negative $X$) is exponentially suppressed. The relation restricts the form of the function $P(X)$ in such a way that it relates the left and the right half of the distribution. In other words, if the function for positive $X$ is known, it automatically determines the function for all negative $X < 0$.

The strong fluctuation theorem has to be distinguished from the so-called weak fluctuation theorem, which relates two different probability densities $P(X)$ and $P^t(X)$ defined on the same random variable $X$:

$$P(X) = e^{X}P^t(-X).$$

(6.33)

Obviously, the weak fluctuation theorem includes the strong one as a special case.
It can be shown easily (left as an exercise to the reader) that the weak fluctuation theorem (and therefore also the strong one) implies the equality
\[ \langle e^{-X} \rangle = 1 \]  
(6.34)
or in full form
\[ \int_a^b P(X) e^{-X} \, dX, \]  
(6.35)
where \([a, b]\) is the definition range of the distribution. Since this equation involves an integration, it is also known as integral fluctuation theorem. By using Jensen’s inequality for convex functions the integral fluctuation theorem implies that the first moment of the distribution is positive:
\[ \langle X \rangle \geq 0 \]  
(6.36)
This is the “second law” associated with the random variable \(X\). As we will see below, the conventional second law of thermodynamics is obtained by setting \(X = \Delta H^{\text{tot}}\).

**Remember:** There are three types of fluctuation theorems (FT’s):
- **Strong FTs** relate right and left half of the same probability density \(P\).
- **Weak FTs** relate two different probability densities \(P\) and \(P^\dagger\).
- **Integral FTs** tell us that the expectation value \(\langle e^{-X} \rangle = 1\).
- **Second Law:** Integral FTs imply the inequality \(\langle X \rangle \geq 0\) for the first moment.

The excitement about the fluctuation theorem stems from the fact that it provides an equation rather than an inequality and that it replaces the well-known second law after more than 150 years (the first one to formulate the second law was Rudolf Clausius in 1854, who taught from 1867-1869 at the University of Würzburg).

**Derivation of the strong fluctuation theorem for \(H^{\text{tot}}\)**

Let us now derive the fluctuation theorem for the total entropy of a stationary system. To this end we consider stochastic path
\[ \Gamma : s_0 \rightarrow s_1 \rightarrow s_2 \rightarrow \ldots \rightarrow s_{N-1} \rightarrow s_N \]  
(6.37)
starting at time 0 and ending at the time \(\tau\). Along this path the system jumps to the configuration \(s_j\) at time \(t_j\). What is the probability to find this particular stochastic path? In order to compute this probability, we first have to know the probability to find the system in the initial configuration \(s_0\), i.e., the resulting expression will be proportional to \(P_{s_0}(0)\). Furthermore, the probability to find this path will be proportional to the product of the rates corresponding to the transitions along the path. However, between the transitions the system stays in a certain configuration for some while which also introduces the statistical weight.

**Remark:** Weight of “doing nothing”
Suppose that the system is currently in the configuration \(s_j\). During an infinitesimal time span \(\Delta t \ll 1\) the probability of a transition is proportional to the sum of all the outgoing rates, i.e., \(dP = w_{s_j} \Delta t\), where \(w_{s_j} = \sum_{s_{j+1}} w_{s_{j+1} \rightarrow s_j}\) is the escape rate (see also Eq. (6.4)). The infinitesimal probability of “doing nothing” is therefore \(1 - \Delta t\). Waiting for a longer period in configurations \(s_j\), let’s say for the time \(N\Delta\), the corresponding probability of “doing
"Nothing" is simply the product \((1 - \Delta P)^N\) over all small time intervals in between. Since the exponential function can be represented in form of the limit \(e^x = \lim_{N \to \infty} (1 + x/N)^N\), it follows immediately that the probability of doing nothing between two transitions is given by

\[
P_{\text{doing nothing}} = \exp\left[-\int_{t_i}^{t_{i+1}} dt \, w_{c \to c'}(t)\right]
\]

Putting all things together, the probability to find the stochastic path \(\Gamma\) is given by

\[
P_{\Gamma} = P_{s_0}(0) \exp\left(-\sum_{i=0}^{N} \int_{t_i}^{t_{i+1}} dt \, w_{c \to c'}^\text{esc}(t)\right) \prod_{i=1}^{N} w_{s_{i-1} \to s_i}(t_i)
\]

(6.39)

where \(w_{c \to c'}^\text{esc}(t) = \sum_{c'} w_{c \to c'}(t)\) is the escape rate and where we formally defined \(t_0 = 0\) and \(t_{N+1} = \tau\).

We now consider the reverse path

\[
\Gamma^\dagger : s_N \xrightarrow{\tau-t_N} s_{N-1} \xrightarrow{\tau-t_{N-1}} s_{N-2} \to \ldots \to s_1 \xrightarrow{\tau-t_1} s_0 .
\]

(6.40)

This path runs exactly the opposite direction. Note that time is still running from 0 to \(\tau\) and that the sequence of events is just reflected during this time interval. The corresponding statistical weight is given by

\[
P_{\Gamma^\dagger} = P_{s_N}(0) \exp\left(-\sum_{i=0}^{N} \int_{\tau-t_i}^{\tau-t_{i+1}} dt \, w_{c \to c'}^\text{esc}(t)\right) \prod_{i=1}^{N} w_{s_i \to s_{i-1}}(\tau - t_i)
\]

(6.41)

Substituting \(t \to \tau - t\) in the integral this turns into

\[
P_{\Gamma^\dagger} = P_{s_N}(0) \exp\left(-\sum_{i=0}^{N} \int_{t_i}^{t_{i+1}} dt \, w_{c \to c'}^\text{esc}(\tau - t)\right) \prod_{i=1}^{N} w_{s_i \to s_{i-1}}(\tau - t_i)
\]

(6.42)

The common goal of all types of fluctuation theorems is to make some sense out of the logarithmic quotient \(\ln\left(P_{\Gamma}/P_{\Gamma^\dagger}\right)\) of the forward and the backward probability. Usually this is possible only if the terms for "doing nothing" cancel out. In the present case this requires that the rates are the same in both expressions, meaning that the rates are time-independent.

Moreover, computing the logarithmic quotient one has to find a suitable interpretation of \(\ln\left(P_{s_0}(0)/P_{s_N}(0)\right)\). This contribution reminds of the change of the system entropy \(\ln\left(P_{s_0}(0)/P_{s_N}(\tau)\right)\), the only difference being in the evaluation time of the denominator. However, if the system is stationary, this interpretation is indeed possible.

Therefore, in the following we shall make the assumption that the system

- has time-independent rates
- is in the stationary state
6.3 Fluctuation theorem

Obviously, this includes equilibrium states as well as so-called non-equilibrium steady states (NESS), as discussed before.

Under these assumptions the contributions for “doing nothing” are exactly identical in both expressions and the quotient of the two probabilities is given by

$$\frac{P_\Gamma}{P_{\Gamma^*}} = \frac{P_{S_0}}{P_{S_N}} \prod_{i=1}^{N} \frac{w_{s_{i-1} \rightarrow s_i}}{w_{s_i \rightarrow s_{i-1}}}.$$ (6.43)

Now we can easily recognize that the logarithmic quotient just gives the total entropy production:

$$\Delta H^\text{tot}_\Gamma = \ln \left( \frac{P_\Gamma}{P_{\Gamma^*}} \right) \prod_{i=1}^{N} \ln \left( \frac{w_{s_{i-1} \rightarrow s_i}}{w_{s_i \rightarrow s_{i-1}}} \right) = \Delta H^\text{sys} + \Delta H^\text{env}. $$ (6.44)

This formula allows us to compute the entropy changes for a given path $\Gamma$. Moreover, we know the probability density to find such a path. Combining these two pieces of information, we can compute the probability for observing the certain increase or decrease of the total entropy. To this end, one has to integrate over all possible paths, i.e., one has to perform a path integral $\int D\Gamma \ldots$. Although the path integral formalism is beyond the scope of these lecture notes, let us assume that it basically works like an ordinary integration. This allows us to express the probability of finding a certain entropy change as

$$P(\Delta H^\text{tot}) = \int D\Gamma \ P_\Gamma \delta \left( \Delta H^\text{tot} - \Delta H^\text{tot}_\Gamma \right)$$

$$= \int D\Gamma \ P_\Gamma \delta \left( \Delta H^\text{tot} - \ln \frac{P_\Gamma}{P_{\Gamma^*}} \right).$$ (6.45)

Since the integration can be understood as a sum over all possible paths, including the reversed ones, we can simply replace $\Gamma$ by $\Gamma^*$ in the integrand:

$$P(\Delta H^\text{tot}) = \int D\Gamma \ P_{\Gamma^*} \delta \left( \Delta H^\text{tot} - \ln \frac{P_{\Gamma^*}}{P_\Gamma} \right) = \int D\Gamma \ P_{\Gamma^*} \delta \left( \Delta H^\text{tot} + \Delta H^\text{tot}_\Gamma \right). $$ (6.46)

**Remark:**
This is basically the same as space replacing $\int_{-\infty}^{+\infty} f(x) \ dx = \int_{-\infty}^{+\infty} f(-x) \ dx$ on the real line.

Finally, we use the relation $\Delta H^\text{tot}_\Gamma = \ln \frac{P_\Gamma}{P_{\Gamma^*}}$ in order to express the backward probability $P_{\Gamma^*}$ begin in terms of the the forward probability $P_\Gamma$:

$$P(\Delta H^\text{tot}) = \int D\Gamma \ e^{-\Delta H^\text{tot}_\Gamma} P_\Gamma \delta \left( \Delta H^\text{tot} + \Delta H^\text{tot}_\Gamma \right) = e^{\Delta H^\text{tot}} P(\Delta H^\text{tot}).$$

These few lines complete the proof of the strong fluctuation theorem for the total entropy.

In thermal equilibrium the entropies of the system and the environment still fluctuate, but the condition of detailed balance ensures that the entropy changes are opposite
6.4. Heat and work

In conventional thermodynamics one distinguishes two fundamentally different kinds of energy transfer between subsystems. The first one is heat, denoted as $Q$. According to Clausius, heat flows from the warm to the cold reservoir, changing their entropy by $dS = \beta dQ$. Heat is some kind of disordered energy transfer. Although heat carries the unit of energy, it is only defined in a thermodynamical context, – you cannot find the notion of heat in any textbook on classical mechanics. The usefulness of heat is limited. Of course, heat can be used for heating, but in order to convert heat into a usable form of energy, a heat engine in combination with a cold reservoir is needed. As everyone knows, the efficiency of such a machine is limited by the formula according to Carnot.

The other fundamental form of energy transfer is work, denoted by $W$. Work may
be thought of as a usable directed form of energy, such as the mechanical motion of the piston, a torque generated by a motor or a flowing electric current. By its directed nature, work can be used to full extent in a secondary process, i.e., the maximal efficiency is always 100%.

Usually heat and work are used in the context of energy transfer, i.e., one is interested in differences $\Delta Q$ and $\Delta W$ rather than absolute values. There is much confusion about the sign of heat and energy changes. However, in recent years it became some kind of standard to define the sign of heat and work always with respect to the system itself. More specifically, heat flowing into the system is positive while heat flowing out of the system is negative. Likewise, work done on the system is positive while work performed by the system on the environment is considered to be negative.

**Distinguishing heat and work in stochastic Markov processes**

In the context of stochastic Markov processes – the cartoon of a system jumping between different configurations – there’s nothing like heat and work. In fact, there isn’t even the notion of energy. Therefore, heat and work seem to emerge only in the presence of energy exchange, e.g., in the canonical equilibrium ensemble. Thus, in order to study heat and work, let us first consider the canonical case.

In the canonical ensemble each configuration carries the stationary Boltzmann weight $p_{s}^{\text{stat}} \propto e^{-\beta E_s}$, where $E_s$ is the energy associated with the configuration of the system. Therefore, whenever the system jumps from $s \rightarrow s'$, it undergoes an energy change of $\Delta E = \ln(p_{s}^{\text{stat}} / p_{s'}^{\text{stat}})$. It should be emphasized that the corresponding probabilities have to be the stationary ones; if the system still relaxes towards the stationary state, the actual probabilities may be different. Since energy is conserved, $\Delta E$ has to come from somewhere, and clearly it comes from the environment. Since the system continuously jumps forth and back in its own configurations space, energy continues to fluctuate between the system in a random-walk-like manner. If the system is not yet in equilibrium, it may happen that there is a net flow from the system to the reservoir or vice versa, and this average flow of stochastic energy transfer is called heat.
7. Spin glasses

7.1. Glass phases

The Ising model on a square lattice is highly homogeneous in various respects. On the one hand, the underlying lattice is regular, without any distortions and defects. On the other hand, the nearest neighbors are mutually coupled with the same intensity everywhere, i.e., the coupling constant $J$ space does not depend on the position. Experimentally realized ferromagnets fulfill these conditions approximately.

The extreme opposite of such a highly regular ferromagnet is a so-called spin glass. A spin glass is a strongly disordered magnet. This disorder could be caused by frustrated interactions or by stochastic positions of the spins on the lattice. Another possibility is that ferromagnetic and anti-ferromagnetic bonds are randomly distributed on the lattice. The term “glass” comes from an analogy with the positional disorder in conventional glasses.

A conventional magnet has a high-temperature paramagnetic phase, whether local magnetization $\langle s_i \rangle$ at site $i$ vanishes, and a low-temperature ordered phase, where $\langle s_i \rangle > 0$. A spin glass has a third phase in between, termed spin glass phase, where the total magnetization

$$M = \frac{1}{N} \sum_{i=1}^{N} \langle s_i \rangle = 0 \quad (7.1)$$

Figure 7.1.: Left: Schematic cartoon of the phase structure of a homogeneous magnetic system such as the Ising model. Right: Corresponding phase structure of a spin glass.
still vanishes while the so-called Edwards-Anderson order parameter

\[ q = \frac{1}{N} \sum_{i=1}^{N} \langle s_i \rangle^2 \neq 0. \tag{7.2} \]

Illustratively stated the spins a locally magnetized in random orientations in such a way that the macroscopic field is still zero.

The hallmark of a spin glass is a very complex landscape of the free energy. For this reason a spin glass does not easily find its equilibrium state, instead it is frequently captured in intermediate metastable states. Consequently the relaxation is extremely slow.

**Edwards-Anderson model**

The simplest spin glass model is the *Edwards-Anderson model*. In this model, we have spins arranged on a regular \( d \)-dimensional lattice with nearest neighbor interactions in the same way as in the Ising model. However, in the present case, each interacting pair of spins has its individual (time independent) coupling constant \( J_{ij} \):

\[ E_s = - \sum_{<ij>} J_{ij} s_i s_j. \tag{7.3} \]

The coupling constants may take any value, even negative ones. A negative value of \( J_{ij} \) denotes an anti-ferromagnetic type of interaction between the spins while the positive value denotes a ferromagnetic one. It is important to note that the disordering the coupling constants is quenched, i.e., time-independent.

To simplify the notation, we will denote the set of all coupling constants \( \{ J_{ij} \} \) by the bold letter \( \mathbf{J} \):

\[ \mathbf{J} := \{ J_{ij} \} \tag{7.4} \]

Like the Ising model, the spin glass is assumed to be in equilibrium with a thermal heat bath. As such, it maximizes the potential \( \mathcal{V} = \ln Z \) which now depends on the temperature as well as on the set of frozen coupling constants

\[ \mathcal{V}(\beta, \mathbf{J}) = \ln Z(\beta, \mathbf{J}), \tag{7.5} \]

where

\[ Z(\beta, \mathbf{J}) = \sum_{s \in \Omega^{\mathbf{J}}} \exp \left( + \beta \sum_{<ij>} J_{ij} s_i s_j \right) \tag{7.6} \]

is the usual partition sum. From that we can compute various physical quantities of interest, e.g., the total heat capacity

\[ C(\beta, \mathbf{J}) = \frac{\partial^2 \mathcal{V}(\beta, \mathbf{J})}{\partial \beta^2}. \tag{7.7} \]

However, it is important to note that this heat capacity is only valid for one particular choice of the coupling constants. In fact, if these coupling constants are chosen ran-
domly, what is needed is a second average over the \( J_{ij} \), denoted as \( \langle \ldots \rangle_J \), giving

\[
C(\beta) = \langle C(\beta, J) \rangle_J
\]  
(7.8)

Of course, when averaging and we have to specify certain distribution over which the average is carried out. Here, one usually assumes that the coupling constants are mutually uncorrelated and distributed randomly according to a normal distribution

\[
P(J_{ij}) = \frac{1}{\sqrt{2\pi J^2}} \exp \left( -\frac{J_{ij}^2}{2J^2} \right),
\]  
(7.9)

where \( J \) controls the variance of the distribution. Another common choice is the bimodal distribution

\[
P(J_{ij}) = \frac{1}{2} \delta(J_{ij} - J_0) + \frac{1}{2} \delta(J_{ij} + J).
\]  
(7.10)

In both cases, the average \( \langle J_{ij} \rangle = 0 \) vanishes while the variance is given by \( \langle J_{ij}^2 \rangle = J^2 \).

Remark: This is a very subtle point for the understanding of spin glasses: there are actually two averages to be carried out. The first one is the average over the fluctuating spins, as described by the Boltzmann weights and the partition sum. This average is carried out before the partition sum is evaluated. The second average is carried out over the randomly quenched coupling constants \( J_{ij} \). This average is performed after evaluation of the partition sum.

For computing the average \( \langle \ldots \rangle_J \) we have to integrate over all coupling constants, i.e. we have to carry out as many integrals as there are coupling in the system. For an arbitrary quantity \( X(J) \) this average may be written as

\[
\langle X(J) \rangle_J := \left( \prod_{<ij>} \int_{-\infty}^{+\infty} dJ_{ij} P(J_{ij}) \right) X(\{J_{ij}\})
\]  
(7.11)

for which we will use the compact notation

\[
\langle X(J) \rangle_J := \int \mathcal{D}J \, X(J)
\]  
(7.12)

Replica trick

The main difficulty in evaluating the second average over the randomly quenched coupling constants lies in the circumstance that the physical quantities of interest are usually given in terms of partial derivatives of the logarithm of \( Z \). For example, the total heat capacity is given by

\[
C(\beta) = \langle \frac{\partial^2 \ln Z(\beta, J)}{\partial \beta^2} \rangle_J
\]  
(7.13)

Since the logarithm is highly nonlinear, it is impossible to commute the average with a logarithm, i.e.

\[
\langle \frac{\partial^2 \ln Z(\beta, J)}{\partial \beta^2} \rangle_J \neq \frac{\partial^2 \ln \langle Z(\beta, J) \rangle_J}{\partial \beta^2}
\]  
(7.14)
**Example:** Let us, for example, consider the average \( \langle f(E) \rangle \), where \( f \) is some function. If \( f \) is linear it commutes with the process of averaging:

\[
\langle f(E) \rangle = \sum_{s \in \Omega} p_s f(E_s) = f\left( \sum_{s \in \Omega} p_s E_s \right) = f \langle E \rangle
\]  

(7.15)

This is because the arithmetic average itself is linear. Obviously, this does not work if the function is nonlinear.

So the main problem with spin glasses is to get rid of the nonlinear logarithm in front of the partition sum. If you try to evaluate the right-hand side of Eq. (7.14) you will see that this is a highly nontrivial problem.

In fact, this problem can be solved with a very elegant mathematical technique, called *replica trick*. This mathematical technique is based on the formula

\[
\ln x = \lim_{n \to 0} \frac{x^n - 1}{n}
\]

(7.16)

**Sketch of a proof:** To prove this formula let us take the usual representation of the exponential function

\[
e^y = \lim_{n \to \infty} \left( 1 + \frac{y}{n} \right)^n
\]

Setting \( y = \ln x \) this turns into

\[
x = \lim_{n \to \infty} \left( 1 + \frac{\ln x}{n} \right)^n
\]

or

\[
\lim_{n \to \infty} \left[ x^{1/n} - \left( 1 + \frac{\ln x}{n} \right) \right] = 0.
\]

Hence

\[
\ln x = \lim_{n \to \infty} \frac{x^{1/n} - 1}{1/n} = \lim_{n \to 0} \frac{x^n - 1}{n}.
\]

Alternatively, this relationship can be proven using l’Hospital’s rule.

Applying this formula to the partition sum

\[
\ln Z = \lim_{n \to 0} \frac{Z^n - 1}{n}
\]

(7.17)

we can formally express quantities such as the total heat capacity in Eq. (7.19) by

\[
C(\beta) = \left\langle \frac{\partial^2 \ln Z(\beta, J)}{\partial \beta^2} \right\rangle_J = \lim_{n \to 0} \left\langle \frac{\partial^2 Z^n(\beta, J)}{\partial \beta^2} \right\rangle_J
\]

(7.18)

because the constant contribution drops out upon differentiation. Assuming that the derivative commutes with the average we get

\[
C(\beta) = \frac{\partial^2}{\partial \beta^2} \lim_{n \to 0} \left\langle Z^n(\beta, J) \right\rangle_J
\]

(7.19)

Therefore, we can express the heat capacity in terms of the averages \( \langle Z^n \rangle \). These averages are still nonlinear in the partition sum, but now they attain a new interpretation. The main observation is the following: if we take \( n \) independent copies of the same system, so-called *replicas*, then the partition sum of the composed system is obviously given by \( Z^n \). This means that we simply have to compute the heat capacity of \( n \) uncor-
related copies of the same system and then to take the limit \( n \to 0 \).

The good news is that this computation is in principle feasible because it boils down to calculating integral moments of a Gaussian distribution. However, the main obstacle with this method is that the number of copies \( n \) can only take integer values \( n = 1, 2, 3, \ldots \), while the formula given above requires to perform a continuous limit \( n \to 0 \). This is a serious problem, but the magic part of the replica method, at least for physicists, is to ignore this problem.

What we actually do is to compute the quantities of interest for integer values, then to guess the general formula which holds for any value of \( n \), and finally to carry out the formal limit \( n \to 0 \) as if the guessed formula was valid for any \( n \in \mathbb{R}_+ \). It is, of course, by no means clear whether such an approach is mathematically consistent, but physicists are already happy if it yields a reasonable result and leave this problem to the mathematicians.

### Computing the spin glass partition sum using the replica method

As we are going to see in the subsection, the replica trick unfolds its power only in combination with the saddle point method and another assumption called replica symmetry. To explain this tedious calculation, we proceed step-by-step. The reader should not be intimidated by the complexity of the lengthy formulas, instead it is important to see the essential steps of the derivation.

According to Eq. (7.19) we have to compute the two-fold average \( \langle Z^n \rangle_J \). To this end, we first compute the partition sum \( Z^n \) of \( n \) replicas of the system enumerated by \( a = 1, \ldots, n \) with the same given set of coupling constants \( J \) in all copies. Since the energy of \( n \) replicas is simply the sum over all individual energies, this partition sum is given by

\[
Z^n = \left( \prod_{a=1}^{n} \sum_{s(a) \in \Omega_{sys}^a} \right) \exp \left( -\beta \sum_{a=1}^{n} E_s(a) \right). \tag{7.20}
\]

Denoting the set of all configurations by

\[
s = \{ s^{(1)}, \ldots, s^{(1)} \} \in \Omega_{sys}^n.
\]

we may shortly write

\[
Z^n = \sum_{s} e^{-\beta \sum_{a} E_s(a)} \tag{7.22}
\]

Having computed \( Z^n \), the partition sum of the \( n \) copies is averaged over the Gaussian distribution \( P(J_{ij}) \) given in Eq. (7.9). Written explicitly this average is given by

\[
\left\langle Z^n (\beta, J) \right\rangle_{J} = \left( \prod_{i,j>0} \int_{-\infty}^{+\infty} dJ_{ij} P(J_{ij}) \right) \left( \prod_{a=1}^{n} \sum_{s(a) \in \Omega_{sys}^a} \right) \exp \left( -\beta \sum_{a=1}^{n} E_s(a) \right) \tag{7.23}
\]

with weighted average over all \( l_{ij} \) configuration sum.
or, using the compact notations in Eqs. (7.12) and (7.21), by

\[ \langle Z^n \rangle = \int D_J \sum_s e^{-\beta \sum E_s(a)} . \] (7.24)

Since both the average over the coupling constants and the sum of all configurations consists only of summations and integrations, we can (disregarding possible mathematical subtleties) exchange the configurational sum and the average:

\[ \langle Z^n \rangle_J = \sum_s \int D_J e^{-\beta \sum E_s(a)} . \] (7.25)

Since the product in front of the integral runs over all pairs of coupled sites, we do not need a separate sum over coupled sites in the argument of exponential, rather we can simply rewrite the expression as

\[ \langle Z^n(\beta, J) \rangle_J = \sum_s \prod_{<ij>} \int_{-\infty}^{+\infty} dJ_{ij} P(J_{ij}) \exp \left( \beta \sum_{a=1}^n \sum_{<ij>} J_{ij} s_i^{(a)} s_j^{(a)} \right) \] (7.26)

\[ = \sum_s \prod_{<ij>} \int_{-\infty}^{+\infty} dJ_{ij} P(J_{ij}) \prod_{<ij>} \exp \left( \beta \sum_{a=1}^n J_{ij} s_i^{(a)} s_j^{(a)} \right) \] (7.27)

\[ = \sum_s \prod_{<ij>} \int_{-\infty}^{+\infty} dJ_{ij} P(J_{ij}) \exp \left( \beta \sum_{a=1}^n J_{ij} s_i^{(a)} s_j^{(a)} \right) \] (7.28)

Now the coupling constants \( J_{ij} \) stands in front of the sum in the argument of the exponential function, allowing us to carry out the integration over the Gaussian distribution in Eq. (7.9). Using the well-known formula

\[ \int_{-\infty}^{+\infty} dJ_{ij} P(J_{ij}) e^{J_{ij}^2 \lambda^2} = \exp \left( \frac{J^2 \lambda^2}{2} \right) \] (7.29)

we arrive at the expression

\[ \langle Z^n(\beta, J) \rangle_J = \sum_s \exp \left( \frac{\beta^2}{2} \sum_{<ij>} \sum_{a,b=1}^n s_i^{(a)} s_j^{(a)} s_i^{(b)} s_j^{(b)} \right) \] (7.30)

which is no quartic in the spin variables. Experience tells us that exponentials with quartic arguments are not easy to integrate.

**Mean field approximation: Sherrington-Kirkpatrick model:**

In order to proceed, we consider the mean field limit of the Edwards-Anderson model, which is known as the *Sherrington-Kirkpatrick model* in the literature. The idea is that not only nearest neighbors, but all possible pairs of sites are mutually coupled. This means
that we can replace the sum $\sum_{<ij>}$ by a sum $\sum_i \sum_j$ with independently running indices. However, this means that the number of couplings scales with $N^2$ instead of $N$, where $N$ is the total number of sites. Consequently, the integrated coupling strength in such a mean field model would be much larger than in the original one. In order to compensate this different type of scaling, the coupling constants have to be rescaled as

$$J_{ij} \rightarrow J_{ij}/\sqrt{N}. \quad (7.31)$$

As a result, we obtain

$$\langle Z^n(\beta,J) \rangle_J = \sum_s \exp \left( \frac{\beta^2 J^2}{2N} \sum_{a,b=1}^n \left[ \sum_i s_i^{(a)} \cdot s_i^{(b)} \right] \left[ \sum_j s_j^{(a)} \cdot s_j^{(b)} \right] \right). \quad (7.32)$$

Let us now reorder the sum over the replica indices by taking out the diagonal part and letting the sum run over $a < b$, which yields an additional constant contribution in front of the sum:

$$\langle Z^n(\beta,J) \rangle_J = \sum_s \exp \left( \frac{\beta^2 J^2 N}{2} \sum_{a,b=1}^n \left[ \sum_i s_i^{(a)} \cdot s_i^{(b)} \right] \left[ \sum_j s_j^{(a)} \cdot s_j^{(b)} \right] \right) + \text{const}. \quad (7.33)$$

This constant contribution can be pulled out as a prefactor in front of the configurational sum:

$$\langle Z^n(\beta,J) \rangle_J = e^{\frac{\beta^2 J^2 N}{2} \sum_{a,b} \left[ \sum_i s_i^{(a)} \cdot s_i^{(b)} \right] \left[ \sum_j s_j^{(a)} \cdot s_j^{(b)} \right]} \sum_s \exp \left( \frac{\beta^2 J^2}{2N} \sum_{a,b=1}^n \left[ \sum_i s_i^{(a)} \cdot s_i^{(b)} \right] \left[ \sum_j s_j^{(a)} \cdot s_j^{(b)} \right] \right). \quad (7.34)$$

Obviously, this expression is still quartic in the spin variables, and the goal would be to reduce it to the quadratic expression. To achieve this goal, one uses the so-called Hubbard-Stratanovich identity, which is nothing but the linearly shifted Gaussian integration:

$$e^{\frac{\lambda r^2}{2}} = \left( \frac{\lambda}{2\pi} \right)^{1/2} \int_{-\infty}^{+\infty} dq \, e^{-\lambda q^2/2 + \lambda r q}. \quad (7.35)$$

Replacing the scalar variables $r$ and $q$ by $d$-dimensional vectors $\mathbf{r}$ and $\mathbf{q}$, the Hubbard-Stratanovich identity becomes

$$\exp \left( \frac{\lambda}{2} \mathbf{r} \cdot \mathbf{r} \right) = \left( \frac{\lambda}{2\pi} \right)^{d/2} \prod_{\mu=1}^d \int_{-\infty}^{+\infty} dq_{\mu} \exp \left( -\frac{\lambda}{2} \mathbf{q} \cdot \mathbf{q} + \lambda r \cdot \mathbf{q} \right). \quad (7.36)$$

Now let us identify these variables with those of our spin glass problem. Setting

$$\lambda = N\beta^2 J^2, \quad d = \frac{n(n-1)}{2}, \quad r_{\mu} = r_{ab} = \frac{1}{N} \left[ \sum_j s_j^{(a)} \cdot s_j^{(b)} \right] \quad (7.37)$$

and introducing the abbreviation

$$\int D_{\mathbf{q}} := \left( \frac{N\beta^2 J^2}{2\pi} \right)^{n(n-1)/4} \prod_{a<b} \int_{-\infty}^{+\infty} dq_{ab} \quad (7.38)$$
the exponential term in Eq. (7.34) turns into
\[
\exp \left( \frac{\beta^2 J^2}{2N} \sum_{a<b} \left[ \sum_i s_{i(a)}^a s_{i(b)}^a \right] \left[ \sum_j s_{j(a)}^b s_{j(b)}^b \right] \right)
\]

\[
= \int D_q \exp \left( -\frac{N\beta^2 J^2}{2} \sum_{a<b} q_{ab}^2 + \beta^2 J^2 \sum_{a<b} q_{ab} \left[ \sum_i s_{i(a)}^a s_{i(b)}^b \right] \right)
\] (7.39)

which is now quartic in the spin variables. Putting all things together and commuting
the configurational sum to the right, we arrive at
\[
\langle Z^n \rangle = \frac{e^{\frac{\beta^2 J^2 N}{2}}} \sum_s \exp \left( \frac{\beta^2 J^2}{2N} \sum_{a<b} \left[ \sum_i s_{i(a)}^a s_{i(b)}^a \right] \left[ \sum_j s_{j(a)}^b s_{j(b)}^b \right] \right)
\]

\[
= \frac{e^{\frac{\beta^2 J^2 N}{2}}} \int D_q \exp \left( -\frac{N\beta^2 J^2}{2} \sum_{a<b} q_{ab}^2 \right) \sum_s \exp \left( \beta^2 J^2 \sum_{a<b} q_{ab} \left[ \sum_i s_{i(a)}^a s_{i(b)}^b \right] \right).\] (7.40)

This expression is not only quartic in the spin variables, the spin variables also decou-
ple from each other. This allows us to recast the configurational sum as a sum over
individual spins raised to the power \( N \). The resulting expression reads
\[
\langle Z^n \rangle = \int D_q \exp \left( -\frac{N\beta^2 J^2}{2} \sum_{a<b} q_{ab}^2 \right) \left[ \prod_{\substack{a=1 \atop \mu(a) = \pm 1}}^{n} \sum_{\nu(a) = \pm 1} \exp \left( \beta^2 J^2 \sum_{a<b} q_{ab}^{u(a)} u(a)^b \right) \right]^n\] (7.41)

Introducing the function
\[
L(q,u) = \beta^2 J^2 \sum_{a<b} q_{ab}^{u(a)} u(a)^b
\]

and using the notation
\[
\sum_u := \prod_{\substack{a=1 \atop \mu(a) = \pm 1}}^{n} \sum_{\nu(a) = \pm 1} \sum_{\nu(b) = \pm 1} \ldots \sum_{\nu(n) = \pm 1}
\] (7.42)

this can be written in a more compact form as
\[
\langle Z^n \rangle = \int D_q \exp \left( -\frac{N\beta^2 J^2}{2} \sum_{a<b} q_{ab}^2 \right) + N \ln \sum_u e^{L(q,u)} \right),\] (7.43)

**Saddle point method:**

At this point it is important to note that the argument of the exponential function scales
linearly with a number of sites \( N \). Therefore, in the limit \( N \to \infty \) the integral will
be dominated by the maximum of the function inside the exponential, allowing us to
approximate the value of the integral. This is known as the so-called **saddle point method.**

**Recall:** Method of steepest descent:
Let \( f(x) \) and \( g(x) \) are smooth functions. Then the integral

\[
I = \int_{-\infty}^{+\infty} dx \, e^{-Nf(x)} g(x)
\]

(7.44)
can be estimated for large \( N \to \infty \) by

\[
I \approx e^{-Nf(x_0)} g(x_0) \sqrt{\frac{2\pi}{Nf''(x_0)}} \left( 1 + O(N^{-1/2}) \right).
\]

(7.45)

Here \( x_0 \) is the point where the function \( f(x) \) is maximal in the integration range, i.e. \( f'(x_0) = 0 \) and \( f''(x_0) < 0 \). More generally, if \( x \in \mathbb{R}^d \) is a vector and \( f(x), g(x) \) smooth functions on the vector space, then we may approximate the \( d \)-dimensional integral in the limit \( N \to \infty \) by

\[
I = \int d^d x \, e^{-Nf(x)} g(x) \approx \left( \frac{2\pi}{N} \right)^{d/2} e^{-Nf(x_0)} g(x_0) \frac{1}{||H||^{1/2}},
\]

(7.46)

where \( H_{ij} = (\partial^2 f/\partial x_i \partial x_j) \) is the Hessian matrix of the function \( f \) and \( ||H|| \) the absolute value of its determinant.

The saddle point method is in fact the most important step in the simplification of our formula. Applying it to Eq. (7.43) with

\[
f(q) = \frac{\beta^2 J^2}{2} \left( n - \sum_{a < b} q_{ab}^2 \right) + \ln \sum_u e^{L[u,q,u]}
\]

(7.47)

we find that this function is maximal at the point \( q \) where

\[
\frac{\partial f(q)}{q_{ab}} = 0 \Rightarrow q_{ab} = \frac{\prod_{a=1}^{n} \sum_{(a) = \pm 1} f^{L[1|q,u]|u^{(a)|u^{(b)}}}}{\prod_{b=1}^{n} \sum_{(b) = \pm 1} f^{L[1|q,u]|u^{(a)|u^{(b)}}}} = \langle u^{(a)|u^{(b)}} \rangle_L
\]

(7.48)

Moreover, one can show that the Hessian matrix is diagonal and that its determinant is taking the value \( ||H||^{1/2} = (\beta J)^d \), where \( d = n(n-1)/2 \). Note that this is an implicit equation since \( q \) enters also on the right hand side as an argument of \( L \).

Doing the calculation we obtain the expression

\[
\langle Z^n \rangle = \left( \frac{N\beta^2 J^2}{2\pi} \right)^{d/2} \left( \frac{2\pi}{N\beta^2 J^2} \right)^{d/2} \exp \left( \frac{N\beta^2 J^2}{2} \left( n - \sum_{a < b} q_{ab}^2 \right) + N \ln \sum_u e^{L[u,q,u]} \right)
\]

(7.49)

where the prefactors cancel one another. This expression is expected to become exact and the thermodynamic limit \( N \to \infty \).

As a last step, we have to calculate the free energy per lattice site

\[
\frac{F}{N} = -\frac{1}{N\beta} \lim_{n \to \infty} \frac{1}{n} \left[ \langle Z^n \rangle - 1 \right]
\]

(7.50)

giving

\[
\frac{F}{N} = -\frac{1}{\beta} \lim_{n \to \infty} \left( \frac{\beta^2 J^2}{2} \left[ 1 - \frac{1}{n} \sum_{a < b} q_{ab}^2 \right] + \frac{1}{n} \ln \sum_u e^{L[u,q,u]} \right)
\]

(7.51)
Recall that the $q_{ab}$ have to be determined from the condition
\[ \frac{\partial F}{\partial q_{ab}} = 0 \quad \forall ab. \] (7.52)

**Replica symmetry:**

The determination of the constants $q_{ab}$ is still a very hard problem. Therefore, an additional assumption is needed, which is suitable to simplify the problem. This is the postulate of replica symmetry. The idea is that all replicas are equally important, i.e. they play a symmetric role with respect to another. Therefore, the solution of the maximization problem should be symmetric as well, meaning that all variables in the extremum coincide:
\[ q_{a,b} \equiv q \] (7.53)

Of course, this is a nontrivial assumption, and in fact there are various cases where this assumption does not hold (replica symmetry breaking).

Postulating replica symmetry the free energy per lattice site reduces to
\[ \frac{F}{N} = -\frac{1}{\beta} \lim_{n \to 0} \left( -\frac{\beta^2 J^2}{4} \left[ 2 - (n - 1) q^2 \right] + \frac{1}{n} \ln \sum_u e^{L[q,u]} \right) \] (7.54)
where the function $L$ is now proportional to $q$:
\[ L[q, \{u\}] = \beta^2 q^2 \sum_{a,b} u^{(a)} u^{(b)} = \beta^2 q^2 \left( \sum_{a=1}^{n} u^{(a)} \right)^2 - n \] (7.55)

Inserting this function and taking out the constant part we get
\[ \frac{F}{N} = -\frac{1}{\beta} \lim_{n \to 0} \left( -\frac{\beta^2 J^2}{4} \left[ 2 + 2q - (n - 1) q^2 \right] + \frac{1}{n} \ln \sum_u e^{\beta^2 q^2 \left( \sum_{a=1}^{n} u^{(a)} \right)^2} \right) \] (7.56)

To get rid of the square in the exponential, we again apply Eq. (7.35), namely
\[ e^{\lambda t^2/2} = \left( \frac{\lambda}{2\pi} \right)^{1/2} \int_{-\infty}^{\infty} dv e^{-\lambda v^2/2 + \lambda vr}, \] (7.57)
setting
\[ \lambda = \beta^2 J^2 q, \quad r = \sum_{a=1}^{n} u^{(a)}. \] (7.58)
Since cosh

In this expression the denominator is just equal to 1. Because of

d integration

Now the 1/

Remark: Alternatively, starting from Eq. (7.60) we could try to first carry out the replica limit $n \to 0$ and then to determine the constant $q$. To this end, let us again consider Eq. (??),
Figure 7.2.: Generic phase diagram of the Sherrington-Kirkpatrick model.

\[
\frac{F}{N} = -\frac{1}{\beta} \lim_{n \to 0} \left( -\frac{\beta^2 J^2}{4} \left[ 2 + 2q - (n - 1)q^2 \right] + \frac{1}{n} \ln \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dz e^{-\frac{1}{2}z^2} \left[ 2 \cosh(\beta \sqrt{qz}) \right]^n \right)
\]

Since
\[
\lim_{n \to 0} \frac{1}{n} \ln \frac{1}{\sqrt{2\pi}} \int dz e^{-\frac{1}{2}z^2} = \lim_{n \to 0} \frac{1}{n} \ln \frac{1}{\sqrt{2\pi}} \int dz e^{-\frac{1}{2}z^2} (1 + n \ln x + O(n^2))
\]
\[
= \lim_{n \to 0} \frac{1}{n} \ln \left[ 1 + \frac{1}{\sqrt{2\pi}} \int dz e^{-\frac{1}{2}z^2} (n \ln x + O(n^2)) \right]
\]
\[
= \frac{1}{\sqrt{2\pi}} \int dz e^{-\frac{1}{2}z^2} \ln x
\]
we get
\[
\frac{F}{N} = -\frac{1}{\beta} \left( -\frac{\beta^2 J^2}{4} \left[ 2 + 2q + q^2 \right] + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dz e^{-\frac{1}{2}z^2} \ln \left[ 2 \cosh(\beta \sqrt{qz}) \right] \right). \quad (7.65)
\]

Solving the condition \( \partial F/\partial q = 0 \) we get
\[
-\frac{\beta^2 J^2}{2} (q + 1) = \quad (7.66)
\]

For simplicity, we have carried out this calculation without an external magnetic field \( h \). However, including such an external field is straightforward and leads essentially to slightly more general equations of the same structure. In fact, if we had included this field, we would find that the global magnetization for symmetrically distributed coupling constant vanishes at any temperature, i.e., a ferromagnetic phase does not exist as long as the \( J_{ij} \) are symmetrically distributed around zero.\(^1\)

As shown before, the free energy is
\[
xxx \quad (7.67)
\]

To investigate the properties of the system near the critical point, where the spin glass order parameter \( q \) is small, we expand the right-hand side of the free energy as
\[
xxx \quad (7.68)
\]

---

\(^1\)An even more elaborate calculation reveals that the ferromagnetic phase exists in the case of non-symmetrically distributed coupling constants.
7.1 Glass phases

The spin glass phase transition is expected to take place at the point where the second order term \( q^2 \) vanishes. Hence, we can conclude that the spin glass phase transition is located at \( T = J \), corresponding to the horizontal line in Fig. ??.

Note that the coefficient in front of \( q^2 \) is negative for \( T > J \), meaning that the paramagnetic solution \( q = 0 \) at high temperatures maximizes the free energy. Similarly, one can find that this spin class solution with \( q > 0 \) maximizes the free energy in the low-temperature phase. This pathological behavior is a consequence of the replica method in itself which can be explained as follows. As one can see from the expression for the free energy before the limit \( n \to 0 \) is carried out, the coefficient in front of \( q^2 \) changes sign at \( n = 1 \). Therefore, going from the physical situations with integer \( n \) to the unphysical limit \( n \to 0 \) we change the sign of the coefficient, causing maximization instead of minimization of the free energy.

For very low temperatures the solution derived above becomes unstable, the reason being that the assumption of replica symmetry does no longer apply in this case. Instead, one has to search for a solution where the replica symmetry is spontaneously broken. One can show that the instability of the replica-symmetric solution is also reflected in a negative entropy. However, a detailed analysis of such aspects is beyond the scope of this lecture. For further reading we refer to the large variety of existing textbooks in this field.
8. Neural networks

In this chapter we give a brief introduction to the theory of neural networks. The study of artificial neural networks is an important interdisciplinary field with the broad range of applications. In these lecture notes we are going to summarize some basic elements which are relevant in the context of equilibrium spin models discussed above. The presentation follows handwritten notes by Michael Biehl [?] and Ref. [?].

8.1. Biological background

The human brain (see Fig. 8.1) is certainly the most advanced information-processing unit existing on earth. It is in many respects superior to a computer. For example, we can easily enter room, look for an empty seat and sit down, – but none of the existing computers can accomplish such an easy task. In fact, the brain differs from a computer in many respects:

- Computers have a central processing unit (CPU) controlled by a serial sequence of commands. These commands are executed extremely fast ($\approx 10^9$ per second). Information processing is purely deterministic. If the hardware or the software are partially damaged the computer does not work anymore.

- The brain works decentralized by the co-operation of an enormous number of elementary units called neurons. There is no program which controls the information processing in detail. The timescale of information processing is less than a millisecond, i.e., rather slow compared to a computer. The low speed is compensated by a high degree of parallelization. The brain is extremely robust, i.e., in the case of partial damage it still functions properly.

Figure 8.1.: Left: Human brain. Right: Schematic view of the human brain. [Wikimedia commons]
The largest part of the human brain is the cortex. Although the functions of different parts of the cortex are highly differentiated, the cortex in itself looks similar everywhere, suggesting that the brain is running the same type of “hardware” in all of its parts. This means that there is a common concept of information processing throughout the whole brain. This is the concept of neurons, which play the role of elementary information-processing units.

**Neurons:**

A biological neuron is a cell which receives and transmits short electric pulses, called *action potentials*. The neurons are mutually coupled by so-called *synapses*. In the brain there are roughly 10,000 synapses per neuron, i.e. approximately $10^{14}$ in total. The interaction strength of the synapses is variable and can be positive (excitatory) or negative (inhibitory).

The figure on the right side shows two communicating neurons. Each neuron has a tree of *dendrites* which serve as input channels. The electric potential coming from the dendrites is continuously integrated inside the cell. When the potential surpasses a certain critical threshold, a complex biochemical process is released, leading to a significant increase of the voltage by almost hundred millivolts. This pulse then propagates along the so-called *axon*. The axon can be enormously long, reaching lengths of about 1 m. At the terminal point of the axon is connected to the dendrites of another neuron by synaptic contacts. The synaptic contact can be excitatory or inhibitory, increasing or decreasing the electric potential in the target cell.

Any kind of mental human activity, let it be thinking, feeling, loving and hating – all that is represented in the brain in the form of electric currents. This applies in particular to the phenomenon of consciousness, which is heavily debated in these days. Likewise, any mental status, for example, our knowledge, memory and our personality is encoded in terms of synaptic contacts. In particular, the process of learning manifests itself in an adaption of the synaptic contacts. Therefore, all mental and psychological elements of human life have a well-defined physical correlate in which they are represented. Since the number of neurons and synapses is finite, it follows trivially that our mental capability is finite as well.
8.1 Biological background

Synapses and learning:
The synapses are the “coupling constants” of the brain. Depending on the incoming electrical potential, they determine how much electricity is released in the target cell. There are various types of synapses which are well understood. The most important one is the chemical synapse. Chemical synapses are always directed, i.e., they have a well-defined input (the presynaptic axon) and a well-defined output (the postsynaptic dendrite). Inside the chemical synapse little vesicles are formed which contain special substances called neurotransmitters. When an action potential arrives, the vesicles move to the cell membrane, attach to the membrane and emit the neurotransmitters into the so-called synaptic cleft between the synapse and the dendrite of the target cell. The neurotransmitters in turn triggers certain ion channels of the dendrite’s cell membrane, increasing or decreasing the electric potential inside the postsynaptic cell.

The concept of synapses allow the neurons to form a complex network without joining the cell bodies, i.e., they are still functioning as separate cells, which is apparently important to keep them alive. On the other hand, the synapses allow the brain to store information and to control its overall reactivity. As mentioned above, the key property is the ability of the synapses to adapt their permeability, a property called plasticity. This adaption depends on the information crossing this synapse. Roughly speaking, we strengthen a synapse by using it frequently. Or, in other words:

“Cells which fire together wire together.”

Moreover, the use of neurotransmitters allows the brain to regulate its average activity globally by means of various chemical substances, controlling, for example, excitement, sleep, and various states of emotion. Consuming artificial drugs distorts this balance, modifying the synaptic transmission rate in a more or less specific way and affecting this delicate regulation mechanism. The synapses would immediately try to compensate these changes caused by the external drug intake, leading to habituation and drug addiction.

Neural coding:
In neurosciences and enormous progress was made thanks to modern imaging techniques, in particular, magnetic resonance imaging, which allows the activity of neurons...
to be monitored in real-time. In other words, we cannot yet see what we think but we
can at least see where we think.

MRI imaging nicely illustrates that most of the neurons are usually at rest, not firing
at all or only at a very low rate. Contrarily, if a neuron is excited, it emits. Not only one
action potential, but usually a sequence of many pulses, a so-called \textit{burst} or \textit{spike train}.

How neuronal information is actually coded is still debated. Some people argue that only the
density of spikes matters, while others propose that the actual timing of the spike sequence is rel-
evant as well. It seems that the truth is somewhere in between, depending on the specific situa-
tion. Clearly, most sensory cells encode their information only in the density of spikes. However,
the sense of hearing definitely transmits information encoded in the spike phases.

\textbf{Artificial neural networks}

Artificial neural networks are simplified models which mimic features of biological neural networks to some extent. As usual, such models are based on a number of simplifications. For example, one usually ignores the possibility of phase coding, assuming that the information is encoded in the density of spikes. Using this assumption, the state of each neuron can be characterized by a single real-valued variable $s \in \mathbb{R}$ describing its actual firing rate. For simplicity, we will assume that the range is limited to the interval $s \in [-1, 1]$, where $s = 1$ stands for the maximal firing rate while $s = -1$ means absolute rest.

In addition one has to set up a suitable network architecture. Among a large variety of possibilities there are two extreme architectures, namely:

(a) \textbf{Fully connected networks}, where every neuron is coupled to every other neuron.

(b) \textbf{Feedforward-networks}, where the neurons are coupled unidirectionally, forming a hierarchical sequence of layers.

\section*{8.2. Magnetically inspired neural networks}

\textbf{Associative memory}

Traditional computers store information at certain locations in their memory. To recall information an explicit address is provided, sent to the memory, and the content at this address is retrieved. Thus, conventional computers work with content-addressable memory.

Recalling information in a traditional computer requires precise knowledge of the
memory address, and a large part of data-base management is concerned with the handling of addresses, for example, in index files.

Associative memory, in contrast, does not require an address, instead it requires a data set which looks similar to the one which we would like to find. The memorized data is not stored in a specific location, instead it is stored non-locally distributed over the whole system. The realization of an associative memory is one of the easiest and most natural tasks to be accomplished by a neural network.

To be more specific, let us assume that the data to be stored is given in the form of \( p \) different patterns \( \vec{\xi}^1, \ldots, \vec{\xi}^p \), each of them consisting of \( N \) binary values \((\xi_1^p, \ldots, \xi_N^p)\) with \( \xi_i^p = \pm 1 \). If we present a new pattern \( \vec{v} = (v_1, \ldots, v_N) \) to the associative memory, it would be the goal to recall the stored pattern \( \vec{\xi} \) which most strongly resembles the presented pattern. This means that \( \vec{v} \) and \( \vec{\xi} \) should differ in as few places as possible, i.e., the Hamming distance, defined as the mean square deviation

\[
h = \sum_{i=1}^{N} (\xi_i - v_i)^2 = ||\vec{\xi} - \vec{v}||^2 \tag{8.1}
\]

should be minimal. Since

\[
(\xi_i - v_i)^2 = \xi_i^2 - 2\xi_i v_i + v_i^2 = 2 - 2\xi_i v_i \tag{8.2}
\]

the search for the minimum of the Hamming distance is equivalent to the search for the maximum of the Euclidean scalar product

\[
\vec{v} \cdot \vec{\xi}^k = \sum_{i=1}^{N} v_i \xi_i^k \tag{8.3}
\]

The Hopfield model

According to Hopfield (1982), the task of finding the maximum can be accomplished by a dynamical evolution of the network according to the following rules. We consider a network of \( N \) neurons (spins) \( s_i = \pm 1 \). This network evolves by random sequential updates, i.e., a neuron \( i \) is randomly selected. This neuron collects electric potential from all other neurons via synaptic contacts \( J_{ij} \), that is, the incoming electric potential is

\[
h_i = \sum_j J_{ij} s_j \tag{8.4}
\]

This electric potential is then shifted linearly by a local threshold parameter \( \theta_i \), which controls the firing threshold of the neuron. Depending on this shifted potential we assign a new value

\[
s_i := \text{sign}(\sum_j J_{ij} s_j - \theta_i) \tag{8.5}
\]

to the neuron \( i \). In many cases, the situation is further simplified by setting the threshold \( \theta_i \) to zero.

Obviously, this update procedure can be used in any network architecture. In the
simplest version of the Hopfield model, the network is recursive and fully connected. By “recursive“ we mean that the set of neurons is coupled with itself. Moreover, the coupling constants are often assumed to be symmetric:

$$J_{ij} = J_{ji}$$  \hspace{1cm} (8.6)

Here it is useful (although not necessary) to set the diagonal couplings to zero.

**Hebb’s learning rule:**

The immediate problem consists in choosing these synaptic coupling constants $J_{ij}$ in dependence of the stored patterns $\vec{\xi}_1, \ldots, \vec{\xi}_p$ in such a way that the network evolves from the presented pattern $\vec{s} := \vec{v}$ into the most similar stored pattern by virtue of its own inherent dynamics.

To start with, let us consider the case of a single stored pattern $\vec{\xi}$. The configuration corresponding to this pattern remains invariant under the network dynamics (with threshold $\theta = 0$) if the local field $h_i = \sum_j J_{ij} s_j$ has the same sign as $\xi_i$. As can be shown easily, this condition is satisfied by the simple choice

$$J_{ij} = \frac{1}{N} \xi_i \bar{\xi}_j.$$  \hspace{1cm} (8.7)

In fact, for $s_i = \xi_i$ the local field is then given by

$$h_i = \sum_j J_{ij} s_j = \frac{1}{N} \sum_j \xi_i \bar{\xi}_j \bar{\xi}_j = \xi_i,$$  \hspace{1cm} (8.8)

hence the pattern $\vec{s} = \vec{\xi}$ is stable under the dynamics.

It is, of course, not sufficient to show that the stored pattern is invariant under the dynamics, in addition we have to show that the dynamics is attractive. To see this, we assume that the network starts its evolution not exactly in the memorized pattern, but in a slightly different configuration, where $n < N$ of the bits have the wrong value. Without loss of generality, we can assume that these bits are the first $n$ elements of the pattern, meaning that we start with

$$s_i := \begin{cases} -\xi_i & \text{for } i = 1, \ldots, n \\ \xi_i & \text{for } i = n + 1, \ldots, N \end{cases}$$  \hspace{1cm} (8.9)

Then the synaptic potentials are given by

$$h_i = \sum_j J_{ij} s_j = -\frac{1}{N} \sum_{j=1}^n \xi_i \bar{\xi}_j + \frac{1}{N} \sum_{j=n+1}^n \xi_i \bar{\xi}_j \bar{\xi}_j = s_i \frac{(N-n) - n}{N} = \left(1 - \frac{2n}{N}\right) s_i.$$  \hspace{1cm} (8.10)

Therefore, for $n < N/2$, the network evolves into the correct pattern after a single update of all neurons, i.e., the stored pattern acts as an attractor for the network dynamics.

We are now going to consider the general case, where $p$ patterns $\vec{\xi}_1, \ldots, \vec{\xi}_p$ are to be stored simultaneously. If all patterns are equally important, it is near at hand to gener-
alize Eq. (8.7) by a simple superposition of the form

$$J_{ij} = \frac{1}{N} \sum_{k=1}^{p} \xi_k \xi_j.$$  \hspace{1cm} (8.11)

This is what is usually referred to as Hebb’s learning rule. This rule can be understood as follows: if two spins are aligned for a majority of patterns, the synapse between them becomes excitatory, otherwise it becomes inhibitory.

Let us first verify the invariance of a presented pattern $\vec{s} := \xi^\ell$. Here the local field is given by

$$h_i = \xi_i^\ell + \frac{1}{N} \sum_{k \neq \ell} \xi_k \sum_j \xi_k^\ell \xi_j^\ell.$$  \hspace{1cm} (8.12)

As can be seen, this field is not automatically aligned with the pattern since the sum in the second term could overcompensate the first term. However, if we assume that the stored patterns are mutually uncorrelated, then, according to the laws of statistics\(^1\) for large $N$ and $p$, its value will typically be of size $\sqrt{(p - 1)/N}$. As long as $p \ll N$, meaning that the number of stored patterns is much smaller than the total number of neurons in the network, the additional terms will most likely not affect the sign of the local field so that the pattern is stable under the evolution dynamics.

Next, let us verify the stability of the fixed point. Again, let us assume that $n$ neurons start out in the wrong state. Then it is straightforward to show that

$$h_i = \left(1 - \frac{2n}{N}\right) \xi_i^\ell + \frac{1}{N} \sum_{k=1}^{p} \xi_k \xi_k^\ell = \left(1 - \frac{2n}{N}\right) \xi_i^\ell + O\left(\sqrt{(p - 1)/N}\right).$$  \hspace{1cm} (8.13)

Again, under the condition $n, p \ll N$ we still have $\text{sign}(h_i) = \xi_i^\ell$ so that the network configuration will converge to the desired pattern within a single, global update. However, if the number of stored patterns becomes comparable to the number of neurons, the second random term becomes of order one so that the patterns can no longer be recalled reliably. As one can show by a replica calculations, this undesirable case occurs when the number of stored patterns exceeds about 14% of the number of neurons. It is important to note that this result holds only for random mutually uncorrelated patterns. In the case of correlated pattern, the recognition capability is even worse. Contrarily, if the patterns happen to be orthogonal on each other (meaning that their mutual scale a product vanishes), the capacity of the network can be higher than 14%.

**Magnetic interpretation:**

As we have seen so far, the Hopfield model is a paradigmatic minimalistic model of an associative memory. It stores $p$ prototype patterns as fixed-point attractors, which are encoded in the coupling matrix $\mathbf{J}$. Loosely speaking, the coupling matrix is chosen in such a way that the network feels comfortable in configurations representing the learned pattern. The stored patterns can be specified by direct computation, as for

\(^1\)This is the main consequence of the central limit theorem.
example, by Hebb’s learning rule, or, as will be discussed below, by a dynamic update scheme.

The neurons are updated asynchronously (\(=\)random-sequentially) according to the dynamical rule

\[
s_i := \text{sign} \left( \sum_j I_{ij}s_j - \theta_i \right). \tag{8.14}
\]

After learning \(p\) prototype patterns \(\xi_1, \xi_2, \ldots, \xi_p\), the network may be used to recognize similar patterns associatively. To this end a similar-looking pattern is applied to the network and the resulting output state is then recursively fed back until the network stabilizes.

The Hopfield model can be interpreted as some kind of spin glass with the energy functional

\[
E_s = -\sum_{i \neq j} J_{ij}s_is_j - \sum_i \theta_is_i, \tag{8.15}
\]

where the sum runs over all connected neurons. Writing this energy functional again as

\[
E_s = -\sum_i h_is_i, \quad h_i = \sum_{j \neq i} I_{ij}s_j - \theta_i \tag{8.16}
\]

we can apply the heat bath algorithm

\[
s_i^{\text{new}} = \text{sign}(z - p_i), \tag{8.17}
\]

where \(z \in [0, 1]\) is a random numbers drawn from a flat distribution and

\[
p_i = \frac{e^{\beta h_i}}{e^{\beta h_i} + e^{-\beta h_i}}. \tag{8.18}
\]

Thus, the update rule in Eq. (8.14) is just the zero temperature limit of heat bath dynamics.

**Energy in the attractor basins:**

If the coupling constants \(I_{ij}\) are determined according to Hebb’s learning rule, and if the configuration of the system corresponds to one of the stored patterns \(\xi^\ell\), then the energy functional takes the value

\[
E_{\xi^\ell} = -\frac{1}{N} \sum_{i,j=1}^N \sum_{k=1}^p \xi_i^\ell \xi_k \xi_j^\ell \xi_j = -\frac{1}{N} \left[ N^2 + \sum_{k \neq \ell} \left( \sum_{i=1}^N \xi_i^\ell \xi_i^k \right)^2 \right]. \tag{8.19}
\]

Assuming that the patterns are uncorrelated and that the values \(\xi_i^k = \pm 1\) are equally likely, one finds that the term \(\left( \sum_{i=1}^N \xi_i^\ell \xi_i^k \right)^2\) in the bracket is of the order \(\sqrt{N}\). Therefore, the energy becomes

\[
E_{\xi^\ell} \approx -\frac{1}{2N} \left[ N^2 + \mathcal{O}((p-1)N) \right] \approx -\frac{N}{2} + \mathcal{O}((p-1)/2). \tag{8.20}
\]
Thus, in the fixed-point of the stored pattern, the energy is basically equal to $-E/2$. The influence of all other patterns causes a slight shift of the total energy, since it is typically caused by fluctuations around the ground state of the system.

Let us begin investigate what happens if the first $n$ bits of the pattern are faulty. In this case, the fluctuating terms replaced by another fluctuating term, which does not result in an essential modification of its intensity. However, the term with $k = \ell$ changes substantially: if the first $n$ bits are incorrectly oriented ($s_i = -\xi_i$, $i = 1, \ldots, n$), its value changes to

$$
\sum_{i,j=1}^{N} \xi_i \xi_j s_i s_j = \left( \sum_{i=1}^{N} \xi_i s_i \right)^2 = (N - 2n)^2
$$

(8.21)

that is, the energy of the configuration rises in leading order proportional to the extent of the deviation from the stored pattern:

$$
E_s \approx E_\xi - 2n^2 N.
$$

(8.22)

This illustrates that the memorized patterns are local minima of the energy landscape.

A more detailed analysis shows that the energy functional has, in addition, infinitely many other local minima. However, all these additional minima are less pronounced than those given by the stored patterns. Therefore, the stored patterns can be interpreted as the global minima of the energy surface $E_s$, at least for moderate values of the storage capacity $p/N$.

**Parallel versus sequential dynamics**

For vanishing temperature $T = 0$, a thermodynamic system in contact with a heat bath minimizes its energy. The update rule of the neural networks has to be designed in such a way that it performs this minimization as time proceeds. In fact, it is easy to see that the energy functional Eq. (8.14) decreases continually with time if random sequential updates are used. To see this, let us consider the energy contribution of neuron $i$ under the assumption of symmetric coupling constants:

$$
E_i(t) = -s_i(t) \left[ \sum_{j \neq i}^{N} J_{ij} s_j(t) \right] = -s_i(t) h_i(t).
$$

(8.23)

If this neuron is updated random-sequentially the energy changes to

$$
E_i(t+1) = -s_i(t+1) \left[ \sum_{j \neq i}^{N} J_{ij} s_j(t) \right] = -\text{sign}[H - I(t)] h_i(t)
$$

$$
= -|h_i(t)| \leq -s_i(t) h_i(t) = E_i(t).
$$

(8.24)

In other words, the energy contribution of a single neuron subjected to random sequential updates never increases with time. Since the energy is bounded from below, this implies that the network dynamics must reach a stationary state corresponding to a minimum of energy functional.
This does not apply to parallel updates (synchronous dynamics). Since here all neurons assumed new states in parallel at the same moment, the contribution of an individual neuron to the energy functional cannot be considered in isolation. It turns out that algorithms with parallel updates are also suitable to minimize the energy, but they might end up in a state with the same configuration of the entire network is repeated every second time step.

### 8.3. Hierarchical networks

The Hopfield networks we have discussed so far may be defined on an arbitrary lattice, e.g., a square lattice with nearest neighbor couplings or a fully connected network. The high connectivity requires a certain update scheme, e.g., random sequential updates. Otherwise, if all neurons would react instantaneously, the behavior of such a network would be uncontrolled.

In hierarchical networks, the situation is different. These networks are built in such a way that closed feedback loops of information flow are forbidden. This allows one to update all neurons instantaneously, defining its action as a map rather than a dynamical procedure. The most important class of models with this property is that of so-called feedforward networks, which will be discussed in the following.

#### Feedforward networks

Feedforward networks are characterized by a layered structure and directed information processing. Usually a feedforward network consists of an input layer, one or several intermediate hidden layers, and an output layer. In the following let us enumerate the layers by an index $\ell$.

Since closed feedback loops are forbidden, neurons within the same layer are not mutually coupled. For this reason, the state of neuron $i$ in the layer $\ell$ is fully and exclusively determined by the configuration of the preceding layer. For example, using the update rule introduced above, the state of the neuron $i$ in layer $\ell$ could be given by

$$ s_i^{(\ell)} = \text{sign} \left( \sum_{j=1}^{N^{(\ell-1)}} J_{ij}^{(\ell)} s_j^{(\ell-1)} - \theta_i^{(\ell)} \right). \quad (8.25) $$

Thus, for a given coupling constants $J_{ij}^{(\ell)}$ and firing thresholds $\theta_i^{(\ell)}$ a feedforward network establishes a direct functional relationship between input and output. The information processing within the network is instantaneous since the definition of the network does not involve a specific dynamics. Since the output layer renders a binary string of neuron states $s = \pm 1$, the feedforward network is performing a binary classification of the input data. This classification, however, depends on the specific choice of the coupling constants.
The perceptron:
The perceptron is a minimal model of a neuron with $N$ input channels (dendrites), each of them reading a real number. Thus, the input data can be understood as a real-valued vector $\vec{\xi} \in \mathbb{R}^N$. Each input channel is weighted with an individual coupling constant, which can also be thought of as forming a vector $\vec{J} \in \mathbb{R}^N$.

The output of the perceptron (the axon) is a single real number. In the simplest version of the perceptron, this output is binary, taking the values $s = \pm 1$. In this case the output value is given by the relation

$$s = \text{sign} \left( \sum_j J_j \xi_j - \theta \right) = \text{sign} (\vec{J} \cdot \vec{\xi} - \theta) . \quad (8.26)$$

To understand how a binary perceptron works, let us consider the following geometric interpretation. The vector of coupling constants $\vec{J}$ defines a hyperplane

$$\{ \vec{\xi} \in \mathbb{R}^N | \vec{J} \cdot \vec{\xi} = \theta \} \quad (8.27)$$

in the $N$-dimensional space. For $\theta = 0$ this hyperplane runs through the origin, while for $\theta > 0$ it is shifted away from it, forming an affine subspace. Since the vector $\vec{J}$ is oriented perpendicular to the hyperplane, a positive value of $\vec{J} \cdot \vec{\xi} - \theta$ indicates that the point $\vec{\xi}$ lies above the hyperplane, while a negative value indicates a point below. Therefore, the hyperplane separates the vector space linearly into two parts. If the point $\vec{\xi}$ is located on one side of the hyperplane, the perceptron responds with a $+1$, otherwise it responds with $-1$. In other words, the perceptron can classify input data with respect to a hyperplane determined by $\vec{J}$ and $\theta$.

Classifications:
Let us consider $p$ sets of input data, called patterns or examples, enumerated by $\nu = 1, \ldots, p$. These patterns are represented by $p$ vectors $\vec{\xi}_v \in \mathbb{R}^N$. The goal is to design a perceptron in such a way that it responds to each of the input patterns $\vec{\xi}_v$ with a well-defined given answer $s^v = \pm 1$.

A map from $p$ input patterns onto $p$ binary outputs is called a classification. In the following we denote a classification by the symbol $C^p_N$:

$$C^p_N = \{ \vec{\xi}_v \in \mathbb{R}^N \mapsto s^v = \pm 1 \}_{v=1,\ldots,p} \quad (8.28)$$

A classification can be thought of as a set of $p$ points in a $N$-dimensional vector space, which are marked by two different colors, say red and green.

A classification is said to be linearly separable if the red and the green dots can be separated by a hyperplane. Obviously not all classifications are linearly separable. On the other hand, every linearly separable classification can be modeled by a perceptron. Therefore, two important questions arise:

(a) Is a given classification $C^p_N$ linearly separable?

(b) If so, how can we find the corresponding coupling constants $\vec{J}$ and the shift $\theta$?
The perceptron storage problem:

To answer the questions posed above we have to understand how generic or special linear separability is. To this end, we would like to compute the number $\Omega(p, n)$ of linearly separable classifications of $p$ patterns in $N$ dimensions. Equivalently, we could ask for the probability that a randomly selected classification is linearly separable.

For a given classification $\Omega$ one may find a whole variety of hyperplanes, as illustrated in Fig. ?? . In order to compute the number $\Omega(p, n)$ let us now add another pattern in this situation, going from $p$ to $p + 1$. Doing so we have to distinguish two cases:

(a) If the new pattern is represented by a vector $\vec{\xi}(p+1)$ which does not touch any of the hyperplanes, all the realizations of the classification (in terms of hyperplanes) can be associated with the same output. This contribution to $\Omega(p + 1, n)$ is denoted as $\Omega_1(p, n)$.

(b) If the new pattern touches one of the hyperplanes the multitude of hyperplanes can be split into two parts, either with the new point on the one side or on the other. In this case, the classification is ambiguous. If $\Omega_2(p, n)$ denotes the number of such ambiguous classifications, the contribution to $\Omega(p + 1, n)$ will be given by $2\Omega_2(p, n)$.

\[ \Rightarrow \Omega(p + 1, N) = \Omega_1(p, N) + 2\Omega_2(p, N) = \Omega(p, N) + \Omega_2(p, N). \quad (8.29) \]

will be continued
A. Appendices

A. Perron-Frobenius Theorem

Here we briefly summarize the Perron-Frobenius theorem and discuss its consequences with respect to stochastic evolution operators.

Let $A = (a_{ij})$ be an $n \times n$ positive matrix: $a_{ij} > 0$ for $1 \leq i, j \leq n$. Then the following statements hold:

- There is a positive real number $r$, called the “Perron root” or the “Perron–Frobenius eigenvalue”, such that $r$ is an eigenvalue of $A$ and the absolute value of any other eigenvalue $\lambda$ (which may be complex) is strictly smaller than $r$. Thus, the spectral radius $\rho(A)$ is equal to $r$.

- The eigenspace associated to $r$ is one-dimensional, i.e. the Perron root is non-degenerate.

- There exists a right eigenvector $\vec{v} = (v_1, v_2, \ldots, v_N)$ of $A$ with eigenvalue $r$ such that all components of $\vec{v}$ are positive. Similarly, there exists a left eigenvector with the same property.

- All other eigenvectors must have at least one negative or non-real component.

- $\lim_{k \to \infty} A^k$ is a projection operator on $\vec{v}$, i.e., in a high power of $A$ the eigenvector $\vec{v}$ corresponding to the eigenvalues $r$ is the only one to survive asymptotically.

- The row sums of the matrix $A$ provide bounds on $r$, i.e.

$$\min_i \sum_j a_{ij} \leq r \leq \max_i \sum_j a_{ij}. \quad (A.1)$$

The eigenvector $\vec{v}$ to the eigenvalue $r$ can be normalized in such a way that the (non-negative) exponents add up to 1. Therefore, the components of such a normalized vectors can be interpreted as probabilities.

Application to Liouville operators  Let us now apply the Perron-Frobenius theorem to the Liouville operator $\mathcal{L}$. According to Eq. (2.7) its matrix elements are given by

$$\mathcal{L}_{cc'} = -w_{c \to c'} + \delta_{c,c'} \sum_{c''} \underbrace{w_{c \to c''}}_{\varepsilon_{c'}}. \quad (A.2)$$

© Hay Hinrichsen — Physics of Complex Systems — CC BY-SA 4.0
The sums $E_c$ in the last term are the so-called escape rates and they are listed as positive entries along the diagonal. The individual transition rates, on the other hand, appear as negative numbers or zeros in the off-diagonal elements. To establish a connection with the Perron-Frobenius theorem, we invert the sign of all entries and add a multiple of a unit matrix, i.e. we set
\[ A = s\mathbb{1} - \mathcal{L}, \]  
where the prefactor is large enough so that all matrix elements become non-negative. In this way we obtain a real-valued (generally non-symmetric) matrix with non-negative entries.

Let us for now assume that all entries are positive, meaning that there are no vanishing rates so that the transition network is fully connected. Then $A$ has an upper Perron-root $r$, hence $\mathcal{L}$ has a lower Perron-root, i.e. there exists a real eigenvalue such that the real part of all other eigenvalues is strictly larger. The corresponding eigenvector is the only one to have non-negative entries.

Now we use the fact that $\mathcal{L}$ is constructed in such a way that it preserves probability. As shown above, this implies that $\langle 1 | \mathcal{L} = 0$, where $\langle 1 | = (1, 1, \ldots)$. Clearly, this is a left eigenvector to the eigenvalue zero, and since we know that all other eigenvectors have at least one negative component, this will be the eigenvector corresponding to the Perron root. Consequently we know that the root eigenvalue of the Liouvillian is zero, implying $s = r$.

The corresponding right eigenvector has generally different components, but we know that they are also non-negative and can be normalized to add up to 1, hence allowing a stochastic interpretation. As argued in the Chapter 1, this is the stationary state of the system. Moreover, in the case of a fully connected transition network, we can also conclude that the stationary state is non-degenerate and unique. Since the real part of all other eigenvalues is positive, they describe relaxational modes which tend to zero in the long-time limit $t \to \infty$.

Partially connected transition networks may decompose into disconnected clusters of transitions. In this case the above considerations can be applied to each of these clusters. Consequently, such a system has several stationary states, namely, as many as disconnected clusters.

### B. Tensor products

#### Definition

Let $V_1$ and $V_2$ be two vector spaces. These vector spaces can be used to create a new vector space $V_1 \otimes V_2$, called the tensor product of $V_1$ and $V_2$. This space can be constructed most easily by defining a suitable basis. To this end let $\{|e_i\rangle\}$ and $\{|f_j\rangle\}$ be a basis of $V_1$ and $V_2$, respectively. Then the basis of $V_1 \otimes V_2$ can be thought of as the set of all possible ordered combinations of these basis vectors. For example, if $\{|e_1\rangle, |e_2\rangle, |e_3\rangle\}$ is a basis of $V_1 = \mathbb{R}^3$ and if $|f_1\rangle, |f_2\rangle$ is a basis of $V_2 = \mathbb{R}^2$, then the basis of the tensor product
given by the six vectors consisting of all possible combinations
\[
\{ |e_1\rangle|f_1\rangle, |e_1\rangle|f_2\rangle, |e_2\rangle|f_1\rangle, |e_2\rangle|f_1\rangle, |e_3\rangle|f_1\rangle, |e_3\rangle|f_2\rangle \},
\]
(A.4)
where we used the compact notation $|e_i\rangle|f_j\rangle := |e_i\rangle \otimes |f_j\rangle$. Therefore, the dimension of a tensor space is the product of the dimensions of its tensor factors, i.e. the dimension is multiplicative with respect to $\otimes$.

Usually the combinations of the basis vectors are sorted lexicographically, meaning that the index of the rightmost tensor factor is the first to be incremented.

Having defined a basis in $V_1 \otimes V_2$ we can use bilinearity to derive the tensor product of arbitrary vectors. Bilinearity means that scalar prefactors in both tensor components can be pulled out in front of the tensor product, i.e.
\[
\left( \lambda |e_i\rangle \right) \otimes \left( \mu |f_j\rangle \right) = \lambda \mu \left( |e_i\rangle \otimes |f_j\rangle \right), \quad \forall \lambda, \mu \in \mathbb{R}
\]
(A.5)

With this rule it is straight-forward to show that the components of the tensor product of two vectors $|a\rangle \otimes |b\rangle$ are given by
\[
|c\rangle = |a\rangle \otimes |b\rangle = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \otimes \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 b_1 \\ \lambda_1 b_2 \\ \lambda_2 b_1 \\ \lambda_2 b_2 \\ \lambda_3 b_1 \\ \lambda_3 b_2 \end{pmatrix}.
\]
(A.6)

Note that the tensor product is an ordered product, i.e. the tensor factors do not commute with each other.

Likewise it is straight-forward to show that the tensor product of two linear operators $A$ and $B$ is given by
\[
C = A \otimes B = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11} b_{11} & a_{11} b_{12} & a_{12} b_{11} & a_{12} b_{12} & a_{13} b_{11} & a_{13} b_{12} \\ a_{11} b_{21} & a_{11} b_{22} & a_{12} b_{21} & a_{12} b_{22} & a_{13} b_{21} & a_{13} b_{22} \\ a_{21} b_{11} & a_{21} b_{12} & a_{22} b_{11} & a_{22} b_{12} & a_{23} b_{11} & a_{23} b_{12} \\ a_{21} b_{21} & a_{21} b_{22} & a_{22} b_{21} & a_{22} b_{22} & a_{23} b_{21} & a_{23} b_{22} \\ a_{31} b_{11} & a_{31} b_{12} & a_{32} b_{11} & a_{32} b_{12} & a_{33} b_{11} & a_{33} b_{12} \\ a_{31} b_{21} & a_{31} b_{22} & a_{32} b_{21} & a_{32} b_{22} & a_{33} b_{21} & a_{33} b_{22} \end{pmatrix}.
\]
(A.7)

Note that a tensor product has no direct geometric interpretation because the smallest vector space which can be decomposed into two tensor factors is $\mathbb{R}^4$. In particular, the tensor product must not be confused with the direct sum, where the vector components

\[\footnote{In the mathematical literature the tensor product of linear operators is referred to as the Kronecker product. The same applies to Mathematica\textsuperscript{\textregistered} : Here you should use KroneckerProduct instead of TensorProduct. Both commands yield the same matrix entries but the resulting list structure is different.}
are simply concatenated.

**Formal properties of tensor products**

Let us now summarize the most important formal properties of the tensor product. At first we can always pull out scalars $\mu, \nu \in \mathbb{R}$ as ordinary products in front of the tensor product, i.e.,

$$
\left( \mu |a\rangle \otimes |b\rangle \right) = \mu \langle a | \otimes | b \rangle,
\left( \mu A \otimes \nu B \right) = \mu \nu \left( A \otimes B \right).
$$

(A.8)

In the case of complex vector spaces it is important to note that tensor products differ from scalar products in so far as the scalar belonging to the left factor must not be complex conjugated.

If a tensor product of operators is applied to a tensor product of vectors, each operator acts on each vector separately, i.e.

$$
(A \otimes B)(|a\rangle \otimes |b\rangle) = A|a\rangle \otimes B|b\rangle.
$$

(A.9)

Consequently, the successive application of several factorizing operators can be carried out separately in every tensor component:

$$
(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1A_2) \otimes (B_1B_2).
$$

(A.10)

Performing the adjoint ‘†’ one has to conjugate all tensor components individually, keeping the order of the tensor factors:

$$
(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger,
\langle a | \otimes | b \rangle)^\dagger = \langle a | \otimes \langle b |.
$$

(A.11) (A.12)

This differs from ordinary operator products $(AB)^\dagger = B^\dagger A^\dagger$, where the order is reversed. The same applies to operator products within the tensor components, which are reversed as well, since in this case, the usual rules for matrix products apply:

$$
\left( (A_1A_2) \otimes (B_1B_2B_3) \right)^\dagger = (A_2^\dagger A_1^\dagger) \otimes (B_3^\dagger B_2^\dagger B_1^\dagger).
$$

(A.13)

The tensor products of two scalars $\lambda, \mu \in \mathbb{C}$ is defined in a formally consistent manner as an ordinary multiplication in $\mathbb{C}$:

$$
\lambda \otimes \mu \equiv \lambda \mu.
$$

(A.14)

For example we have

$$
\left( \langle a_1 | \otimes \langle b_1 | \right) \left( A \otimes B \right) \left( |a_2\rangle \otimes |b_2\rangle \right) = \langle a_1 |A|a_2\rangle \langle b_1 |B|b_2\rangle \in \mathbb{C}.
$$

(A.15)

In particular, the norm of a tensor product is simply given by the product of the norms of its factors:

$$
| |a \rangle \otimes | b \rangle | = ||a \rangle ||b \rangle |.
$$

(A.16)
C. Properties of the Shannon entropy

Invariance under reordering

The Shannon entropy

\[ S = - \sum_{c \in \Omega} P_c \ln P_c \]  

(A.21)

is a sum over a continuously differentiable function \( P_c \ln P_c \) of the probabilities \( P_c \). Since the summands can be commuted freely, it is clear that the Shannon entropy is invariant under re-ordering (permutations) of the sample space \( \Omega \).

Convexity of the individual entropy

A geometrical object is said to be convex if each connection line between two arbitrary points lies in the interior of the object. For example, a function \( f(x) \) is convex on the interval \((a, b)\) if every straight line between two points on the curve lies entirely above the curve. Parameterizing this straight line by a parameter \( \lambda \in [0,1] \), convexity can be expressed as

\[ \forall x_1, x_2 \in (a, b) \forall \lambda \in [0,1] : \quad f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) . \]  

(A.22)
If the curve touches this straight line only at the ending points at \( \lambda = 0, 1 \), then the function is said to be \textit{strictly convex}. Conversely, a function \( f(x) \) is called \textit{concave} (strictly concave) if \( -f(x) \) is convex (strictly convex).

Clearly, a two-fold differentiable function is (strictly) convex if and only if its second derivative is a nonnegative (positive). For the entropy we can show that it is a strictly convex function since the second derivative is always positive:

\[
\frac{\partial^2}{\partial P_c^2} H_c = \frac{\partial^2}{\partial P_c^2} \left(-\ln P_c\right) = \frac{1}{P_c^2} > 0. \tag{A.23}
\]

The convexity property is very important and allows us to derive various inequalities for information-theoretic quantities. An important tool in this context is \textit{Jensens inequality} for convex functions, as will be discussed in the following.

\[\text{Jensens inequality}\]

Suppose that we have a collection of numbers \( x_1, \ldots, x_n \) and let \( \bar{x} \) be their average, as sketched in the figure. Furthermore, let \( f(x) \) be a convex function. \textit{Jensens inequality} basically tells us that the function of the average \( f(\bar{x}) \) is always smaller than the average of the function values \( f(x) \). In fact, the claim of Jensens inequality is stronger and remains valid even if the values \( x_1, \ldots, x_n \) are weighted with non-uniform probabilities. Thus, viewing the \( x_i \) as results of a random variable \( X \), Jensens inequality states that

\[
f\left(\mathbb{E}(X)\right) \leq \mathbb{E}\left(f(X)\right). \tag{A.24}
\]

\textbf{Proof:} In order to prove this relation let us denote the possible results of the random variable by \( x_1, \ldots, x_n \) and the corresponding probabilities by \( P_1, \ldots, P_n \). Using the convexity property (A.22) one can show that

\[
f\left(\mathbb{E}(X)\right) = f\left(\sum_{i=1}^{n} P_i x_i\right) = f\left(P_1 x_1 + (1 - P_1) \frac{1}{(1 - P_1)} \sum_{i=2}^{n} P_i x_i\right) \tag{A.25}
\]

\[
\leq P_1 f(x_1) + (1 - P_1) f\left(\frac{1}{(1 - P_1)} \sum_{i=2}^{n} P_i x_i\right)
\]

\[
= P_1 f(x_1) + (1 - P_1) f\left(P_2 \frac{x_2}{1 - P_1} + (1 - P_2) \frac{1}{(1 - P_1)(1 - P_2)} \sum_{i=3}^{n} P_i x_i\right)
\]

\[
\leq P_1 f(x_1) + P_2 f(x_2) + \ldots + (1 - P_1)(1 - P_2) f\left(\frac{1}{(1 - P_1)(1 - P_2)} \sum_{i=3}^{n} P_i x_i\right)
\]

This estimation can be applied consecutively to each of the remaining summands. After
\[ f\left( E(X) \right) \leq \sum_{i=1}^{n} P_i f(x_i) = \mathbb{E}\left( f(X) \right) \quad \text{(A.26)} \]

As an immediate corollary of Jensen’s inequality, we can derive that the Shannon entropy \( H \) is bounded from above by the entropy \( H_{\text{max}} \) of the uniform probability distribution. To this end, we set \( X_c = \frac{1}{P_c} \) and \( f(x) = \ln x \), noting that the logarithm is a concave function. In this context, Jensen’s inequality takes the form

\[ \ln \mathbb{E}(X) \geq \mathbb{E}(\ln X) \quad \text{(A.27)} \]

or explicitly

\[ \ln |\Omega| = \ln \sum_c P_c \frac{1}{P_c} \geq \sum_c P_c \ln \frac{1}{P_c} = -\sum_c P_c \ln P_c \quad \text{(A.28)} \]

meaning that

\[ H \leq \ln |\Omega|. \quad \text{(A.29)} \]

This inequality is easy to understand. It tells us that the information of a system combined with partial prior knowledge about the system in the form of a probability distribution cannot be larger than the corresponding entropy without prior knowledge. In other words: Partial knowledge in the form of a probability distribution can only reduce the information content of a system.

Another important consequence: Knowing nothing means to assume a uniform probability distribution, corresponding to the maximum possible entropy \( \ln |\Omega| \).

### D. Moments and Cumulants

**Moments:**

For a given normalized probability distribution \( P_c \) and a given ‘observable’ \( X_c \), which can be thought of as a random variable, the arithmetic average is defined as

\[ \langle X \rangle = \sum_c P_c X_c. \quad \text{(A.30)} \]

Likewise, we can compute higher moments

\[ \mu_k = \langle X^k \rangle = \sum_c P_c X_c^k. \quad \text{(A.31)} \]

For example, we may study the average energy \( \langle E \rangle \) and the corresponding second moment, the average squared energy \( \langle E^2 \rangle \). However, usually we are not interested in the second moment itself, rather we would like to know the fluctuations of the energy around its mean. The intensity of these fluctuations is characterized by the variance

\[ \sigma^2 = \langle E^2 \rangle - \langle E \rangle^2, \quad \text{(A.32)} \]

where the offset of the average has already been subtracted.
Moment-generating function:

For a given sequence of numbers \( \{a_0, a_1, \ldots\} \), the associated generating function is defined as

\[
g(t) = \sum_{k=0}^{\infty} \frac{a_k t^k}{k!}. \tag{A.33}
\]

Such a generating function can be defined for any sequence, but when dealing with generating functions, the actual radius of convergence has to be carefully taken into account. Moreover, there are several variants of generating functions, depending on whether the factorial in the denominator is included or not, for a survey have a look at Wikipedia.

If the series \( g(t) \) converges, the numbers \( a_k \) can be retrieved by computing the \( k \)th derivative at \( t = 0 \):

\[
a_k = \frac{\partial^k g(t)}{\partial t^k} \bigg|_{t=0}. \tag{A.34}
\]

Note that \( t \) is set to zero only after taking the \( k \)th derivative.

In the present case of a random variable \( X : c \mapsto X_c \), it is useful to encode all moments \( m_0, m_1, m_2, \ldots \) in terms of a moment-generating function by defining:

\[
M_X(t) = \sum_{k=0}^{\infty} \frac{\langle X^k_c \rangle t^k}{k!} = \sum_{k=0}^{\infty} \frac{X^k_c t^k}{k!} = \langle e^{X_c t} \rangle, \tag{A.35}
\]

where, as before, the angle bracket \( \langle \ldots \rangle \) denote the average over all configurations \( c \in \Omega \) weighted by \( P_c \). In the discrete and continuous case we therefore have

\[
M_X(t) = \sum_c P_c e^{X_c t}, \quad M_X(t) = \int p(z) e^{X(z) t} \, dz \tag{A.36}
\]

In both cases the moments can be retrieved by taking the derivative \( \mu_k = \frac{\partial^k M_X(t)}{\partial t^k} \bigg|_{t=0} \).

The power of moment-generating function is related that for many standard distributions it can be expressed in a simple closed form. For example, for the discrete binomial distribution \( B(n, p) = \frac{n! p^k (1-p)^{n-k}}{k!(n-k)!} \) the corresponding moment-generating function is simply given by \( M_X(t) = (1 - p + pe^t)^n \).

Cumulants:

A systematic mathematical framework, in which the variance appears as a special case, is the concept of cumulants. Roughly speaking, cumulants are like moments, but much better. The easiest way to start with cumulants is to define their generating function which is just the logarithm of the moment-generating function:

\[
K_X(t) := \ln M_X(t). \tag{A.37}
\]
The cumulants can then be calculated by taking the \( k \)th derivative
\[
\kappa_k = \left. \frac{\partial^k K_X(t)}{\partial t^k} \right|_{t=0} .
\] (A.38)

Alternatively, they can also be derived from the moments \( \mu_n \) by the recursion relation
\[
\kappa_n = \mu_n - \sum_{m=1}^{n-1} \binom{n-1}{m-1} \kappa_m \mu_{n-m} .
\] (A.39)

Thus, the cumulants can be expressed as linear combinations of the moments. The first few are given by
\[
\begin{align*}
\kappa_1 &= \mu_1 \\
\kappa_2 &= \mu_2 - \mu_1^2 \\
\kappa_3 &= \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 \\
\kappa_4 &= \mu_4 - 4\mu_3\mu_1 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^4 \\
\kappa_5 &= \mu_5 - 5\mu_4\mu_1 - 10\mu_3\mu_2 + 20\mu_3\mu_1^2 + 30\mu_2^2\mu_1 - 60\mu_2\mu_1^3 + 24\mu_1^5
\end{align*}
\] (A.40)

and there are similar relations in opposite direction. As can be seen, \( \kappa_1 \) is the mean and \( \kappa_2 \) is the variance of the quantity under investigation. The higher-order cumulants are related to more complex statistical properties such as the kurtosis \( \gamma_2 = \kappa_4 / \kappa_2^2 \), which is a measure for the “peakedness” of the distribution with respect to the variance. The third cumulant typically measures the tilt of a distribution.

E. The Schnakenberg Formula for Entropy Production

Eq. \([6.31]\) is known as the celebrated Schnakenberg formula because it was derived for the first time in 1976 by Jürgen Schnakenberg [7]. Today the Schnakenberg formula is accepted as a general expression which quantifies the entropy production in the environment, irrespective of the specific structure of the environment. This means that it is generally assumed that the entropy production, no matter whether the environment is equilibrated or not, is given by Eq. \([6.31]\). Most researchers use this formula as it is without paying attention to its range of validity.

It is therefore interesting to investigate how this formula was actually derived. Schnakenberg noticed that the master equation has exactly the same structure as the law of mass action for chemical reactions. This analogy can be established by associating with each configuration \( s \) a chemical substance \( X_s \):

\[
s \iff X_s
\]

These fictitious chemical substances are then mixed in reactor where they react with one
Another. The chemical reaction rates $k_{ss'}$ in turn are chosen according to the transition rates of the master equation. Finally, the entropy production of the chemical reaction is studied on the basis of conventional thermodynamics.

Let us briefly sketch this derivation in more detail. In physical chemistry, an important quantity is the so-called extent of reaction $\xi_{ss'}$, defined as the accumulated number of forward reactions $X_s \rightarrow X_{s'}$ minus the number of backward reactions $X_{s'} \rightarrow X_s$. This quantity evolves in time according to

$$\frac{d\xi_{ss'}}{dt} = N_s k_{ss'} - N_{s'} k_{s's}, \quad (A.41)$$

where $N_s$ is the number of molecules of type $X_s$.

In conventional thermodynamics, each thermodynamic flux is associated with a conjugate thermodynamic force defined as the partial derivative of the free energy with respect to the flux. In the present case, the temporal change of the extent of reaction defines the thermodynamic flux. The conjugate thermodynamic force is the so-called chemical affinity

$$A_{ss'} = \left. \frac{\partial F}{\partial \xi_{ss'}} \right|_{V,T} \quad (A.42)$$

This implies that the chemical reaction changes to three energy according to

$$\dot{F} = \sum_{ss'} A_{ss'} \dot{\xi}_{ss'} = \sum_{ss'} A_{ss'} (\dot{N}_{s'} - \dot{N}_s). \quad (A.43)$$

This has to be compared with a chain rule

$$\dot{F} = \sum_s \frac{\partial F}{\partial N_s} \dot{N}_s = \sum_s \mu_s \dot{N}_s \quad (A.44)$$

where $\mu_s$ denotes the chemical potential of the substance $X_s$. As a result, we obtain that the affinities are basically given by the chemical potential difference

$$A_{ss'} = \mu_{s'} - \mu_s. \quad (A.45)$$

However, the chemical potentials may change as the reaction proceeds, i.e., they depend on the actual molecule numbers $N_s$. To take this additional change into account, we note that the free energy is usually given by an expression of the form

$$F = \sum_s \left( const_s N_s + k_B T N_s \ln N_s \right). \quad (A.46)$$

For example, the free energy of an ideal gas can be written in this form. This implies that the chemical potentials can be written as

$$\mu_s = \mu_s^{(0)} + k_B T \ln N_c. \quad (A.47)$$

Inserting this result the affinities are given by

$$A_{ss'} = \mu_{s'}^{(0)} - \mu_s^{(0)} + k_B T \ln \frac{N_{s'}}{N_s}. \quad (A.48)$$
In order to access the difference of the two constants on the right-hand side, let us consider the stationary equilibrium state. Here, the extent of reaction and therewith the affinity is zero, i.e. \( A_{ss'}^{(0)} = 0 \), hence
\[
\mu_s^{(0)} - \mu_{s'}^{(0)} = k_B T \ln \frac{N_{ss}^{eq}}{N_{s's'}^{eq}} = k_B T \ln \frac{k_{ss'}}{k_{s's'}},
\] (A.49)
where we used the balance condition \( N_s k_{ss'} = N_{s'} k_{s's'} \) in the last step. This leads us to
\[
A_{ss'} = k_B T \ln \frac{k_{ss'}}{N_s k_{s's'}}. \quad \text{(A.50)}
\]
Putting all things together we arrive at the free energy change
\[
\dot{F} = k_B T \sum_{ss'} \xi_{ss'} \ln \frac{N_s k_{ss'}}{N_s k_{s's'}}. \quad \text{(A.51)}
\]
Introducing the concentrations \([X_s] := N_s / N\) and recalling that the free energy is nothing but the trick to quantify entropy changes of the total system, consisting of laboratory system and environment, we get
\[
\frac{dS_{\text{tot}}}{dt} = -\beta \dot{F} = \sum_{ss'} \xi_{ss'} \ln \frac{[X_s]}{[X_{s'}]} \frac{k_{ss'}}{k_{s's'}}. \quad \text{(A.52)}
\]
We now identify the chemical problem with the original master equation by setting
\[
\xi_{ss'} \approx P_s(t) w_{s \rightarrow s'}(t), \quad [X_s] \approx P_s(t), \quad k_{ss'} \approx w_{s \rightarrow s'}(t) \quad \text{(A.53)}
\]
The courageous corollary by Schnakenberg is to identify the chemical entropy production with a total entropy production of the master equation:
\[
\frac{dS_{\text{tot}}}{dt} = \sum_{ss'} P_s(t) w_{s \rightarrow s'}(t) \ln \frac{P_s(t) w_{s \rightarrow s'}(t)}{P_{s'}(t) w_{s' \rightarrow s}(t)}. \quad \text{(A.54)}
\]
This can be split into two parts
\[
\frac{dS_{\text{tot}}}{dt} = \sum_{ss'} P_s(t) w_{s \rightarrow s'}(t) \ln \frac{P_s(t)}{P_{s'}(t)} + \sum_{ss'} P_s(t) w_{s \rightarrow s'}(t) \ln \frac{w_{s \rightarrow s'}(t)}{w_{s' \rightarrow s}(t)}. \quad \text{(A.55)}
\]
which can be identified with the system at the environment.

The environmental part derived by Schnakenberg can be written as
\[
\frac{dS_{\text{env}}}{dt} = \sum_{ss'} J_{s \rightarrow s'}(t) \ln \frac{w_{s \rightarrow s'}(t)}{w_{s' \rightarrow s}(t)}. \quad \text{(A.56)}
\]
Obviously, his expression the logarithm is averaged over the current flowing from configuration \( s \) to configuration \( s' \). In 2005 this led Seifert and Speck to the conjecture that the entropy in the environment increases discontinuously whenever the system jumps...
to a different configuration:

$$\Delta S_{s \rightarrow s'}^{\text{env}} = \ln \frac{w_{s \rightarrow s'}(t)}{w_{s' \rightarrow s}(t)}.$$  \hspace{1cm} (A.57)

Today, this formula is commonly accepted as valid expression for the entropy production in the surrounding medium. It is remarkable that this expression does not depend any way on the physical properties of the environment. Does the environment really generate only that little entropy for the microscopic transition from $s$ to $s'$? Could it be that the real entropy production in an inefficient setup is actually higher?

This is an important question which to my knowledge has not been addressed so far. In fact, both derivations given above, the one with the fully connected transition network and Schnakenberg’s chemical variant, implicitly assume that the environment equilibrates almost instantaneously on a much faster timescale than the intrinsic dynamics of the system. If it does not equilibrate directly after each transition the actual entropy production may be higher. Probably Schnakenberg’s formula provides only a lower bound.
List of Acronyms

In these lecture notes we frequently use the following acronyms, here sorted in alphabetical order:

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACS</td>
<td>Algebraic computer system</td>
</tr>
<tr>
<td>ASEP</td>
<td>Asymmetric exclusion process</td>
</tr>
<tr>
<td>CP</td>
<td>Contact process</td>
</tr>
<tr>
<td>DP</td>
<td>Directed percolation</td>
</tr>
<tr>
<td>FT</td>
<td>Fluctuation theorem</td>
</tr>
<tr>
<td>IFT</td>
<td>Integral fluctuation theorem</td>
</tr>
<tr>
<td>l.h.s.</td>
<td>Left hand side of equation</td>
</tr>
<tr>
<td>MF</td>
<td>Mean field approximation</td>
</tr>
<tr>
<td>PASEP</td>
<td>Partially asymmetric exclusion process</td>
</tr>
<tr>
<td>PCPD</td>
<td>Pair contact process with diffusion</td>
</tr>
<tr>
<td>r.h.s.</td>
<td>Right hand side of equation</td>
</tr>
<tr>
<td>SDE</td>
<td>Stochastic differential equation</td>
</tr>
<tr>
<td>SEP</td>
<td>Symmetric exclusion process</td>
</tr>
<tr>
<td>SPDE</td>
<td>Stochastic partial differential equation</td>
</tr>
<tr>
<td>TASEP</td>
<td>Totally asymmetric exclusion process</td>
</tr>
</tbody>
</table>
Bibliography