## Foundations of Quantum Mechanics

Lecturer: Prof. C. Wetterich

Winter term 2021/22, Heidelberg University

## Foreword

This script is aimed to facilitate students to follow the lecture on foundations of quantum mechanics<sup>1</sup>. It is based on my handwritten notes and done by a group of engaged students, Julien Haagner, David Ridinger, Christos Karapoulitidis, Georgios Theologou, Simon Wagner, Johannes Hägele. Please let them know if you find errors. Up to now I did not do any corrections or writing of text - this may change in the future. For the moment it seems most helpful to make the notes available rapidly, even though not everything is perhaps already perfect. Thanks a lot to the writing team for this great effort and result! Christof Wetterich

<sup>&</sup>lt;sup>1</sup>The recorded lectures can be found on: https://www.thphys.uni-heidelberg.de/~wetterich/fqm21/. One may also follow the paper: https://inspirehep.net/literature/1828609

# Contents

1	Basic properties of quantum mechanics and classical statistics				
	1.1 Formalism of quantum mechanics	4			
	1.2 Formalism of classical statistics	9			
	1.3 Comparison between QM and CS	12			
	1.4 The density matrix in QM	13			
<b>2</b>	babilistic particles				
3	Explanation of some "Quantum mysteries"				
	3.1 Particle wave duality	22			
	3.2 Entanglement	23			
	3.3 Phases in quantum dynamics	26			
	3.4 Planck's constant	29			
4	Ising chain and wave functions in classical statistics 30				
	4.1 Ising chain	30			
	4.2 Classical wave functions	35			
	4.3 Transfer matrix and evolution law for wave functions	39			
	4.4 Step evolution operator	43			
	4.5 Operators of observables	45			
	4.6 Solution of the initial value problem	45			
	4.7 Classical density matrix	50			
	4.8 Non-commuting operators	51			
	4.9 Quantum formulation for classical statistics	53			
<b>5</b>	Probabilistic cellular automata	<b>54</b>			
	5.1 Deterministic cellular automata	54			
	5.2 Probabilistic (cellular) automaton	56			
	5.3 Clock automata	57			
6	Generalised Ising model for fermions 5				
	6.1 Diagonal Ising model	59			
	6.2 Particle-hole conjugation and complex structure	61			
	6.3 Vacuum	64			
	6.4 One-particle wave function	65			

	$\begin{array}{c} 6.5 \\ 6.6 \end{array}$	Continuum limit       67         Interactions       72         6.6.1       Quantum Field Theory for fermions       73
7	Rec	luction of the wave function and conditional probabilities 74
	7.1	Reduction of the wave function
	7.2	Conditional probabilities
	7.3	Measurement correlation
	7.4	Non-uniqueness of conditional probabilities
	7.5	Ideal measurements
	7.6	Unequal time quantum correlations
	7.7	Decoherent ideal measurements and reduction of the wave function
	7.8	Bell's inequalities

### Chapter 1

# Basic properties of quantum mechanics and classical statistics

The aim of this lecture is to show that a quantum description of our world follows from classical probability theory, so that quantum mechanics can be viewed as a subfield of classical statistics. A special aim will be to emphasise that there are no further laws or axioms of quantum mechanics needed, it is already included in the classical statistics. But given that aim, we should recall some basic properties from both quantum mechanics and classical statistics.

#### **1.1** Formalism of quantum mechanics

The formalism of quantum mechanics should be familiar. Especially in connection with the Copenhagen interpretation, which is a set of axioms we want to overcome by viewing quantum mechanics as natural subfield of classical statistics. However, the fundamentals of quantum mechanics as a probabilistic theory should be outlined once again.

#### 1. Wave function $\psi$ :

In quantum mechanics, physical states  $|\psi\rangle$ , which contain all information about a system we know, are elements of a Hilbert space. Given the state of quantum mechanical system and the corresponding Hamiltonian, the time evolution is fully determined and all possible outcomes of measurements of an observables are clear, this will be recalled shortly.

Directly connected to quantum mechanical states is another important object, the wave function  $\psi$ . The wave function is a representation of the state  $|\psi\rangle$  in position space (or in momentum space):

$$\psi(t, \vec{x}) := \langle \vec{x} | \psi(t) \rangle \tag{1.1}$$

where the time evolution  $|\psi(t)\rangle$  is determined by Schrödinger's equation. The wave function acts as a probability amplitude, while the absolute square  $|\psi|^2$  then describes the probability distribution of the system, which is the reason for quantum mechanics being a probabilistic theory. Indeed, the time evolution of a state or a wave function is deterministic, by Schrödinger's equation, but both states and the wave equations are mathematical objects which are not observable. Further, recall that states are in general superpositions of eigenstates of the observable we want to measure, a superposition of states has no physical interpretation. We deal with probability distributions containing all possible outcomes of a measurement of an observable. Hence, physical reality is then generated by such an outcome, not by a superposition of eigenstates or a wave function itself.

In order to use states or wave functions as probabilistic objects they need to be normalised:

$$\langle \psi | \psi \rangle = \int d^3 x |\psi(\vec{x})|^2 = 1 \tag{1.2}$$

During this lecture we will focus on a discrete description of quantum mechanics. For example let A be an operator, as defined below, with a discrete spectrum and orthonormal eigenbasis  $\{\alpha_i\}$ , then any state in the Hilbert space can be written as:

$$|\psi\rangle = \sum_{i} c_i |a_i\rangle \tag{1.3}$$

where  $c_i \in \mathbb{C}$ . The normalisation then reads:

$$\sum_{i} |c_i|^2 = 1 \tag{1.4}$$

Thus,  $|c_i|^2$  denotes the probability to measure the eigenvalue of the eigenstate  $|\alpha_i\rangle$ , if the system was prepared in the state  $|\psi\rangle$ . Since we are dealing with Hilbert spaces, continuous situations can be viewed as a continuum limit, the formalism of quantum mechanics remains the same.

From now on we will no longer use Dirac's notation and simply denote states by  $\psi$  so that:

$$\langle \psi | \phi \rangle \to \psi^{\dagger} \phi$$
 (1.5)

The difference between states and wave functions should be clear.

#### 2. Observables:

Observables A in quantum mechanics correspond to operators  $\hat{A}$ . We will ignore the  $\hat{\cdot}$  throughout the lecture. Such operators A are elements of the space of linear maps  $\mathscr{H} \to \mathscr{H}$ , where  $\mathscr{H}$  denotes the Hilbert space, which are further Hermitian  $(A = A^{\dagger})$ . Being Hermitian the operators have a eigenbasis with real eigenvalues. Real eigenvalues are needed, because the set of possible outcomes of a measurement of an observable equals the spectrum of the operator corresponding to the observable. The term "possible outcome" implements that measuring an eigenvalue is fundamentally random! The probability of measuring a certain eigenvalue, however, depends on the initial state of the system and its representation in the eigenbasis. Assume a system prepared in the state  $\psi$ , the probability to measure it in the eigenstate  $\alpha_i$ can expressed by the scalar product  $\alpha_i^{\dagger}\psi$ :

$$p(\alpha_i) = |\alpha_i^{\dagger} \psi|^2 \tag{1.6}$$

**Example 1** (Spin operator). Assume the simple case of a spin-1/2-system, which can be described in a two-dimensional complex Hilbert space. It is convenient to choose  $\mathbb{C}^2$ . The observable of interest is then the spin (orientation), with the corresponding spin operator  $\vec{S} = (S_1, S_2, S_3)^{\intercal}$ . Choosing the basis  $\{(1, 0)^{\intercal}, (0, 1)^{\intercal}\}$ , the representation of its components is:

$$S_i = \frac{\hbar}{2} \tau_i \tag{1.7}$$

where the  $\tau_k$  denote the Pauli matrices:

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
(1.8)

Obviously all of the spin operators' components, and therefore the spin operator itself, are Hermitian. Moreover,  $\{(1,0)^{\intercal}, (0,1)^{\intercal}\}$  is the eigenbasis of  $S_3$  with the eigenvalues  $\pm \hbar/2$ , which are usually denoted  $\{|\uparrow\rangle, |\downarrow\rangle\}$ .

Assume a general spin-1/2 state in the basis as above:

$$|\psi\rangle = \alpha |\uparrow\rangle + \beta |\downarrow\rangle \tag{1.9}$$

where  $\alpha, \beta \in \mathbb{C}$ :  $|\alpha|^2 + |\beta|^2$ . Thus,  $|\alpha|^2$  stands for the probability to measure a quantum mechanical system prepared in the state  $|\psi\rangle$  with spin up in z-direction, the analogue holds for  $|\beta|^2$  and spin down. The spin in the x- and y-direction can be determined by the scalar product of  $|\psi\rangle$  with one of the two eigenvectors of  $S_1$  or  $S_2$ .

#### 3. Non-commutativity:

Another speciality of quantum mechanics is that non-commutativity of operators, corresponding to observables, leads to the inability of measuring both observables exact simultaneously. From linear algebra it known that:

> $[A, B] = 0 \Leftrightarrow A$  and B are diagonisable simultaneously  $\Leftrightarrow A$  and B share a common eigenbasis

for two operators A,B. The probably most famous pair of non-commutating operators is (q,p), position and momentum of a particle cannot be measured exact simultaneously.

**Example 2** (Spin components). As another example of non-commuting operators choose the components of the spin-operator, which satisfy:

$$[S_i, S_j] = i\epsilon_{ijk}S_k \tag{1.10}$$

while:

$$[\vec{S}^2, S_i] = 0 \; \forall i \in \{1, 2, 3\} \tag{1.11}$$

So it is possible to know the absolute value  $\vec{S}^2$  of the spin, but only one of its directions at the same time.

#### 4. Unitary time evolution:

It has already been mentioned that the time evolution of a quantum mechanical state/ wave function is deterministic, as long as the system can evolve without a disturbance from the outside, like a measurement. Recall that there are different pictures of quantum mechanics, the time evolution of states or wave functions is naturally described in the Schrödinger picture, using Schrödinger's equation.

Due to our probabilistic approach of quantum mechanics, it is necessary that the normalisation of a wave function is conserved under time evolution:

$$\psi_0^{\dagger}\psi_0 = \psi^{\dagger}(t)\psi^{\dagger}(t) \tag{1.12}$$

This eventually leads to the definition of the unitary time operator.

#### (a) Schrödinger Picture:

In the Schrödinger picture a state/ wave function is time dependent:

$$\psi = \psi(t) \tag{1.13}$$

while operators are stationary. In quantum mechanics the expectation value of an observable, with corresponding operator A, therefore changes due to time dependence of the probability amplitude:

$$\boxed{\langle A(t)\rangle = \psi^{\dagger}(t)A\psi(t)}$$
(1.14)

This expression is defined analogue to the expectation value in classical statistics. The discrete version reads:

$$\langle A(t) \rangle = p_{\alpha}(t)A_{\alpha} \tag{1.15}$$

where  $p_{\alpha}(t)$  denotes the probability to measure the value  $A_{\alpha}$  at the time t.

#### (b) Unitary evolution law:

The time evolution of a state/ wave function is determined by a unitary operator, the time evolution operator, which is defined as follows:

$$\psi(t_2) = U(t_2, t_1)\psi(t_1)$$
(1.16)

 $U(t_2, t_1)$  can be represented by a unitary matrix:

$$U^{\dagger}(t_2, t_1)U(t_2, t_1) = 1 \tag{1.17}$$

so that inverse time evolution satisfies:

$$U^{\dagger}(t_2, t_1) = U(t_1, t_2) \tag{1.18}$$

which easily proven inverting equation (1.17):

$$\psi(t_1) = U(t_1, t_2)\psi(t_2) = U^{\dagger}(t_2, t_1)\psi(t_2)$$
(1.19)

There is also the trivial case of  $t_1$  being equal to  $t_2$ :

$$U(t_1, t_1) = 1 \tag{1.20}$$

Performing multiple time evolutions can be simplified to a single evolution from the initial timepoint to the last one:

$$U(t_3, t_2)U(t_2, t_1) = U(t_3, t_1)$$
(1.21)

#### (c) Infinitesimal evolution:

We define an infinitesimal timestep  $\epsilon$ :

$$\epsilon = t_2 - t_1, \ \epsilon \to 0$$

We can use this newly defined variable to relate the time evolution operator U to the Hamiltonian H:

$$U(t_2, t_1) = 1 - iH(t_1)(t_2 - t_1) = 1 - i\epsilon H(t_1)$$
(1.22)

For the Hamilton operator  $H = H^{\dagger}$  applies, which can be carried over to U being unitary:

$$U^{\dagger}U = (1 + i\epsilon H)(1 - i\epsilon H) = 1 + \epsilon^2 H^2$$
(1.23)

which is equal to 1 for  $\epsilon = 0$ .

#### (d) Step evolution operator:

The step evolution operator gives another possibility of time evolution of a state/ the wave function in case of a finite small  $\epsilon$ :

$$U(t + \epsilon, t) = exp(-i\epsilon H(t))$$
(1.24)

As before the following properties apply:  $U^{\dagger}U = 1$  and  $H^{\dagger} = H$ . Define a fixed reference time  $t_0$ :  $U(t) = U(t, t_0)$ , for differentiable a U(t) we can write the Hamiltonian H as:

$$H(t) = i(\partial t \ U(t))U^{\dagger}(t)$$
(1.25)

Proof:

$$\partial t U(t) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} (U(t+\epsilon) - U(t))$$
$$= \lim_{\epsilon \to 0} \frac{1}{\epsilon} (U(t+\epsilon,t) - 1)U(t)$$
$$\Rightarrow (\partial t U(t))U^{\dagger}(t) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} (exp(-i\epsilon H(t)) - 1)$$
$$= -iH(t)$$

We should check whether our time evolution law is indeed unitary. The time evolution being unitary is equivalent to the Hamilton operator satisfying  $H^{\dagger} = H$ . A simple calculation yields:

$$H^{\dagger} = -iU(\partial t \ U^{\dagger}) = i(\partial t \ U)U^{\dagger} = H$$
(1.26)

so that:

$$U^{\dagger}(t+\epsilon,t) = exp(+i\epsilon H^{\dagger}) = (U(t+\epsilon,t))^{-1}$$
(1.27)

#### (e) Schrödinger's equation:

The time evolution of states/ wave functions is fully determined by the Schrödinger equation:

$$i\partial t \ \psi(t) = H(t)\psi(t) \tag{1.28}$$

This immediately follows using how the time evolution operator is related to the Hamilton operator under small time steps (1.24):

$$\partial t \ \psi(t) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} [\psi(t+\epsilon) - \psi(t)]$$
$$= \lim_{\epsilon \to 0} \frac{1}{\epsilon} [U(t+\epsilon) - U(t)]\psi_0$$
$$= (\partial t \ U(t))U^{\dagger}(t)\psi(t)$$
$$= -iH(t)\psi(t)$$

#### **1.2** Formalism of classical statistics

Here, the foundation of classical statistics shall be outlined. This formalism will later we used to derive the known quantum laws and is therefore of great importance for the whole lecture.

#### 1. States

In the viewpoint of classical statistics, a system always is in a certain state, which can be described by corresponding variables. In thermodynamics, one usually uses the term 'microstate' instead, while probability theory uses the word 'event'.

**Example 3** (states). (a) For an ensemble of n bits, which each can take the values 0 or 1, the configuration can be in  $2^n$  possible states.

- (b) The phase space in classical mechanics describes the state of a particle *i* with its coordinates  $(\vec{q_i}, \vec{p_i})$ .
- (c) The in thermodynamics, a system can be in different microstates characterized be the variables of state.

We label the states with the greek letters  $\omega$  or  $\tau$ . For a system with 5 bits, we could for instance define the state  $\tau$  to be

$$\begin{aligned} \tau &= (1,0,0,1,1) \\ &= 19 \\ &= (+,-,-,+,+) \end{aligned}$$

or likewise

 $= \{n_{\gamma}\}, n_{\gamma} = 0, 1, \gamma = 1, ..., 5$ 

A system can have discrete or continuous states. In the following, we will concentrate on the discrete case and encounter sums over different states. In the continuous case, these sums need to be replaced by Lebesgue integrals.

#### 2. Observables

As opposed to the formalism in quantum mechanics, observables are simply described by real numbers in classical statistics. In a given state  $\tau$ , the observable A takes the value  $A_{\tau}$ . The spectrum of A is then given by  $\{A_{\tau}\}$  for all possible states  $\tau$ .

#### 3. Probability distribution

In classical statistics, every possible state  $\tau$  is associated with a probability  $p_{\tau}$  to find the system in this state. The probabilities need to meet the requirements of non-negativity

$$p_{\tau} \ge 0 \quad \forall \ \tau \tag{1.29}$$

and normalisation.

$$\sum_{\tau} p_{\tau} = 1 \tag{1.30}$$

#### 4. Expectation value of observables

The expectation value of an observable A with a spectrum  $\{A_{\tau}\}$  for states  $\tau$  associated with their probabilities  $p_{\tau}$  is simply given by the *basic law of classical statistics*:

$$\langle A \rangle = \sum_{\tau} p_{\tau} \cdot A_{\tau} \tag{1.31}$$

#### 5. Time evolution

In this section we want to investigate, whether we can find a general description of how a system in classical statistics evolves in time. In other words, we are searching for an equivalent to the Schrödinger equation in quantum dynamics. Our first step towards such a description is to recognise that time becomes an additional state variable.

For example, an ensemble of five bits at the positions  $x_i$ , which can change its configuration within the time steps  $t_i$  (i = 0, ..., 4), can be found in a total of N=2<sup>25</sup> states  $\omega = \{n(t, x)\}$ with n=0,1. These  $\omega$  cover all times and the associated probabilities  $p_{\omega}$  give the overall probability distribution.

If we are interested in the probability of the system being in a certain configuration at a selected time (e.g.  $t_0$ ), we need the *time-local probability distribution*. In this case, we only want to consider the *subsystem* of the  $N_{time-local} = 2^5$  events  $\tau = \{n(t_0, x))\}$ . It is worth mentioning that here,  $\tau$  is always a time-local event, while  $\omega$  covers all times. The time-local probability, to find the system in a state  $\tau$ , is then given by the sum over all overall probabilities of events  $\omega$  for which the configuration at  $t=t_0$  matches  $\tau$ :

$$p_{\tau} = \sum_{\{\omega' \mid \{n(t=t_0, x)\} = \tau\}} p_{\omega'}$$
(1.32)

**Example 4.** For an even simpler example, we consider only one bit at the times  $t_i$  for i = 0, ..., 4. We therefore have  $2^5$  states

$$\omega = \{n(t)\} = \{n(t_0), ..., n(t_4)\}$$
(1.33)

If we now consider the time  $t_2$ , we only have 2 possible states  $\tau = n(t_2) \in \{0,1\}$ . The time-local probability to find the bit with the configuration 1 or 0 at  $t=t_2$ , is then given by:

$$p_1(t_2) = \prod_{t' \neq t_2} \sum_{n(t')} p\left(n\left(t_0\right), n\left(t_1\right), n\left(t_2\right) = 1, n\left(t_3\right), n\left(t_4\right)\right)$$
(1.34)

or:

$$p_0(t_2) = \prod_{t' \neq t_2} \sum_{n(t')} p\left(n\left(t_0\right), n\left(t_1\right), n\left(t_2\right) = 0, n\left(t_3\right), n\left(t_4\right)\right)$$
(1.35)

respectively. The notation of the configuration sum hereby must not be interpreted as a product of individual sums, but rather as a product of the summation signs:

$$\prod_{i=1}^{3} \sum_{n_i=0}^{1} p(n_1, ..., n_3) = \sum_{n_1=0}^{1} \sum_{n_2=0}^{1} \sum_{n_3=0}^{1} p(n_1, ..., n_3)$$
(1.36)

If we now generalize this simple example to a lattice of bits with spatial coordinates x, we obtain the time-local probability

$$p_{\tau}(\bar{t}) = p\left(\{n(\bar{t}, x)\}\right) = \prod_{t' \neq \bar{t}} \prod_{x} \sum_{n(t', x) = 0}^{1} p\left(\{n(t', x), n(\bar{t}, x)\}\right)$$
(1.37)

at the time  $\bar{t}$ . The interpretation is simple, the probability for the time-local event  $\{n(\bar{t}, x)\}$  to occur, equals the sum of all probabilities assigned to an overall event including this time-local event.

We now want to consider a time-local observable A, which takes the values

$$A_{\tau}(\bar{t}) = A(\{n(\bar{t}, x)\})$$
(1.38)

for the time-local events  $\tau = \{n(\bar{t}, x)\}$ . According to the basic law of classical statistics (1.31), the expectation value for this observable at a time  $\bar{t}$  is then given by:

$$\langle A\left(\bar{t}\right)\rangle = \sum_{\omega} p_{\omega} \cdot A_{\omega} \tag{1.39}$$

By separating the configuration sum into a sum over all times but  $\bar{t}$  (denoted with the product sign) and a sum dependent on  $\bar{t}$ , one obtains an expression simply containing time-local

entities:

$$\begin{split} \langle A\left(\bar{t}\right) \rangle &= \sum_{\omega} p_{\omega} \cdot A_{\omega} \\ &= \prod_{t,x} \sum_{n(t,x)} p\left(\{n\left(t,x\right)\}\right) \cdot A\left(\{n\left(\bar{t},x\right)\}\right) \\ &= \left(\prod_{x} \sum_{n\left(\bar{t},x\right)}\right) \left(\prod_{t' \neq \bar{t}} \prod_{x} \sum_{n\left(t',x\right)}\right) p\left(\{n\left(t',x\right)\}, \{n\left(\bar{t},x\right)\}\right) \cdot A\left(\{n\left(\bar{t},x\right)\}\right) \\ &= \left(\prod_{x} \sum_{n\left(\bar{t},x\right)}\right) A\left(\{n\left(\bar{t},x\right)\}\right) \left(\prod_{t' \neq \bar{t}} \prod_{x} \sum_{n\left(t',x\right)}\right) p\left(\{n\left(t',x\right)\}, \{n\left(\bar{t},x\right)\}\right) \\ &= \prod_{x} \sum_{n\left(\bar{t},x\right)} A\left(\{n\left(\bar{t},x\right)\}\right) \cdot p\left(\{n\left(\bar{t},x\right)\}\right) \\ &= \sum_{\tau} A_{\tau}\left(\bar{t}\right) \cdot p_{\tau}\left(\bar{t}\right) \end{split}$$

We see that the expectation value of A at the time  $\bar{t}$  can be calculated from the time-local probabilities  $\{p_{\tau}(t)\}$ , which are defined by the overall probability distribution,  $\{p_{\omega}\}$ , and the time-local values of A:

$$\langle A\left(\bar{t}\right)\rangle = \sum_{\tau} A_{\tau}\left(\bar{t}\right) \cdot p_{\tau}\left(\bar{t}\right)$$
(1.40)

This is what we expected intuitively, the law for expectation values can be transferred to time-local events.

To get to an *evolution law*, however, we are interested in computing the time-local probabilities  $\{p_{\tau}(t_2)\}$  at a time  $t_2$  based on the the time-local probabilities  $\{p_{\tau}(t_1)\}$  at an earlier point in time  $t_1$ . Sometimes, but *not always*, this can be achieved, examples are the Liouville equation or Markow chains. Later, we will see a large class of systems where the evolution law can be expressed in terms of a density matrix. The density matrix formalism will be a main point to connect classical statistics and quantum mechanics. It is time for a first comparison between classical statistics and quantum mechanics.

#### 1.3 Comparison between QM and CS

Until now, we have seen that both QM and CS are probabilistic theories. There are states and observables which can take different values depending on the state the system is in. Probabilities enter with a spectrum of possible measurement values. But, beyond there are many points where classical statistics and quantum mechanics seem to differ at first sight. Which is the reason why quantum world is commonly viewed as mysterious, even though it is just a probabilistic theory in the end. We shall name some of these points:

1. the existence of wave function, which is the analogue of a probability amplitude, but is not observable

- 2. complex quantities
- 3. operators for observables
- 4. non-commutativity
- 5. quantum law for expectation values
- 6. unitary time evolution

Recall the aim of this lecture is to show that Quantum Mechanics emerge from Classical Statistics, and that there is no mystery to it. All quantum laws can be found in classical statistics. We will show that all Quantum laws follow from basic laws of Classical Statistics. In the end we are interested in expectation values of our observables for a given overall probability distribution  $\{p_{\omega}\}$ :

$$\langle A \rangle = \sum_{\omega} p_{\omega} \cdot A_{\omega} \tag{1.41}$$

Preview:

- 1. The description of the world is fundamentally probabilistic.
- 2. The basic laws of Classical Statistics are sufficient to describe the quantum world.

Our conclusion will be that the Foundations of Quantum Mechanics are Classical Statistics. This approach is conceptually very simple, without adding complicate (philosophical) concepts ("many worlds, "reality of wave function", "role of the observer", "contextuality" etc.), but resolving the mysterious view on quantum mechanics.

#### 1.4 The density matrix in QM

A very important object for quantum mechanics is the so called density matrix, it contains all probabilistic information about a system we can gather. Expectation values can be immediately calculated from it, and time evolution is included as well. The density matrix will play an important part in showing that quantum mechanics emerge from classical statistics.

1. Basic formulation of QM mechanics in terms of  $\varrho(t)$ :

For a finite-dimensional function space, the most general density operator is of the form:

$$\varrho \equiv \sum_{i} p_{i} \psi_{i} \psi_{i}^{\dagger} \tag{1.42}$$

where the coefficients  $p_i$  are non-negative and add up to one, and  $|\psi_i\rangle \langle \psi_i|$  is an outer product written in bra-ket notation. This can be interpreted as ensemble of states  $\psi_i$  with the probabilities  $p_i$  to find the ensemble the state i. In quantum mechanics the notion of mixed states is used.

Quantum mechanics can be fully described using the density matrix formalism, we list the main concepts:

(a) Properties:

$$tr(\varrho) = 1 \tag{1.43}$$

$$\varrho^{\dagger} = \varrho \tag{1.44}$$

(b) Quantum law for expectation values:

$$\langle A(t) \rangle = \sum_{j} \langle j | \varrho(t) | A | j \rangle = tr(\varrho(t) | A)$$
(1.45)

(c) Unitary evolution:

$$\varrho(t_2) = U(t_2, t_1)\varrho(t_1)U^{\dagger}(t_2, t_1)$$
(1.46)

#### 2. Pure states:

This is the case of ordinary quantum mechanics which the state of the system described by a one-dimensional subspace in the Hilbert space. Our ensemble reduces to a single subspace, which is visited with a probability of 1. In this case the density matrix can be written as:

$$\varrho = \psi \psi^{\dagger} \tag{1.47}$$

We recognize that the structure of density matrix in pure state is exactly the same like with a projection operator, but here with a projection on the 1-dimensional subspace  $\psi$ . An essential property is:

$$\varrho^2 = \varrho \tag{1.48}$$

Proof:

$$\varrho^2 = \varrho \ \varrho = \psi \psi^{\dagger} \psi \psi^{\dagger} = \psi \psi^{\dagger} = \varrho \tag{1.49}$$

In order to proof the opposite we need a general matrix property which is beyond the scope of this course.

Now, we will try to formulate the expectation value in terms of the Density matrix

$$\langle A(t) \rangle = \psi^{\dagger}(t) A \psi(t) \tag{1.50}$$

We assume that the eigenstates of the operator A are the  $\alpha_j$ , then the eigenequation reads:

$$A\alpha_j = a_j \alpha_j \tag{1.51}$$

Thus:

$$\langle A \rangle = \sum_{j} p_{j} a_{j} = \sum_{j} |\alpha_{j}^{\dagger} \psi|^{2} a_{j}$$
(1.52)

$$|\alpha_j^{\dagger}\psi|^2 = (\alpha_j^{\dagger}\psi)(\psi^{\dagger}\alpha_j) \tag{1.53}$$

we can rewrite the above mean value:

$$\begin{split} \langle A \rangle &= \sum_{j} (\alpha_{j}^{\dagger} \psi) (\psi^{\dagger} a_{j} \alpha_{j}) \\ &= \sum_{j} (\alpha_{j}^{\dagger} \psi) (\psi^{\dagger} A \alpha_{j}) \\ &= \sum_{j} \alpha_{j}^{\dagger} (\varrho A) \alpha_{j} \end{split}$$

This is the quantum law for expectation values in the density matrix formalism:

$$\boxed{\langle A(t)\rangle = tr(\varrho(t)A)}$$
(1.54)

The evolution law for the pure state is the following:

$$\psi(t_2) = U(t_2, t_1)\psi(t_1), \qquad \psi^{\dagger}(t_2) = \psi(t_1)U^{\dagger}(t_2, t_1)$$
 (1.55)

$$\psi_{\alpha}(t_2) = U_{\alpha\gamma}(t_2, t_1)\psi_{\gamma}(t_1), \qquad \psi_{\beta}^*(t_2) = \psi_{\delta}^*(t_1)U_{\delta\beta}^{\dagger}(t_2, t_1)$$
(1.56)

And with this we can formulate the evolution of the density matrix  $\rho_{\alpha\beta}$ :

$$\rho_{\alpha\beta}(t_2) = \psi_{\alpha}(t_2)\psi^*_{\beta}(t_2) = U_{\alpha\gamma}(t_2, t_1)\psi_{\gamma}(t_1)\psi^*(t_2)U^{\dagger}_{\delta\gamma}$$
$$= U_{\alpha\gamma}(t_2, t_1)\psi_{\gamma}(t_2)\psi^*_{\delta}(t_2)U_{\delta\beta}(t_2)$$

So that:

$$\rho(t_2) = U(t_2, t_1)\rho(t_2)U^{\dagger}(t_2, t_1)$$
(1.57)

#### 3. Mixed states:

Different pure states  $\psi^{(j)}$ , with the corresponding pure state density matrices  $\varrho^{(j)}$ 

$$\varrho = \sum_{j} w_{j} \rho^{(j)}, \qquad w_{j} \ge 0, \qquad \sum_{j} w_{j} = 1$$
(1.58)

Here  $w_j$  is the probability to realize the pure state j. Further properties of the density matrix are:

$$tr\varrho = \sum_{j} w_j tr\varrho^{(j)} = \sum_{j} w_j = 1, \qquad \varrho^{\dagger} = \varrho, \qquad \varrho \text{ is positive}$$
(1.59)

 $\rho$  is positive if:  $w_i \ge 0$ . Every density matrix can be represented in this way with  $w_j$ . The Quantum rules for expectation values and the unitary evolution law are <u>linear</u> in  $\rho$ . Pure states are special cases of general density matrices.

Using:

#### 4. Change of basis:

• Transformation of the wave function:

$$\psi \to \psi' = D\psi, \qquad DD^{\dagger} = 1$$
 (1.60)

The change of basis is a unitary transformation.

$$\psi^{\dagger} \to \psi^{\prime \dagger} = \psi^{\dagger} D^{\dagger} \tag{1.61}$$

The norm is conserved under this transformation:

$$\psi^{\dagger}\psi^{\prime} = \psi^{\dagger}D^{\dagger}D\psi = \psi^{\dagger}\psi \tag{1.62}$$

• Transformation of the density matrix:

$$\varrho \to \varrho' = D \varrho D^{\dagger} \tag{1.63}$$

The usual properties of  $\rho$  apply:  $tr\rho' = 1, \rho'^t = \rho'$ , positivity of  $\rho'$ . Our density matrix formalism can be used in a random basis. In the next part we will see how this simplifies the probabilistic interpretation.

• Transformation of operators:

$$A \to A' = DAD^{\dagger} \tag{1.64}$$

This transformation law needs to hold for the expectation value to be invariant under change of basis:

$$\langle A \rangle = \psi^{\dagger} A^{\prime} \psi^{\prime} = \psi^{\dagger} D^{\dagger} D A D^{\dagger} D \psi = \psi^{\dagger} A \psi \tag{1.65}$$

Or in density matrix formulation:

$$\langle A \rangle = tr(\varrho'A') = tr(D\varrho D^{\dagger}DAD^{\dagger}) = tr(D\varrho AD^{\dagger}) = tr(D^{\dagger}D\varrho A = tr(\varrho A)$$
(1.66)

Since we are dealing with a unitary transformation, the spectrum of A, the set of eigenvalues, is independent under change of basis as well. This is another crucial property for the probabilistic interpretation.

#### 5. Probabilistic interpretation:

With introducing the density matrix formalism in quantum mechanics we should discuss its probabilistic interpretation. In order to check whether we are approaching classical statistics further.

For that assume an observable A, without loss of generality (as proven in the last part) choose the eigenbasis  $\{\alpha_j\}$  (eigenstates  $A_{\alpha_j}$ ) of A in the Hilbert space, such that the assigned operator A' becomes diagonal. The quantum law for expectation values then reads:

$$\langle A(t) \rangle = tr(\varrho'(t)A')$$

$$= \sum_{j} \alpha_{j}^{\dagger} \varrho'(t) A' \alpha_{j}$$

$$= \sum_{j} \alpha_{j}^{\dagger} \varrho'(t) \alpha_{j} A_{\alpha_{j}}$$

$$= \sum_{j} \varrho'_{jj}(t) A_{\alpha_{j}}$$

$$\Rightarrow \langle A(t) \rangle = \sum_{j} \varrho'_{jj}(t) A_{\alpha_{j}}$$

$$(1.67)$$

According to the classical law for expectation values we write:

$$\langle A(t) \rangle = \sum_{j} p_{\alpha_j}^{(A)}(t) A_{\alpha_j} \tag{1.68}$$

By comparison with the above equation we realise:

$$\varrho_{jj}'(t) = p_{\alpha_j}^{(A)}(t)$$
(1.69)

Given the eigenbasis of an operator A, the diagonal elements  $\varrho'_{jj}(t)$  of the density matrix can be interpreted as the probabilities  $p^{(A)}_{\alpha_j}(t)$  to measure the (eigen)value  $A_{\alpha_j}$  of the observable A.

This interpretation still holds for more than one observables if the observables commute. So there exists a common eigenbasis, which allows the interpretation of diagonal elements as probabilities for all observables. However, if we assume non-commuting observables, a problem arises. Non-commuting operators have no common eigenbasis. We can always find to bases such that:

$$\begin{split} \mathbf{A}^{'} & \text{diagonal} \Rightarrow \varrho_{jj}^{\prime}(t) = p_{\alpha_{j}}^{(A)}(t) \\ \mathbf{B}^{''} & \text{diagonal} \Rightarrow \varrho_{jj}^{\prime\prime}(t) = p_{\beta_{j}}^{(B)}(t) \end{split}$$

but if we choose the eigenbasis of one observable, the probabilistic information of the other observable clearly moves into the off-diagonal elements.

We conclude that the density matrix contains the probabilistic information for all observables, but only in the eigenbasis of a certain observable the interpretation of probabilities is possible. Choosing the eigenbasis of an observable makes the probabilistic information of other observables, non-commuting with this observable, inaccessible. The probabilistic information is then contained in the off-diagonal elements, but never lost.

#### 6. Incomplete statistics:

In the latter part we realised that, in quantum mechanics, there is a problem of incomplete statistics due to non-commuting observables. In the density matrix formalism it is not possible to calculate the probabilistic information of non-commuting observables simultaneously:

$$\nexists p_{\alpha\beta}^{(AB)} \,\forall A, B: \, [A, B] \neq 0 \tag{1.70}$$

But this means that we cannot determine correlations like in classical statistics:

$$\langle AB \rangle = \sum_{\alpha,\beta} p_{\alpha\beta}^{(AB)} A_{\alpha} B_{\beta} \tag{1.71}$$

At first this seems problematic, on the other hand there are also subsystems in classical statistics which yield incomplete statistics. If we want to show that quantum mechanics emerge from classical statistics, we will have to identify such subsystems.

## Chapter 2

## **Probabilistic** particles

In the last section we discussed the basic properties of quantum mechanics and classical statistics as probabilistic theories. By comparing both theories, we found common traits, but most importantly differences which need to be resolved in order to show that quantum mechanics emerge from classical statistics.

Quantum mechanics is a probabilistic theory, the dynamics of observables are described by probabilistic laws. On the other hand the quantum description of our world is based on matter and its interactions, namely by exchange of force carrying bosons. How can the probabilistic picture of quantum laws be connected with particles propagating through spacetime? In quantum field theory, the answer lies in the path integral formalism.

#### 1. Fermions on a lattice (1):

We now want to emphasise the path integral formalism with a discrete formulation. For simplification assume a d=1,1 spacetime and equidistant points (t,x). Like in figure 2.1. Further assume Fermions, the Pauli principle then demands that at a given point (t,x) the occupation number n(t,x) is either one or zero.

Let us begin with one Fermion on our lattice, then the occupation numbers at a fixed time  $\bar{t}$  need to satisfy:

$$\sum_{x} n(\bar{t}, x) = 1 \tag{2.1}$$

So the Fermion can only be at one position  $x(\bar{t})$  at a fixed time. This condition restricts the number of possible states from  $2^{M_x M_t}$  to  $M_x^{M_t}$ , where  $M_x, M_t$  denote the number of available positions and times. An event can then be written as:

$$\omega = \{n(t,x)\}\tag{2.2}$$

This event can also be interpreted to be the path the Fermion propagated along. But in contrast to classical mechanical trajectories, a path is not a tuple of a position at a given time  $(t, x) \rightarrow x(t)$ , but rather the set of all occupation numbers at given points in spacetime. This difference will become clear when we assume two Fermions.



Figure 2.1: Fermion on lattice (d=1,1 discrete spacetime)

#### 2. Time-local probability distribution:

Like in classical statistics, we can define time-local probability distributions in case of Fermions on a lattice. Actually, this is similar to the lattice discussed in (5), where we looked at a bit configuration. In fact, such "yes or no - questions" are sufficient to describe our reality, not only in quantum world, but also in our macroscopic world. Even though reducing a system to "yes or no - questions" may not be the best method to obtain simple descriptions.

Returning to Fermions on a lattice in spacetime, time-local probabilities (fixed time  $\bar{t}$ ) can be written as:

$$p(n(\bar{t}, x) = 1) = \prod_{t \neq \bar{t}} \sum_{x} p(n(\bar{t}, x) = 1; \{n(\bar{t}, x)\})$$
(2.3)

This results has already been calculated in (5), but the interpretation is new. The probability to find a Fermion occupying the state  $(\bar{t}, x)$  equals the sum of the probabilities of the overall occupation configurations where the state  $(\bar{t}, x)$  is occupied. So all occupation configurations, with the state  $(\bar{t}, x)$  being occupied, contribute to the time-local probability distribution of this state.

This interpretation can be expanded to the path integral formalism. In that case the probability for an initial state i to propagate into a final state f equals the sum of the probabilities of *all* possible paths, which yield the transition from the initial state to into the final state. Thus, all paths, not only the classical path, contribute to the transition probability. However, the probabilities are weighted proportional to a factor  $exp(iS/\hbar)$ . It turns out that the classical path is the most probable one.

Keep in mind that in quantum field theory the paths are processes with there is a higher number, which does not have to be not constant, of different particles, like Fermions and Bosons. The interpretations, otherwise, remains the same. We are interested in time-local probabilities, where all possible paths contribute.

#### 3. Fermions on a lattice (2):

In the case of a single Fermion the interpretation of a trajectory like in classical mechanics is still possible. To see the conceptual difference, let us assume two Fermions on our spacetime. Still, the occupation number of a point n(t,x) is either one or zero, so the Fermion cannot occupy the same point, but the condition (2.1) has changed:

$$\sum_{x} n(\bar{t}, x) = 2 \tag{2.4}$$

for a fixed time  $\bar{t}$ . In figure 2.2 a random event of two Fermions occupying the lattice is can be seen. In this case it is no longer clear what a classical trajectory should be. All that remains is a configuration of occupation numbers, which we denote as a path. We are left with a set of 0's and 1's or "yes or no - questions".



Figure 2.2: Fermions on lattice (d=1,1 discrete spacetime)

## Chapter 3

# Explanation of some "Quantum mysteries"

#### 3.1 Particle wave duality

What is the particle aspect of our system? We can define this property of our system as the possible measurement values of observables which are often discrete. The discreteness of these values are reflected on the discrete eigenvalues of the operators for observables. On the other hand we detect also, the wave aspect of our system through the probabilistic information and its time evolution, which are described by continuous quantities.

**Example 5** (Particle in 1d). We examine a particle at:

$$x > 0, n = 1$$
  
 $x < 0, n = 0$ 

The spectrum of the observable (in our case the existence of the particle in the different fractions of space) is (0,1). Of course, in this case the spectrum is discreet, since is just the values 0 and 1. This reflects the particle aspect of the system.

Now, for each possibility that discussed above, we can define the probabilities  $p_0$ ,  $p_1$  which are positive real numbers and obey the normalization relationship:

$$p_0 + p_1 = 1$$

So, the wave aspect will be the continuous nature of the probabilistic information. Of course, QM and CS have not exactly a priori the same formulation. For QM, probabilistic information is formulated through the density matrix  $\varrho(x, y)$  which is a matrix in position space but CS described by  $p(x) = \varrho(x, x)$ , which is a diagonal element of  $\varrho(x, y)$ .

The observable in this example is associated with an operator A(x, y) which is given by:

$$A(x,y) = \delta_{xy}\Theta(y) \tag{3.1}$$

So the expectation value can be calculated to be:

$$\langle A \rangle = \sum_{x,y} \varrho(x,y) A(x,y)$$

$$= \sum_{x} \varrho(x,x) \Theta(x)$$

$$\Rightarrow \langle A \rangle = \sum_{x>0} p(x)$$

$$(3.2)$$

The probability distribution is a completely real and continuous quantity and nevertheless the possible measurement values are discreet, as we showed above, the expectation values are then continuous quantities. It is important to keep in mind where we notice discreteness and continuity. There is no mystery in that!

The mysteries appeared if we try to consider the wave function or the density matrix as an observable. Many approaches, as Bohmian Quantum Mechanics, tried to do that but led to very complicated constructions. It is crucial to accept, that the probabilistic information which is contained either in wave function or in density matrix is not an observable!

#### 3.2 Entanglement

Usually entanglement is considered as central property of QM that cannot be found in CS. But as we will see below, entanglement is nothing more than a statement about the correlations. Correlations or correlation functions of course exists in CS.

We begin by two examples to work out that entanglement corresponds to correlations, which do also appear in classical statistics. We will see the correlations naturally appear if one goes from overall probability distributions to local probability distributions.

**Example 6** (Ising spins and quantum dices). We will restrict ourselves to situations with two local random variables:

#### 1. Ising spins

$$S^{(A)} = \pm 1; \ S^{(B)} = \pm 1$$
 (3.3)

Probability distribution for two Ising spins, there are four possibilities:

$$\{p_{\tau}\} = \{p_{++}, p_{+-}, p_{-+}, p_{--}\}$$
(3.4)

For the case of maximally anti-correlated state:

$$p_{++} = p_{--} = 0, \ p_{-+} = p_{+-} = \frac{1}{2}$$
 (3.5)

Two meanings of the word "state":

(a) state = microstate which refers to the exact spin configuration  $\tau$ 

(b) state = probability distribution  $\{p_{\tau}\}$  or  $\varrho$  or  $\psi$ , it is referred to the "state of the system". When we refer our state as maximally anti-correlated state, this is a statement about probability distribution.

Computation of the values of the spin observables  $S_{\tau}^{(A)}$  and  $S_{\tau}^{(B)}$ :

$$S_{++}^{(A)} = 1, \quad S_{+-}^{(A)} = 1, \quad S_{-+}^{(A)} = -1, \quad S_{--}^{(A)} = -1$$
 (3.6)

$$S_{++}^{(B)} = 1, \quad S_{+-}^{(B)} = -1, \quad S_{-+}^{(B)} = 1, \quad S_{--}^{(B)} = -1$$
 (3.7)

The next step is to compute the corresponding expectation values:

$$\langle S^{(A)} \rangle = p_{++} + p_{+-} - p_{-+} - p_{--} = 0 \tag{3.8}$$

$$\langle S^{(B)} \rangle = p_{++} - p_{+-} + p_{-+} - p_{--} = 0 \tag{3.9}$$

$$\langle S^{(A)}S^{(B)}\rangle = p_{++} - p_{+-} - p_{-+} + p_{--} = -1 \tag{3.10}$$

So, for the maximal anti-correlated state, whenever we measure  $S^{(A)} = 1$ , we will find  $S^{(B)} = -1$ , because the product must be -1 and vice versa, which means that if there are correlations we cannot separate the system into its parts. The whole is more than its parts.

These correlations can be perfectly non-local, there is no need for the correlations to be local. Non-local correlation may have a casual origin. For example  $S^{(A)}$  and  $S^{(B)}$  may be the product of a decay process of a spin-zero particle. Initially they are causally connected, where the correlations are created, but they could end up to very distant places, where they are not anymore causally connected. But the correlations are still there and they influence the measurements of these two particles without any signal propagation. The signal transmission is necessary only if we divide our system into its parts and of course these would be spatially separated. So, since that the system is treated as a whole, with the existence of the correlations, there is not any kind of signal with speed higher than the speed of light.

One can have different spin directions that one and all can be anti-correlated. An example could be to have three spin directions:

$$S_x^{(A)}, \quad S_y^{(A)}, \quad S_z^{(A)},$$
 (3.11)

$$S_x^{(B)}, \quad S_y^{(B)}, \quad S_z^{(B)}$$
 (3.12)

Which will result totally  $2^6$  microstates. Maximally anti-correlated state:

$$\langle S_r^{(A)} \rangle \langle S_l^{(B)} \rangle = 0, \quad \langle S_r^{(A)} S_l^{(B)} \rangle = -\delta_{rl}$$
(3.13)

To construct the probability distribution is leaving as an exercise.

#### 2. Quantum dices

Assume two dices which can each show the numbers  $i, j \in \{1, ..., 6\}$ . We can label the possible states  $\tau = (i, j)$ .

Now, we define two observables simply to be the numbers of the individual dices A and B:

$$A_{ij} = i; \quad B_{ij} = j \tag{3.14}$$

The interesting part now is the probability distribution of the dices. As opposed to "normal" dices with a uncorrelated distribution

$$p_{ij} = p_i \cdot p_j \tag{3.15}$$

we assume a correlated distribution

$$p_{ij} = \frac{1}{6}\delta_{ij} \tag{3.16}$$

Similarly to normal dices, the probability to find a certain number for one of the dices is  $\frac{1}{6}$ :

$$p(A=i) = \sum_{j} p_{ij} = \frac{1}{6}$$
(3.17)

$$p(B=j) = \sum_{i} p_{ij} = \frac{1}{6}$$
(3.18)

But if we now calculate the probability to find two equal numbers, we see the significant difference given by the correlated distribution:

$$p(i=j) = \sum_{i} \sum_{j=i}^{n} \frac{1}{6} \delta_{ij} = \sum_{i}^{n} p_{ii} = \sum_{i}^{n} \frac{1}{6} = 1$$
(3.19)

With a probability of 1, the two dices will show the same number. This is of course vastly different from our classical dices and simply caused by the fact that the probability is correlated and cannot be divided into 2 subsystems.

We conclude that quantum mechanical entanglement is basically described by correlated random variables, or correlated probability distributions. And correlated probability distributions are neither restricted to quantum mechanics nor "mystical" objects. Correlations just describe that there are systems in statistics with probability distributions which are not independent from each other.

However, the problem with entanglement in quantum mechanics arises when one tries to separate the two probability distributions. For correlated random variables:

$$\langle AB \rangle = \langle A \rangle \langle B \rangle \tag{3.20}$$

does simply not hold. But the possibility of separating the two random variables is what our intuition usually demands for, which is why entanglement seems problematic to us.

**Example 7** (EPR paradox). Even Albert Einstein tried to prove that quantum mechanics is not complete by claiming hidden variables. In cooperation with Podolsky and Rosen this claim resulted in the EPR paradox, which was based on quantum mechanical entanglement.

Let us assume a simplified version of the original EPR paradox. Assume a two spin system which is at one point in spacetime prepared such that:

$$\langle \hat{S}_i \hat{S}_j \rangle = (-1)\delta_{ij} \tag{3.21}$$

Hence, measuring one spin orientation in a random direction will yield measuring the second spin in the same direction with an opposite spin orientation.

Now separate the two spins such that the correlation, shown in equation (3.21), remains and a space-like separation is achieved. Measuring one spin orientation (in a random direction) would then immediately fix the spin orientation of the second spin in the same direction, even though the spins are space like separated and measuring one spin should not influence measuring the other spin. Where is the problem? The problem can be found in the assumption that both measurements influenced each other on a non-local basis, this is not true. This assumption is based on viewing both spins to be separated, which they are in spacetime, but the probability distributions of the spins are connected by the correlation. The system was prepared such that a measurement of one spin orientation fixes the spin orientation of the other one in the same direction. But it is important to note that what spin orientation. The correlation of both measurements will just become clear by a local interchange of the results.

#### 3.3 Phases in quantum dynamics

In quantum dynamics, the quantities dealt with, such as the wave function  $\psi$ , the density matrix  $\rho$  or the operators, are complex, in general. Therefore, their phases play an important role, for example in the superposition of 2 possible wave functions. This importance is the reason for statements like "Quantum mechanics is the physics of phases". At first glance, phases do not appear in classical statistics and seem to be a particularity of quantum dynamics. In the following, we will however discover that the complex notation can be perfectly replaced by a more classical, real description.

#### 1. Real quantum mechanics:

We want to map the complex wave function  $\psi$  and operator A to a real representation, which is isomorphic in terms of all relevant combinations of wave functions and operators.

We start with a wave function  $\psi \in \mathbb{C}^M$ . Its component  $\psi_{\alpha}$  can be written in terms of a real and an imaginary part:

$$\psi_{\alpha} = \psi_{R,\alpha} + i \cdot \psi_{I,\alpha} , \qquad \psi_{R,\alpha}, \psi_{I,\alpha} \in \mathbb{R}$$
(3.22)

We now map this vector onto a real vector  $q \in \mathbb{R}^{2M}$ :

$$\psi \to q = \begin{pmatrix} \psi_{R,\alpha} \\ \psi_{I,\alpha} \end{pmatrix}$$
(3.23)

For a complex M×M operator, we similarly decompose its components into real and imaginary part. Since all operators representing observables need to be hermitian  $A^{\dagger} = A$ , the real and imaginary parts have to be symmetric, or antisymmetric, respectively:

$$A_R^T = A_R, \quad A_I^T = -A_I \tag{3.24}$$

Now, the complex operator A is mapped onto a real  $2M \times 2M$  matrix

$$A \to \hat{A} = \begin{pmatrix} A_R & -A_I \\ A_I & A_R \end{pmatrix} = \mathbb{1} \otimes A_R + I \otimes A_I$$
(3.25)

where  $\mathbb{1}$  is the unity matrix and  $I = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ . Due to the condition in equation (3.24), it becomes clear that  $\hat{A}$  is symmetric.

Thus, we found a mapping onto real-valued entities. For this mapping to be useful, it needs to be isomorphic regarding matrix-vector multiplication, matrix multiplication and calculation of expectation values.

#### • matrix-vector multiplication

For multiplication of the real entities, we obtain

$$\hat{A}q = \begin{pmatrix} A_R & -A_I \\ A_I & A_R \end{pmatrix} \begin{pmatrix} \psi_{R,\alpha} \\ \psi_{I,\alpha} \end{pmatrix} = \begin{pmatrix} A_R\psi_R - A_I\psi_I \\ A_I\psi_R + A_R\psi_I \end{pmatrix}$$
(3.26)

If we multiply the complex entities, on the other hand, we get

$$A\psi = (A_R + i \cdot A_I)(\psi_R + i \cdot \psi_I) = A_R\psi_R - A_I\psi_I + i \cdot (A_I\psi_R + A_R\psi_I)$$
(3.27)

$$\Rightarrow \hat{A}q = \begin{pmatrix} \Re \mathfrak{e}(A\psi) \\ \Im \mathfrak{m}(A\psi) \end{pmatrix}$$
(3.28)

We have therefore proven that the mapping is isomorphic regarding matrix-vector multiplication.

#### • Matrix multiplication

For the multiplication of 2 complex M×M matrices A and B we get

$$AB = (A_R + iA_I)(B_R + iB_I) = A_R B_R - A_I B_I + i(A_I B_R + A_R B_I)$$
(3.29)

If  $\hat{A}$  and  $\hat{B}$  are defined analogously to equation (3.25), their multiplication yields

$$\hat{A}\hat{B} = \begin{pmatrix} A_R & -A_I \\ A_I & A_R \end{pmatrix} \begin{pmatrix} B_R & -B_I \\ B_I & B_R \end{pmatrix} = \begin{pmatrix} A_R B_R - A_I B_I & -A_R B_I - A_I B_R \\ A_I B_R + A_R B_I & -A_I B_I + A_R B_R \end{pmatrix}$$
(3.30)  
$$= \begin{pmatrix} \Re \mathfrak{e}(AB) & -\Im \mathfrak{m}(AB) \\ \Im \mathfrak{m}(AB) & \Re \mathfrak{e}(AB) \end{pmatrix}$$
(3.31)

Note that the entries of the matrices  $\hat{A}$  and  $\hat{B}$  are matrices again. Still, the multiplication works alike, which can be seen when writing out single components.

#### • Expectation values

For the expectation value of an observable A, the complex notation gives:

$$\langle A \rangle = \psi^{\dagger} A \psi = (\psi_R^T - i\psi_I^T) (A_R + iA_I) (\psi_R + i\psi_I)$$

$$= (\psi_R^T - i\psi_I^T) (A_R \psi_R - A_I \psi_I + iA_R \psi_I + iA_I \psi_R)$$

$$= \psi_R^T A_R \psi_R - \psi_R^T A_I \psi_I + \psi_I^T A_R \psi_I + \psi_I^T A_I \psi_R$$

$$+ i \cdot [\underbrace{\psi_R^T A_R \psi_I - \psi_I^T A_R \psi_R}_{=0} + \underbrace{\psi_R^T A_I \psi_R}_{=0} + \underbrace{\psi_I^T A_I \psi_I}_{=0}]$$

$$(3.32)$$

In the last step, the symmetry of  $A_R$  and the antisymmetry of  $A_I$  are exploited. This can be compared to the expression with real entities

$$q^{T}\hat{A}q = (\psi_{R}^{T},\psi_{I}^{T}) \begin{pmatrix} A_{R}\psi_{R} - A_{I}\psi_{I} \\ A_{I}\psi_{R} + A_{R}\psi_{I} \end{pmatrix} = \psi_{R}^{T}A_{R}\psi_{R} - \psi_{R}^{T}A_{I}\psi_{I} + \psi_{I}^{T}A_{I}\psi_{R} + \psi_{I}^{T}A_{R}\psi_{I}$$

$$\Rightarrow q^{T}\hat{A}q = \langle A \rangle$$
(3.33)

We have thereby proven that all complex wave functions and operators can be isomorphically mapped onto real entities, which naturally do not have phases.

#### 2. Complex structures:

If we want to find an inverse mapping

$$q \in \mathbb{R}^{2M} \to \psi \in \mathbb{C}^M$$

we notice that the complex notation is only possible for a certain class of observables: A real  $2M \times 2M$  matrix  $\hat{A}$  needs to have a representation with two  $M \times M$  matrices  $A_R$  and  $A_I$  equivalent to equation (3.25) such that the transformation

$$\hat{A} = \mathbb{1} \otimes A_R + I \otimes A_I \to A_R + iA_i \tag{3.34}$$

is well-defined. More generally, one can say that the complex formulation is only possible for operators represented by matrices which are compatible with the complex structure.

For a real vector  $q \in \mathbb{R}^{2M}$ , a general complex structure is defined by a tuple of discrete transformations (K, I) with the following properties:

$$K^{2} = +1$$
  
 $I^{2} = -1$   
 $\{K, I\} = 0$  (3.35)

In our case, we use the definition of I above

$$I = \left(\begin{array}{cc} 0 & -1\\ 1 & 0 \end{array}\right) \tag{3.36}$$

and an operator

$$K = \left(\begin{array}{cc} 1 & 0\\ 0 & -1 \end{array}\right). \tag{3.37}$$

With our definition of q above, the transformation K is equivalent to complex conjugation. For an appropriate mapping of q onto  $\mathbb{C}^M$ , we must therefore demand

$$q \to \psi \iff Kq \to \psi^*$$
 (3.38)

A real  $2M \times 2M$  operator  $\hat{A}$  is compatible with the complex structure if the condition in equation(3.34) is fulfilled. We further notice that a multiplication of a real vector q with I corresponds to a multiplication of a complex vector with i:

$$Iq \to i\psi$$
 (3.39)

To summarize, we can state that all complex matrices and wave functions can be expressed by real entities. The complex formulation of QM is therefore not conceptually different, but already uses the existence of additional complex structures K and I. Often, we will work with real discrete quantum dynamics. Classical quantum dynamics are a particular case for which

- a) a complex structure exists and
- b) the continuum limit for a time evolution exists.

At last we should note that both the complex and the real formulation of quantum mechanics are representations of quantum mechanics in certain spaces. Quantum mechanics are defined by the time evolution law and the probabilistic formalism, including wave functions and expectation values. Thus, the mathematical representation of quantum mechanics is in principle not restricted to a certain space. All spaces satisfying the fixed properties are allowed. Recall that for physics only expectation values, which are independent of the representation, matter. Again, objects like wave functions have no physical meaning.

#### 3.4 Planck's constant

In quantum mechanics, Planck's constant  $\hbar$  is encountered all the time. This raises the question whether an equivalent to it should appear in classical statistics. To answer this, we first have a look on the role of  $\hbar$  in quantum dynamics.

For instance,  $\hbar$  is part of the momentum operator:

$$P = -i\hbar\partial_x \tag{3.40}$$

and also appears in the commutator of the position operator X and the momentum operator:

$$PX = -i\hbar - i\hbar x \partial_x; \quad XP = -i\hbar x \partial_x$$
$$\Rightarrow [P, X] = -i\hbar \tag{3.41}$$

According to equation (3.40), we can understand  $\hbar$  simply as a conversion factor of units. Or, in the case of the commutation relation of position and momentum, as a measure for the smallest possible area in phase space (p,q). Moreover, we could even set it to 1 and measure momentum in units of inverse length, which is called natural units and is the basis for the new international unit convention. Similarly, the velocity of light can be set to 1. Our new momentum operator therefore is

$$P = -i\partial_x \tag{3.42}$$

$$[P,X] = -i \tag{3.43}$$

In the classical limit of quantum mechanics, the effects of commutators are neglected as the action S is assumed to be much larger than  $\hbar$  ( $|S| >> \hbar$ ). This corresponds to the limit of  $\hbar \to 0$ . However, we will *not* use this limit and instead set  $\hbar$  to 1.

### Chapter 4

# Ising chain and wave functions in classical statistics

In section 1.2 it has been mentioned that in general it is not possible to determine a time evolution law. So we are in general not able to calculate the probability distribution at time  $t_2$  given the probability distribution at a time  $t_1 < t_2$ . In contrast quantum mechanics admits a universal time evolution law, described by the time evolution operator. By that the time evolution of the wave function  $\psi(t)$  and therefore the probability distribution  $|\psi(t)|^2$  is fixed. In the current section we want to assume an explicit system, the 1d Ising model, and emphasise that there exist wave functions in classical statistics which yield a time evolution law.



Figure 4.1: 1d Ising model

#### 4.1 Ising chain

The 1d Ising model is a 1d lattice with lattice sites  $m \in \{0, \ldots, M\}$  called Ising spins  $s(m) = \pm 1$ , as visualised in figure (4.1). The 1d Ising model describes a lattice with binary lattice sites. The interaction of the lattice sites will be introduced shortly. Let us begin with notations and definitions. The spin configurations of the 1d Ising model are denoted  $s = \{s(0), s(1), \ldots, s(M)\}$ . For the

configuration sum we write:

$$\prod_{m} \sum_{s(m)=\pm 1} = \sum_{s(0)=\pm 1} \dots \sum_{s(M)=\pm 1} \equiv \int \mathscr{D}s$$
(4.1)

Note that the right side of the equation is typically used to denote path integrals. Since the configurations can be viewed as the discrete limit of a path integral, we may stick to this notation. Ising spins are binary systems, it is possible to transform from the "spin up - spin down" representation into the occupation number representation by:

$$n(m) \equiv n[s(m)] = \frac{s(m) + 1}{2}$$
(4.2)

$$\Rightarrow n[s(m) = +1] = 1 \tag{4.3}$$

$$n[s(m) = -1] = 0 \tag{4.4}$$

Both representations are equivalent in the sense that:

$$\int \mathscr{D}s = \int \mathscr{D}n \tag{4.5}$$

For the probabilistic formalism we need to know how the spin configurations s are weighted. In the canonical ensemble for example, one assigns a Boltzmann factor:

$$\exp\left(\frac{-E[s]}{k_B T}\right) \tag{4.6}$$

where E[s] is the energy corresponding to the spin configuration s. The probability for a certain spin configuration to occur then is:

$$p[s] = \frac{1}{Z} \exp\left(\frac{-E[s]}{k_B T}\right) \tag{4.7}$$

with the partition function:

$$Z = \int \mathscr{D}s \, \exp\left(\frac{-E[s]}{k_B T}\right) \tag{4.8}$$

The generalisation is obvious, we claim a general weight distribution w[s]  $(w[s] \ge 1 \forall s)$  for the spin configurations. Given this weight distribution, the partition function Z then reads:

$$Z = \int \mathscr{D}s \ w[s] \tag{4.9}$$

Thus, the probability distribution p[s] is:

$$p[s] = \frac{1}{Z}w[s] \tag{4.10}$$

$$\int \mathscr{D}s \ p[s] = 1 \tag{4.11}$$

To introduce Ising spin interactions define the weight distribution to be:

$$w[s] = \tilde{f}_f[s(M)] \prod_{m=0}^{M-1} \mathscr{K}(m) f_i[s(0)]$$
(4.12)

with the boundary conditions  $\tilde{f}_i[s(0)]$ ,  $\tilde{f}_f[s(M)]$ , which fix the Ising spins at 0 and M, and the local factors  $\mathscr{K}(m) = K[s(m), s(m+1)]$  describing a next neighbour interaction.



Figure 4.2: 1d Ising model: local factors

It is further convenient to define the action S:

$$S[s] = \sum_{m=0}^{M-1} \mathscr{L}(m) = \sum_{m=0}^{M-1} \mathscr{L}[s(m), s(m+1)]$$
(4.13)

Where the local factors  $\mathscr{K}$ , describing the interaction of two neighbouring Ising spins, are related to the  $\mathscr{L}(m)$  by claiming:

$$\mathscr{K}(m) = \exp(-\mathscr{L}(m)) \ge 0 \tag{4.14}$$

which yields:

$$w[s] = \tilde{f}_f[s(M)] \exp\{-S[s]\} f_i[s(0)]$$
(4.15)

This reformulation seems to be arbitrary, but equation (4.15) states that the weight of a Ising spin configuration is essentially determined by the term exp(-S). Up to a factor of i this coincides with the weight of a path in the path integral formalism.

At last the expectation value for an observable A, with values A[s] for a given Ising spin configuration, can be expressed like before:

$$\langle A[s] \rangle = \frac{1}{Z} \int \mathscr{D}s \ \tilde{f}_f[s(M)] \ \exp(-S[s(m)]) \ f_i[s(0)] \ A[s]$$

$$(4.16)$$

$$= \int \mathscr{D}s \ p[s]A[s] \tag{4.17}$$

Note that the probabilistic formulation of the 1d Ising model always depends on the boundary conditions  $\tilde{f}_f[s(M)]$ ,  $f_i[s(0)]$ . Given a set of boundary conditions then fixes the probabilities and the classical laws like the law for expectation values hold.

**Example 8** (Phase transition). Let us now use our formalism for the 1d Ising model and discuss its phase transition, which is a typical problem for Ising models. For calculating that the configuration sum will be of major importance, which might make it a more intuitive quantity.

Assume that two neighbouring Ising spins interact by:

$$\mathscr{L} = -\beta s(m+1)s(m) + \beta, \ \beta > 0 \tag{4.18}$$

In terms of local factors  $\mathscr{K}(m)$  this means:

$$\mathscr{K}(m) = \begin{cases} 1, \ s(m+1)s(m) = +1\\ \exp(-2\beta), \ s(m+1)s(m) = -1 \end{cases}$$
(4.19)

Under this assumption the Ising spin interactions favor spin configurations where the neighbouring spins are aligned. By defining:

$$\beta = \frac{U}{k_B T}, \ U > 0 \tag{4.20}$$

one can argue for a phase transition/spontaneous symmetry breaking at temperature T=0. The terms phase transition/spontaneous symmetry breaking denote a situation where the Ising spin system spontaneously polarises itself (without external influence):

$$\left\langle \sum_{m=0}^{M} s(m) \right\rangle \neq 0 \tag{4.21}$$

From now on<sup>1</sup> neglect the boundary conditions, we will observe the Ising chain in the limit  $M \to \infty$ . With that the Ising spin correlation reads:

$$\langle s(i)s(i+j) \rangle = \frac{1}{Z} \int \mathscr{D}s \ s(i)s(j) \ \exp(-S[s])$$

$$= \frac{\exp(-M\beta)}{Z} \sum_{s(0)=\pm 1} \dots \sum_{s(M)=\pm 1} s(i)s(i+j) \prod_{m=0}^{M-1} \exp[\beta s(m)s(m+1)]$$

$$= (\dots) \sum_{s(0)=\pm 1} \dots \sum_{s(M)=\pm 1} s(i)s(i+j)s(i)s(i+1)s(i+1)\dots s(i+j-1)s(i+j-1)s(i+j) (\dots)$$

$$(4.22)$$

To evaluate this expression we write  $\mathscr{L}(m) = \beta_m s(m) s(m+1)$  and use the trick:

$$s(i)s(i+j)s(i)s(i+1)s(i+1)\dots s(i+j-1)s(i+j-1)s(i+j)\prod_{m=0}^{M-1}\exp[\beta_m s(m)s(m+1)]$$
  
=  $\partial_{\beta_i}\dots\partial_{\beta_{i+j-1}}\prod_{m=0}^{M-1}\exp[\beta_m s(m)s(m+1)]$  (4.23)

<sup>&</sup>lt;sup>1</sup>This part was not covered in the lecture, but it has didactic value working with the configuration sum. It is based on the corresponding chapter in the "Theoretical Statistical Physics" course of Prof. Ulrich Schwarz from the winter term 2020/21 [https://www.thphys.uni-heidelberg.de/~biophys/PDF/Skripte/StatPhys.pdf].

Commuting the configuration sum and the partial derivatives then results in:

$$\langle s(i)s(i+j) \rangle = \frac{1}{Z} \int \mathscr{D}s \ s(i)s(j) \ \exp(-S[s])$$

$$= \frac{\exp(-M\beta)}{Z} \ \partial_{\beta_i} \dots \partial_{\beta_{i+j-1}} \sum_{s(0)=\pm 1} \dots \sum_{s(M)=\pm 1} \prod_{m=0}^{M-1} \exp[\beta_m s(m)s(m+1)]$$

$$= (\dots) \sum_{s(0)=\pm 1} \dots \sum_{s(M-1)=\pm 1} \prod_{m=0}^{M-2} \exp[\beta_m s(m)s(m+1)] 2 \cosh(\beta_{M-1}s(M-1))$$

$$(4.24)$$

Using  $\cosh(\beta_{M-1} \pm 1) = \cosh(\beta_{M-1})$  then yields:

$$\sum_{s(0)=\pm 1} \dots \sum_{s(M-1)=\pm 1} \prod_{m=0}^{M-2} \exp[\beta_m s(m) s(m+1)] 2 \cosh(\beta_{M-1} s(M-1))$$
  
=  $2^M \prod_{m=0}^{M-1} \cosh(\beta_m)$  (4.25)

So we finally obtain:

$$\langle s(i)s(i+j)\rangle = \frac{1}{Z} \int \mathscr{D}s \ s(i)s(j) \ \exp(-S[s])$$
$$= \frac{\exp(-M\beta)}{Z} \ \partial_{\beta_i} \dots \partial_{\beta_{i+j-1}} \ 2^N \prod_{m=0}^{M-1} \cosh(\beta_m)$$
$$= \frac{2^M \exp(-M\beta)}{Z} \prod_{m=i}^{i+j-1} \sinh(\beta_m) \prod_{m=0, m\notin\{i\dots i+j-1\}}^{M-1} \cosh(\beta_m)$$
(4.26)

It remains to calculate the partition function Z, which is:

$$Z = \exp(-M\beta) \sum_{s(0)=\pm 1} \dots \sum_{s(M)=\pm 1} \prod_{m=0}^{M-1} \exp[\beta_m s(m) s(m+1)]$$
(4.27)

But this has already been done, the result is:

$$Z = 2^{M} \exp(-M\beta) \prod_{m=0}^{M-1} \cosh(\beta_m)$$
(4.28)

Inserting this into the Ising spin correlation gives us our final result:

$$\langle s(i)s(i+j)\rangle = \prod_{m=i}^{i+j-1} \tanh(\beta_m) = \tanh(\beta_m)^j$$
(4.29)

where  $\beta_m = \beta \ \forall m \in \{0, \dots, M-1\}$  was reenforced. The temperature emerges in  $\beta$ :

$$\langle s(i)s(i+j)\rangle = \left[ \tanh\left(\frac{U}{k_BT}\right) \right]^j$$
(4.30)

We observe that:

$$\langle s(i)s(i+j)\rangle = \begin{cases} 1, \ T=0\\ \alpha < 1, \ T\neq 0 \end{cases}$$

$$(4.31)$$

In the limit  $j \to \infty$  we expect:

$$\lim_{j \to \infty} \langle s(i)s(i+j) \rangle = \langle s(i) \rangle \langle s(j) \rangle = \langle s \rangle^2$$
(4.32)

using that the system is homogeneous  $\langle s(m) \rangle = \langle s \rangle \forall m$ . But then:

$$\left\langle \sum_{m} s(m) \right\rangle^2 = (M+1)^2 \langle s \rangle^2 = (M+1)^2 \lim_{j \to \infty} \langle s(i)s(i+j) \rangle = \begin{cases} (M+1)^2, \ T = 0\\ 0, \ T > 0 \end{cases}$$
(4.33)

We conclude that a phase transition/spontaneous symmetry breaking occurs only in the case T=0 for the 1d Ising model, in the limit  $M \to \infty$  without boundary conditions. The 2d Ising model already yields a phase transition at a finite temperature.

#### 4.2 Classical wave functions

Our longtime goal is showing that, by accepting that our world is fundamentally probabilistic, quantum mechanics emerge from classical statistics. For that reason we are interested in finding a time evolution law in classical statistics, which is one of our major problems. For that we introduced the 1d Ising model to demonstrate how to construct a time evolution law. In this section we will work out classical wave functions, including a (time) evolution law. However, this time evolution law will not be an evolution law for probability distributions, but only these wave functions. Later will construct the density matrix formalism of classical statistics, then a time evolution law similar to the one of quantum mechanics will arise.

#### 1. Local Probabilities:

To emphasise classical wave functions, let us return to the 1d Ising model. For an evolution law we need local probability distributions. Recall that an evolution law would connect these local probability distributions. Such that given one local probability distribution another could be calculated. In case of the 1d Ising mode we have:

$$p(m) = p[s(m) = \pm 1] = \prod_{m' \neq m} \sum_{s(m') = \pm 1} p[s, s(m) = \pm 1]$$
(4.34)

For completeness, given a local observable A(m) one can derive the expectation value:

$$\langle A(m) \rangle = \sum_{s(m)=\pm 1} p(m)A(m) \tag{4.35}$$
But given a local probability distribution p(m), it is **not** possible to determine the probability distribution p(m+1) by some kind of evolution law. There is no dependence of p(m-1)implemented in p(m), it is "annihilated" by the configuration sum. We will have to find other objects than local probability distributions to find a general time evolution law.

**Example 9** (Local probability distributions).  $\langle S(m) \rangle = \sum_{s(m)=\pm 1} p_+(m) - p_-(m)$  with  $p(m) = \prod_{m' \neq m} \sum_{s(m')=\pm 1} p[s]$ 

The Evolution law should compute p(m+1) as a function of p(m). If it exists it should start with p(0) and follow the evolution law to compute p(m) and solve the initial value problem. It is however important that the evolution law in terms of local probabilities does not exist! We will now establish an evolution law in terms of the wave functions or the density matrix.

Now to the classical wave function! A difference to QM is that there is a pair of wave functions in CS. A wave function and its conjugate wave function (in QM the conjugate wave function is the complex conjugate of the wave function):

$$f(m) = f[s(m)]$$
  
$$\bar{f}(m) = \bar{f}[s(m)]$$

Those wave functions depend only on the local spins s(m), similarly to  $p_+$  and  $p_-$ . We will now define this pair of wave functions in terms of our overall weight distribution. Therefore we have to remember that the local probabilities are defined as integrating over all spins different from m. Now we will only integrate over the spins smaller than m. The definition of  $\tilde{f}(m)$  is then the following:

$$\tilde{f}(m) = \prod_{m' < m} \sum_{s(m') = \pm 1} \exp(-S_{<}) f_i[s(0)]$$
(4.36)

With  $S_{\leq}$  being :

$$S_{<} = \sum_{m'=0}^{m-1} \mathscr{L}(m')$$
(4.37)

And of course the exponential function of  $S_{<}$  is:

$$\exp(-S_{\leq}) = \prod_{m' \leq m} \mathscr{K}(m) \tag{4.38}$$

The wave function "integrates out the lower region of the spins" and therefore depends only on the spin s(m). Here the integral corresponds to the sum over the configurations  $\{s(m' < m)\}$ . (This variable appears in  $\mathcal{K}(m-1)$  and is not integrated out.) In a similar way the conjugate wave function is defined, just by "integrating out the upper region of the spins":

$$\bar{f}(m) = \prod_{m' > m} \sum_{s(m') = \pm 1} \tilde{f}_f[s(M)] \exp(-S_{>})$$
(4.39)

with:

$$S_{>} = \sum_{m'=m}^{m-1} \mathscr{L}(m')$$
 (4.40)

$$exp(-S_{>}) = \prod_{m' \ge m} \mathscr{K}(m')$$
(4.41)

This means consequently, that the conjugate wave function "integrates out the upper region of he spins". As before  $\bar{f}(m)$  depends on s(m). s(m) appears in  $\mathscr{K}(m)$  and is not integrated.



Figure 4.3: Classical wave functions

Given the definition of the two classical wave functions, we can now try to formulate our probabilistic theory with those. First observe:

$$S_{<} + S_{>} = S$$
 (4.42)

But then, by construction of the two classical wave functions above, we obtain a prior result:

$$\bar{f}(m)\tilde{f}(m) = \left(\prod_{m'>m}\sum_{s(m')=\pm 1}\right) \left(\prod_{m'>m}\sum_{s(m')=\pm 1}\right) \tilde{f}_f \exp(-S) f_i$$
$$= \left(\prod_{m'\neq m}\sum_{s(m')=\pm 1}\right) \tilde{f}_f \exp(-S) f_i = w(m) = Z \cdot p(m)$$
$$\Rightarrow p(m) = \frac{1}{Z} \bar{f}(m)\tilde{f}(m)$$
(4.43)

which is somewhat similar to the term obtained in quantum mechanics. Just for two wave functions, and with the factor 1/Z. From that conclude the expectation value of local observables to be:

$$\langle A(m)\rangle = Z^{-1} \sum_{s(m)=\pm 1} \bar{f}(m) A(m) \tilde{f}(m)$$
(4.44)

Note that by multiplicative rescaling of  $\bar{f}_f$ ,  $f_{in}$  one can achieve Z=1, to obtain the expected result:

$$p(m) = \bar{f}(m)\tilde{f}(m)$$
(4.45)

#### 2. Evolution law:

Given the local factors  $\mathscr{K}(m)$ , which describe the interaction at neighbouring local points, the classical wave functions at m and m+1 are related by:

$$\tilde{f}(m+1) = \sum_{s(m)=\pm 1} \mathscr{K}(m)\tilde{f}(m)$$
(4.46)

One does simply include the spin s(m) in the "lower region" to obtain the wave function  $\tilde{f}(m+1)$ . In order to obtain the wave function  $\bar{f}(m-1)$  one subtracts the spin s(m):

$$\bar{f}(m-1) = \sum_{s(m)=\pm 1} \mathscr{K}(m-1)\bar{f}(m)$$
 (4.47)

Observe that the two evolution laws are linear in  $\tilde{f}(m)$  or  $\bar{f}(m)$  respectively. Similar to the Schrödinger equation in quantum mechanics. And so our theory yields a superposition law.

#### 3. Occupation number basis

Next let us define a basis to get a representation of our abstract wave functions. It is convenient to choose the occupation number basis. For that introduce the two (local) basis functions:

$$h_0(s) = \frac{1}{2}(1 - s(m)) = 1 - n[s(m)]$$
(4.48)

$$h_1(s) = \frac{1}{2}(1+s(m)) = n[s(m)]$$
(4.49)

The label "occupation number basis" is clear looking at the occupation numbers n[s(m)] assigned to a spin configuration s(m):

$$h_0(1) = 0, \ h_0(0) = 1, \ h_1(1) = 1, \ h_1(0) = 0$$

Then, expand the wave functions  $\tilde{f}(m)$  in terms of these basis functions  $h_{\tau}[s(m)], \tau = 0, 1$ :

$$\tilde{f}(m) \equiv \tilde{f}[s(m)] = \tilde{q}_{\tau}(m) \ h_{\tau}[s(m)] = \tilde{q}_{0}(m) \ h_{0}[s(m)] + \tilde{q}_{1}(m) \ h_{1}[s(m)]$$
(4.50)

Thus,  $\tilde{f}(m)$  is represented by:

$$\tilde{q}(m) = \begin{pmatrix} \tilde{q}_1(m) \\ \tilde{q}_0(m) \end{pmatrix}$$
(4.51)

in the chosen occupation number basis. The same applies for the conjugate wave function:

$$\bar{f}(m) = \bar{f}[s(m)] = \bar{q}_{\tau}(m) h_{\tau}[s(m)]$$
(4.52)

$$\bar{q}^T(m) = (\bar{q}_1(m), \bar{q}_0(m))$$
(4.53)

It remains to calculate the local probabilities in the occupation number basis. Representing the local probabilities in the occupation number basis reads:

$$p(m) = p[s(m)] = p_{\tau}(m) h_{\tau}[s(m)]$$
(4.54)

Define:

$$p[s(m) = 1] = p_1(m) = p_+(m), \ p[s(m) = -1] = p_0(m) = p_-(m)$$
(4.55)

We know the local probabilities expressed by the classical wave functions in the abstract sense. In the occupation number basis this becomes:

$$p[s(m)] = f[s(m)]f[s(m)] = q_{\tau}(m) \ h_{\tau}[s(m)] \ \tilde{q}_{\rho}(m) \ h_{\rho}[s(m)]$$
(4.56)

We use:

$$h_{\tau}(s(m))h_{\rho}(s(m)) = \delta_{\tau\rho}h_{\tau}(s(m)) \tag{4.57}$$

The proof is simple by just using the definition of our occupation number basis:

$$h_1^2 = n^2 = n = h_1, \ h_0^2 = (1-n)^2 = (1-n) = h_0, \ h_1 h_0 = n(1-n) = 0$$

And so we can calculate:

$$p[s(m)] = \sum_{\tau,\rho} \bar{q}_{\tau}(m) \tilde{q}_{\rho}(m) \delta_{\rho\tau} h_{\tau}(m)$$
$$= \sum_{\tau} \bar{q}_{\tau}(m) \tilde{q}_{\tau}(m) h_{\tau}(m)$$
$$= \sum_{\tau} p\tau(m) h_{\tau}(m)$$
(4.58)

By comparing the components of the local probability in terms of occupation number basis to the one of the direct computation, we conclude:

$$p_{\tau}(m) = \bar{q}_{\tau}(m)\tilde{q}_{\tau}(m)$$
(4.59)

It is important to remember that we do not sum over  $\tau$  here. The normalisation still holds:

$$\sum_{\tau} \bar{q}_{\tau}(m) \tilde{q}_{\tau}(m) = 1$$

Once more, recall the similarity to QM:  $p_{\tau}(m) = \Psi_{\tau}^{\dagger}(m)\Psi_{\tau}(m)$ . However, there are still problems to solve. The classical wave functions  $\bar{q}$ ,  $\tilde{q}$  are real, and  $\bar{q}$  is in general not equal to  $\tilde{q}^{\dagger}$ .

Anyway, let us now formulate the evolution law in terms of  $\tilde{q}(m)$  and  $\bar{q}(m)$ . An important object will be the so called transfer matrix.

# 4.3 Transfer matrix and evolution law for wave functions

The transfer matrix for the one dimensional Ising chain is simply given by an expansion of the local factor  $\mathscr{K}$  in our basis functions  $h_{\tau}(m)$ . Since it depends on both s(m) and s(m+1), we need a double expansion:

$$\mathscr{K}(m) = \mathscr{K}[s(m+1), s(m)] = T_{\tau\rho}(m) h_{\tau}(s(m+1)) h_{\rho}(s(m))$$
(4.60)

The  $T_{\tau\rho}$  are the components of the 2x2 transfer matrix. Now, we can formulate an evolution law for the classical wave function. The wave function at the point m+1 is connected to the one at m through the local factor  $\mathscr{K}(m)$ :

$$\tilde{f}(m+1) = \sum_{s(m)} \mathscr{K}(m) \tilde{f}(m) = \sum_{s(m)} T_{\tau\rho}(m) h_{\tau}[s(m+1)] h_{\rho}[s(m)] \tilde{q}_{\sigma}(m) h_{\sigma}[s(m)] = T_{\tau\rho}(m) \tilde{q}_{\rho}(m) \sum_{s(m)} h_{\tau}[s(m+1)] h_{\rho}[s(m)]$$
(4.61)

In the next step we use

$$\sum_{s} h_{\rho}(s) = 1 \tag{4.62}$$

which can easily be shown for our basis functions in the occupation number basis:

$$\sum_{s} h_0(s) = \sum_{s} \frac{1}{2}(1-s) = 1$$
$$\sum_{s} h_1(s) = \sum_{s} \frac{1}{2}(1+s) = 1$$

Therefore, we find:

$$f(m+1) = T_{\tau\rho}(m) \, \tilde{q}_{\rho}(m) \, h_{\tau}[s(m+1)] \\ \stackrel{!}{=} \tilde{q}_{\tau}(m+1) \, h_{\tau}(s(m+1))$$
(4.63)

which yields the evolution law:

$$\overline{\tilde{q}_{\tau}(m+1) = T_{\tau\rho}(m) \ \tilde{q}_{\rho}(m)}$$

$$(4.64)$$

or in vector notation:

$$\tilde{q}(m+1) = T(m) \ \tilde{q}(m) \tag{4.65}$$

For the conjugated wave function we obtain the analogue evolution law:

$$\bar{q}_{\tau}(m-1) = \bar{q}_{\rho}(m)T_{\rho\tau}(m-1)$$
(4.66)

Let us use the newly developed formalism on the Ising chain, another step in direction of solving the initial value problem of the Ising chain.

**Example 10** (Transfer matrix of the Ising chain:). Since the Ising chain is a simple but important system in Classical Statistics, we will now see how to calculate its transfer matrix. Nevertheless,

the same steps can be used to find the transfer matrix of any system.

We start with the local factor  $\mathscr{K}(m)$  and expand it in the local basis functions. Since K depends on s(m) and s(m+1), we also need two basis functions:

$$\mathscr{K}(m) = \exp(-\mathscr{L}(m)) = \exp\{\beta[s(m+1)s(m) - 1]\}$$
(4.67)

$$= T_{\tau\rho}h_{\tau}(s(m+1))h_{\rho}(s(m))$$
(4.68)

Proceed in the following steps:

1. **Express**  $\mathcal{L}$  in the occupation number basis:

$$\mathscr{L}(m) = \mathscr{L}_{\tau\rho}(m) h_{\tau}[s(m+1)] h_{\rho}[s(m)]$$
(4.69)

With this, we can find an expression for the coefficients  $T_{\tau\rho}$  by expanding the exponential and using the orthonormality of the basis functions:

$$\exp(-\mathscr{L}(m)) = 1 - \mathscr{L}_{\tau\rho}h_{\tau}(s(m+1))h_{\rho}(s(m)) + \frac{1}{2} \underbrace{\mathscr{L}_{\tau\rho}h_{\tau}(s(m+1))h_{\rho}(s(m))\mathscr{L}_{\alpha\beta}h_{\alpha}(s(m+1))h_{\beta}(s(m))}_{=\mathscr{L}_{\tau\rho}\mathscr{L}_{\alpha\beta}h_{\tau}(m+1)h_{\rho}(m)\delta_{\tau\alpha}\delta_{\delta\beta} = (\mathscr{L}_{\tau\rho})^{2}h_{\tau}(m+1)h_{\rho}(m)} = 1 + \sum_{\tau\rho} \left[ -\mathscr{L}_{\tau\rho} + \frac{1}{2}\mathscr{L}_{\tau\rho}^{2} - \frac{1}{6}\mathscr{L}_{\tau\rho}^{3} + \dots \right] h_{\tau}(m+1)h_{\rho}(m) = 1 + \left( \sum_{\tau\rho} \exp(-\mathscr{L}_{\tau\rho}) - 1 \right) h_{\tau}(m+1)h_{\rho}(m) = \sum_{\tau\rho} \exp(-\mathscr{L}_{\tau\rho})h_{\tau}(m+1)h_{\rho}(m)$$

$$(4.70)$$

By comparison with (4.67) we get

$$T_{\tau\rho} = exp(-\mathscr{L}_{\tau\rho})$$
(4.71)

#### 2. Compute $\mathscr{L}_{\tau\rho}$

It remains to explicitly find the components of  $\mathscr{L}$  in the occupation number basis. We do this by expressing the spin in terms of the occupation number n and then substituting our definition of the basis functions:

$$\mathscr{L} = -\beta s(m+1)s(m) + \beta \tag{4.72}$$

$$= -\beta \{ (2n(m+1) - 1)(2n(m) - 1) - 1 \}$$
(4.73)

$$= -\beta \left\{ \left( n(m+1) - 1 + n(m+1) \right) \left( n(m) - 1 + n(m) \right) - 1 \right\}$$
(4.74)

$$= -\beta \{ (h_1(m+1) - h_0(m+1))(h_1(m) - h_0(m)) - (\underbrace{h_1(m+1) + h_0(m+1)}_{=1})(\underbrace{h_1(m) + h_0(m)}_{=1}) \}$$

By representing the 1 as  $(h_0(m) + h_1(m))(h_0(m+1) + h_1(m+1))$ , we can see that all squared terms drop out and only the terms which mix m and m+1 stay.

$$\mathscr{L} = 2\beta \{h_1(m+1)h_0(m) + h_0(m+1)h_1(m)\}$$
(4.76)

We therefore obtain

$$\mathscr{L}_{00} = \mathscr{L}_{11} = 0, \qquad \qquad \mathscr{L}_{10} = \mathscr{L}_{01} = 2\beta$$
(4.77)

and can write the transfer matrix of the Ising chain as

$$T = \begin{pmatrix} 1 & e^{-2\beta} \\ e^{-2\beta} & 1 \end{pmatrix}$$
(4.78)

Next we will deal with the connection between transfer matrix and partition function. So far, we have used the normalization of the wave function and of the local factors that corresponds to the partition function Z=1. So, we have to go back to the unnormalized quantities in order to extract the partition function Z, which means that we have no longer normalized wave functions and particular normalization for the local factors. Up the normalization the classical wave functions are defined in the same way as before:

$$\tilde{q}(m) = T \ \tilde{q}(m-1) = T^2 \ \tilde{q}(m-2)\dots 
= T^m \tilde{q}(0)$$
(4.79)

For the final time M:

$$\tilde{q}(M) = T^M \tilde{q}(0) \tag{4.80}$$

Now, we can immediately write the expression for the probability:

$$w_{\tau}(M) = T^{M}_{\tau\rho} \,\tilde{q}_{\rho}(0) \,\tilde{q}_{\tau}(M) \tag{4.81}$$

It is clear that, if would have normalized the wave functions,  $w_{\tau}$  would be local probabilities. By now, the  $w_{\tau}$  are not normalized, for that we need the partition function Z.

The product of the initial and the final wave function can be written in terms of the boundary matrix  $\overline{B}$ :

$$\tilde{q}_{\rho}(0) \ \tilde{q}_{\tau}(M) \equiv \bar{B}_{\rho\tau} \tag{4.82}$$

Writing the  $w_{\tau}(M)$  as a trace we get:

$$w_{\tau}(M) = \sum_{\rho} \left( T^M \right)_{\tau\rho} \bar{B}_{\rho\tau} \tag{4.83}$$

With that we can enforce the normalization by defining the partition function:

$$Z = \sum_{\tau} w_{\tau}(M) = tr\{T^M \cdot \bar{B}\}$$
(4.84)

We observe that the normalization by the partition function Z is dependent on the boundary conditions. Let us choose an explicit form and assume the Ising chain periodic which means  $\overline{B} = T$ . Then the partition function reads:

$$Z = tr\{T^{M+1}\}$$
(4.85)

The above relationship gives us the expression of the canonical partition function with respect to transfer matrix, which provides a very elegant method for solving the partition function. We can now go to a basis of the eigenstates of the transfer matrix, in which it would be diagonal, so the trace is pretty simple to compute. Having computed the canonical partition function Z, we could determine the entropy, mean energy, etc. of the system.

# 4.4 Step evolution operator

To take one closer to quantum mechanics we will now choose a particular normalization of the transfer matrix, to define the so called step evolution operator  $\hat{S}$ . The reason to choose a particular normalization arises from the fact that if we have a transfer matrix that has no particular normalization, then it has some eigenvalues, which could be smaller or bigger than one. But, in the case that eigenvalues are bigger than one, the corresponding eigenfunctions will explode exponentially and on the other hand, for smaller than one eigenvalues the eigenfunctions will go to zero. In quantum mechanics we have unitary matrices for the evolution of the systems, with all their eigenvalues having absolute value equal to one.

**Definition 1.** The step evolution operator  $\hat{S}$  is the transfer matrix T in the particular normalization that "the largest" eigenvalue of  $\hat{S}$  obeys:

$$|\lambda| = 1 \tag{4.86}$$

The normalization can be fixed by an additive constant in  $\mathscr{L}(m)$ :

$$\mathscr{L}'(m) = \mathscr{L}(m) + c \tag{4.87}$$

$$\mathscr{K}'(m) = e^{-c} \ \mathscr{K}(m) \tag{4.88}$$

So we get multiplicative normalization of the transfer matrix T and of the partition function Z. If we choose the normalization as above, the evolution law is expressed in terms of the step evolution operator:

$$\tilde{q}(m+1) = \hat{S}(m) \ \tilde{q}(m) \tag{4.89}$$

$$\bar{q}(m+1) = (\hat{S}^{\mathsf{T}})^{-1}(m) \ \bar{q}(m)$$
(4.90)

The case of the wave function components  $\tilde{q}(m+1)$  is trivial, it is obtained from equation (4.64) by replacing the transfer matrix with the step evolution operator. In case of the conjugated wave function  $\bar{q}(m+1)$  we need to assume that the step evolution operator is invertible. Then:

$$\bar{q}^T(m-1) = \bar{q}^T(m) \ \hat{S}(m-1)$$
(4.91)

$$\Rightarrow \bar{q}^T(m) = \bar{q}^T(m+1) \ \hat{S}(m) \tag{4.92}$$

$$\Rightarrow \bar{q}(m+1) = (\hat{S}^{\mathsf{T}})^{-1}(m) \ \bar{q}(m), \text{ if } \exists \ \hat{S}^{-1}$$
(4.93)

It is important to note that we explicitly demanded the step evolution operator to be invertible. After all, this means that we can look at the evolution of the system in both directions. In quantum mechanics, the time evolution is invertible because it is unitary:

$$\hat{U}^{\dagger} \cdot \hat{U} = I \tag{4.94}$$

As our motivation is showing that quantum mechanics emerge from classical statistics, we may restrict ourselves to systems with an invertible time evolution. Yet, our time evolution is not unitary. But, whenever the step evolution operator is an orthogonal matrix, which corresponds to a unitary matrix in the complex formulation, then:

$$(\hat{S}^T)^{-1}(m) = \hat{S}(m) \tag{4.95}$$

This means that  $\tilde{q}(m)$  and  $\bar{q}(m)$  have exactly the same evolution, and that is the essence of the unitary evolution in quantum mechanics. For general classical statistical systems the step evolution operator will not be an orthogonal matrix. As we will see below explicitly in the Ising model, there the step evolution operator is not an orthogonal matrix. This has to do with the fact that one looses information of boundary conditions if one goes to the bulk, in the bulk one often has only the equilibrium state that remains.

**Example 11** (Step evolution operator for the Ising chain). *Recall the transfer matrix of the Ising chain, and its eigenvalues:* 

$$T = \begin{pmatrix} 1 & e^{2\beta} \\ e^{-2\beta} & 1 \end{pmatrix}, \quad \lambda_{\pm} = 1 \pm 2e^{-2\beta}$$
(4.96)

$$tr\{T\} = 2, \quad det(T) = 1 - e^{-4\beta}$$
(4.97)

We notice that the eigenvalues are different to 1 and in particular one eigenvalue is bigger than 1, while the other one is smaller than 1. So, in order to get the step evolution operator, we normalize to the eigenvalue bigger than 1:

$$\hat{S} = \frac{1}{\lambda_{+}}T = \begin{pmatrix} \frac{1}{1+2e^{-2\beta}} & \frac{e^{2\beta}}{1+2e^{-2\beta}} \\ \frac{e^{-2\beta}}{1+2e^{-2\beta}} & \frac{1}{1+2e^{-2\beta}} \end{pmatrix} = \frac{1}{2\cosh(\beta)} \begin{pmatrix} e^{\beta} & e^{-\beta} \\ e^{-\beta} & e^{\beta} \end{pmatrix}, \quad where \quad \frac{1}{\cosh(\beta)} = \frac{1}{e^{\beta} + e^{-\beta}}$$
(4.98)

with the eigenvalues of  $\hat{S}$ :

$$\lambda_{+} = 1, \quad \lambda_{-} = tanh(\beta) \tag{4.99}$$

Notice that in the limit of zero temperature  $\beta \to \infty$   $(T \to 0)$  the two eigenvalues coincide.

$$\lim_{\beta \to \infty} \lambda_{-} = 1 \tag{4.100}$$

Furthermore, in this limit the step evolution operator becomes trivial:

$$\lim_{\beta \to \infty} \hat{S} = \frac{1}{2} \begin{pmatrix} 1 & 0\\ 0 & 1 \end{pmatrix}$$
(4.101)

Once more we observe the critical point at T = 0. The step evolution being trivial means nothing else than all Ising spins being aligned.

# 4.5 Operators of observables

Assume the local observable A(m) = A[s(m)]. Our aim is to compute the expectation value of the operator A, obtaining the expectation value law. We start from the basic law of CS:

$$\langle A(m) \rangle = \sum_{s(m)} A[s(m)] p[s(m)]$$

$$= \sum_{\tau} A_{\tau}(m) p_{\tau}(m)$$

$$= \sum_{\tau} A_{\tau}(m) \tilde{q}_{\tau}(m) \bar{q}_{\tau}(m)$$

$$(4.102)$$

If we associate A with a diagonal operator  $\hat{A}$ :

$$(\hat{A})_{\tau\rho} = A_{\tau}(m)\delta_{\tau\rho} \tag{4.103}$$

then law for expectation values reads:

$$\langle A(m) \rangle = \bar{q}_{\tau}(m) \ \hat{A}_{\tau\rho}(m) \ \tilde{q}_{\rho}(m)$$
(4.104)

Not only can the law for expectation values be written in terms of the classical wave functions, but it is also coincides with the one from quantum mechanics, except for the wave functions.

The proof is simple, insert the assumption of  $\hat{A}$  being a diagonal operator:

$$\bar{q}_{\tau}(m) \ \hat{A}_{\tau\rho}(m) \ \tilde{q}_{\rho}(m) = \sum_{\tau\rho} \bar{q}_{\tau}(m) \ A_{\tau}(m) \ \delta_{\tau\rho} \ \tilde{q}_{\rho}(m)$$
 (4.105)

$$=\sum_{\tau} \bar{q}_{\tau}(m) A_{\tau}(m) \tilde{q}_{\tau}(m)$$
(4.106)

$$=\sum_{\tau} A_{\tau}(m) p_{\tau}(m) \tag{4.107}$$

One important point, which is probably hidden in the above expression, is the fact that the operator  $\hat{A}$  is assumed diagonal in the occupation number basis. Thus, our probabilistic information is encoded in a "diagonal way". If we go to another basis,  $\hat{A}$  will not remain diagonal, but the law for expectation values remains the same. And this important for non-commuting observables, like in quantum mechanics. If we assume two non-commuting operators (corresponding to observables) and a basis such that one operator is diagonal, then the expectation values of both observables can be calculated. But for one of the observables the information is encoded in off-diagonal elements. Soon we will see that non-commuting operators are also part of CS.

# 4.6 Solution of the initial value problem

Now we have collected all concepts to solve the initial value problem of the Ising chain. Remember, we could solved the initial value problem of the Ising chain earlier, by "brute force" (like done in example 8 without boundary conditions). Instead we have seen that we construct a classical wave function formalism with similar traits to the one in quantum mechanics, the main differences being two classical wave functions and a more general time evolution.

#### 1. Initial value problem:

We would like to compute the expectation value  $\langle A(m) \rangle$  of a local observable A(m) (e.g A(m) = s(m)) given arbitrary boundary conditions  $f_i(0)$  and  $\tilde{f}_f(M)$ .

There are 4 conceptual steps to do so:

(a) Input: Impose boundary conditions:

$$\tilde{q}(t_i) = \tilde{q}(0)$$
 ,  $\bar{q}(t_f) = \bar{q}(M)$ 

(b) Evolution: Use the step evolution operator  $\hat{S}$  *m*-times to produce q(m) from q(0):

$$\tilde{q}(m) = \hat{S}(m-1)\tilde{q}(m-1) = \hat{S}(m-1)\left(\hat{S}(m-2)\tilde{q}(m-2)\right) = \dots = \left(\underbrace{\hat{S}(m-1)\hat{S}(m-2)\dots\hat{S}(0)}_{m \text{ factors}}\right)\tilde{q}(0)$$

Since in the Ising model,  $\hat{S}$  is independent of the spin position m we get:

$$\tilde{q}(m) = \hat{S}^m \; \tilde{q}(0)$$

In a similar way

$$\bar{q}(m) = (\hat{S}^T)^{M-m} \bar{q}(M)$$

(c) <u>Normalization</u>: In order to properly normalize our wave functions, we need the partition function Z:

$$Z = \bar{q}^T(m)\tilde{q}(m) =$$

$$= \left( (\hat{S}^T)^{M-m}\bar{q}(M) \right)^T \hat{S}^m \tilde{q}(0) =$$

$$= \bar{q}(M)^T \left( (\hat{S}^T)^{M-m} \right)^T \hat{S}^m \tilde{q}(0) =$$

$$= \bar{q}(M)^T \left( \hat{S}^{M-m} \hat{S}^m \right) \tilde{q}(0) =$$

$$= \bar{q}(M)^T \left( \hat{S}^M \right) \tilde{q}(0)$$

Normalize  $\tilde{q}(0)$ ,  $\bar{q}(M)$  such that Z = 1. For example:

$$\tilde{q}(0) \rightarrow \frac{1}{\sqrt{Z}} \tilde{q}(0) \quad \& \quad \bar{q}(M) \rightarrow \frac{1}{\sqrt{Z}} \bar{q}(M)$$

(d) Expectation value: Given the normalization Z=1, the law for expectation values provides:

$$\langle A(m) \rangle = \bar{q}^T(m)\hat{A}(m)\tilde{q}(m)$$

#### 2. Equilibrium eigenfunctions:

By definition the, the largest eigenvalue of the matrix  $\hat{S}$  obeys  $^{2} \lambda = 1$ . If the largest eigenvalue is unique (multiplicity 1), then we define the equilibrium eigenfunctions as follows.

<sup>&</sup>lt;sup>2</sup>We can always normalize  $\hat{S}$  in order to have  $\lambda = 1$ :  $\hat{Q}\vec{v} = \lambda\vec{v} \Rightarrow \frac{1}{\lambda}\hat{Q} = \frac{\lambda}{\lambda}\vec{v} = \vec{v}$ 

**Definition 2.** The equilibrium wave function  $\tilde{q}_{eq}(\bar{q}_{eq})$  is the eigenvector of the step evolution operator  $\hat{S}(\hat{S}^{\intercal})$  that corresponds to the unique eigenvalue  $\lambda = 1^3$ :

$$\hat{S}\tilde{q}_{eq} = \tilde{q}_{eq} \quad , \quad \hat{S}^T \bar{q}_{eq} = \bar{q}_{eq} \iff \bar{q}_{eq}^T \hat{S} = \bar{q}_{eq}^T$$

The term equilibrium wave function will become clear soon, we will realize that away from the bulks the equilibrium wave functions will be dominant, due to the information loss by going from the fixed boundaries into the bulk.

For the Ising chain at finite  $\beta$ , the largest eigenvalue of  $\hat{S}$  has multiplicity 1. We can explicitly write:

$$\hat{S} = \frac{1}{2 \cosh(\beta)} \begin{pmatrix} e^{\beta} & e^{-\beta} \\ e^{-\beta} & e^{\beta} \end{pmatrix}$$
$$\lambda_{+} = 1, \quad \lambda_{-} = \tanh(\beta)$$

where the corresponding eigenvectors are

$$\tilde{q}_{\rm eq} \equiv \tilde{q}_+ = \begin{pmatrix} (\tilde{q}_{eq})_+ \\ (\tilde{q}_{eq})_- \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad , \quad \tilde{q}_- = \begin{pmatrix} (\tilde{q}_-)_+ \\ (\tilde{q}_-)_- \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

In this case  $\hat{S}$  is symmetric, therefore:

$$\tilde{q}_{\rm eq} = \bar{q}_{\rm eq} \quad , \quad \tilde{q}_- = \bar{q}_-$$

Observe that in equilibrium:

$$p_{+}(m) = (\bar{q}_{eq})_{+} \cdot (\tilde{q}_{eq})_{+} = \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} = \frac{1}{2}$$
$$p_{-}(m) = (\bar{q}_{eq})_{-} \cdot (\tilde{q}_{eq})_{-} = \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} = \frac{1}{2}$$

The probability for Ising spin up and down are equal, like expected in the equilibrium state.

3. Boundary conditions: It is time to include our boundary conditions. The eigenvectors  $\{\tilde{q}_{eq}, \tilde{q}_{-}\}$  correspond to different eigenvalues  $\lambda_{+} \neq \lambda_{-}$ , so that they are orthogonal. As a consequence the set  $\{\tilde{q}_{eq}, \tilde{q}_{-}\}$  is a basis for  $\mathbb{R}^{2}$ . With that in mind we can expand the boundary conditions in this basis:

$$\tilde{q}(0) = c(\tilde{q}_{eq} + a \; \tilde{q}_{-}) \quad , \quad \bar{q}(\mathscr{M}) = \bar{c}(\bar{q}_{eq} + \bar{a} \; \bar{q}_{-})$$

where  $c, \bar{c}, a, \bar{a} \in \mathbb{R}$ .

4. Correlation length (time) A quantity we have already met in example 8 is the correlation length. In example 8 the correlation length entered in the spin-spin-correlation  $\langle S_i S_j \rangle$  of two spins along the Ising chain. For T > 0 the correlation length was finite and the spin-spincorrelation decreased with  $\alpha^{\Delta m}$ ,  $\alpha < 1$ . The larger the distance between to Ising spins, the

<sup>&</sup>lt;sup>3</sup>In general,  $\hat{S}, \hat{S}^T$  have the same eigenvalues but different eigenvectors

smaller their correlation, which corresponds to an information loss along the Ising chain. Now, we will see that the correlation length enters in the solution of the initial value problem as well. And here it will describe the loss of information of the boundary conditions when going into the bulk.

We have seen that every local probability on the Ising chain can be expressed by the boundary conditions  $\tilde{q}(0)$ ,  $\bar{q}(\mathcal{M})$  using the operator  $\hat{S}$ :

$$\begin{split} \tilde{q}(m) &= \hat{S}^m \; \tilde{q}(0) \\ &= \hat{S}^m c(\tilde{q}_{\rm eq} + a \tilde{q}_-) \\ &= c(\hat{S}^m \tilde{q}_{\rm eq} + a \hat{S}^m \tilde{q}_-) \\ &= c(\lambda_+^m \tilde{q}_{\rm eq} + a \lambda_-^m \tilde{q}_-) \end{split}$$

Using that  $\lambda_{+} = 1$ , the final result is:

$$\tilde{q}(m) = c(\tilde{q}_{\rm eq} + a\lambda_{-}^m \tilde{q}_{-})$$

Similarly:

$$\bar{q}(m) = \bar{c}(\bar{q}_{\rm eq} + \bar{a}\lambda_-^{M-m}\bar{q}_-)$$

These are the classical wave functions for the Ising chain including boundary conditions. Note that the difference from the equilibrium wave function decreases if we are far away from the boundaries, as  $\lambda_{-}^{m} < 1$ . If we move from the boundary into the bulk we get closer to the equilibrium state  $\tilde{q}_{m} \rightarrow \tilde{q}_{eq}$ . This motivates defining the correlation length in order to see how fast this transition into the equilibrium state happens. If the transition occurs rapidly,  $\lambda_{-} \ll 0$ , then the state of an Ising spin going into the bulk will be dominated by the equilibrium wave function soon. In contrast the information about the boundary conditions gets lost. On the other hand, if the transition happens slowly,  $\lambda_{-} \approx 1$ , then the information about the boundary conditions will influence the Ising spins far within the bulk. We conclude that the quantity that measures how fast the transition to equilibrium occurs, is  $\lambda_{-}$ . And  $\lambda_{-}$  is linked to the correlation length  $\xi$ .

Let  $\epsilon$  denote the distance of two lattice points. Define the parameter t by  $t = \epsilon m$ , it is the difference from the initial Ising spin s(0). Further denote  $t_f = \epsilon M$ .

**Definition 3.** The correlation length  $\xi$  is defined via the following equations:

$$\lambda_{-}^{m} = exp\left(-\frac{t}{\xi}\right)$$
$$\lambda_{-}^{M-m} = exp\left(-\frac{t_{f}-t}{\xi}\right)$$

For the Ising chain, with  $\lambda_{-} = tanh(\beta)$ , we obtain:

$$\xi_{\text{Ising}} = \frac{\epsilon}{\ln\left(\frac{1}{\lambda_{-}}\right)} = \frac{\epsilon}{\ln\left(\frac{1}{\tanh\beta}\right)}$$

Using that  $tanh(\beta) = 1/coth(\beta)$ :

$$\xi_{Ising} = \frac{\epsilon}{\ln(\coth\beta)}$$

Now we can express  $\tilde{q}(m)$ ,  $\bar{q}(m)$  in terms of the correlation length  $\xi$ :

$$\tilde{q}(t) = c \left( \tilde{q}_{eq} + a \exp\left(-\frac{t}{\xi}\right) \tilde{q}_{-} \right), \ \bar{q}(t) = \bar{c} \left( \bar{q}_{eq} + \bar{a} \exp\left(-\frac{t_f - t}{\xi}\right) \bar{q}_{-} \right)$$

We observe that for  $t \gg \xi$  and  $t_f - t \gg \xi$  we have an effective loss of memory. The wave functions converge into the equilibrium wave function, the correlation length determines how fast the convergence is.

In the limit  $\beta \to \infty \Rightarrow e^{-\beta} \to 0$  multiple Taylor series can be applied:

$$\tanh \beta = \frac{e^{\beta} - e^{-\beta}}{e^{\beta} + e^{-\beta}} = \frac{1 - e^{-2\beta}}{1 + 2e^{-\beta}} = (1 - e^{-2\beta})(1 - e^{-2\beta} + \dots) = 1 - 2e^{-2\beta} + \dots$$
$$\coth \beta = \frac{1}{\tanh \beta} = \frac{1}{1 - 2e^{-2\beta} + \dots} = 1 + 2e^{-2\beta} + \dots$$
$$\ln(\coth \beta) = \ln(1 + 2e^{-2\beta} + \dots) = 2e^{-2\beta} + \dots$$

In the limit  $\beta \to \infty \leftrightarrow T \to 0$  the correlation length diverges:

$$\xi_{\text{Ising}} = \frac{\epsilon}{2}e^{2\beta} + \dots \to \infty$$

In this limit there is no loss of boundary information and all Ising spins have to be aligned to the initial Ising spin. Once more we observe the critical point at T = 0.

#### 5. Local probabilities:

No that we have solved the initial value problem of the Ising chain, we can investigate the local probabilities for an Ising spin to be up or down. The probabilities simply are:

$$p_{+}(t) = (\bar{q}_{eq})_{+}(t) \cdot (\tilde{q}_{eq})_{+}(t) = \frac{c\bar{c}}{2} \left(1 + a \ e^{-\frac{t}{\xi}}\right) \left(1 + \bar{a} \ e^{-\frac{t_{f}-t}{\xi}}\right)$$
$$p_{-}(t) = (\bar{q}_{eq})_{-}(t) \cdot (\tilde{q}_{eq})_{-}(t) = \frac{c\bar{c}}{2} \left(1 - a \ e^{-\frac{t}{\xi}}\right) \left(1 - \bar{a} \ e^{-\frac{t_{f}-t}{\xi}}\right)$$

The normalization demands  $p_+ + p_- = 1$ , which is equivalent to:

$$c\bar{c} = \frac{1}{1+a\ \bar{a}\ e^{-\frac{t_f}{\xi}}}$$

Now assume the Ising spin orientation as local observable, then the expectation value is:

$$\begin{split} \langle S(t) \rangle &= \sum_{s(t)=\pm 1} s(t) p_{s(t)} = p_{+} - p_{-} = c\bar{c} \left( a \ e^{-\frac{t}{\xi}} + \bar{a} e^{-\frac{t_{f}-t}{\xi}} \right) = \frac{a \ e^{-\frac{t}{\xi}} + \bar{a} e^{-\frac{t_{f}-t}{\xi}}}{1 + a \ \bar{a} \ e^{-\frac{t_{f}}{\xi}}} \\ \langle S(0) \rangle &= \frac{a + \bar{a} e^{-\frac{t_{f}}{\xi}}}{1 + a \bar{a} e^{-\frac{t_{f}}{\xi}}} \end{split}$$

$$\langle S(t_f) \rangle = \frac{a \ e^{-\frac{t_f}{\xi}} + \bar{a}}{1 + a\bar{a}e^{-\frac{t_f}{\xi}}}$$

We could further solve for  $a, \bar{a} \dots$  in order to obtain an expression in terms of  $\langle S(0) \rangle$ ,  $\langle S(t_f) \rangle$ . This is left as an exercise for the reader.

With that we close the discussion about the initial value problem of the Ising chain.

# 4.7 Classical density matrix

Having constructed the classical wave functions  $\tilde{q}, \bar{q}$  (in the occupation number basis) it is only natural to ask, whether we can construct a classical density matrix. The answer is yes, and the construction is analogue to the one in quantum mechanics. There we have seen that given a wave function  $\psi$  we can construct the pure state density matrix:

$$\rho = \psi \psi^{\dagger} \tag{4.108}$$

Motivated by this, we define the classical density matrix  $\rho'$  to be:

$$\rho_{\tau\rho}'(t) = \tilde{q}_{\tau}(t)\bar{q}_{\rho}(t)$$
(4.109)

which is a bilinear in the classical wave functions. The diagonal elements of this matrix are our local probabilities:

$$p_{\tau}(t) = \rho_{\tau\tau}'(t) \tag{4.110}$$

Using the density matrix we can calculate the expectation value of an observable A in the following way:

$$\langle A(t)\rangle = tr\{\rho'(t)\hat{A}\}$$
(4.111)

For the proof just use the wave function expression of the density matrix:

 $\langle A$ 

$$\begin{aligned} \langle t \rangle &= \bar{q}_{\tau}(t) \ \hat{A}_{\tau\rho} \ \tilde{q}_{\rho}(t) \\ &= \tilde{q}_{\rho}(t) \ \bar{q}_{\tau}(t) \ \hat{A}_{\tau\rho} \\ &= \rho'_{\rho\tau}(t) \ \hat{A}_{\tau\rho} = tr(\rho'(t)\hat{A}) \end{aligned}$$
(4.112)

More importantly, the evolution law for the density matrix is given by:

$$\rho'(t+\epsilon) = \hat{S}(t)\rho'(t)\hat{S}^{-1}(t)$$
(4.113)

To prove that we need to remember the evolution law for the classical wave functions:

$$\rho'(t+\epsilon) = \tilde{q}(t+\epsilon)\bar{q}^{T}(t+\epsilon)$$

$$= \left(\hat{S}(t)\tilde{q}(t)\right) \left((\hat{S}^{-1})^{T}(t)\bar{q}(t)\right)^{T}$$

$$= \left(\hat{S}(t)\tilde{q}(t)\right) \left(\bar{q}(t)^{T}\hat{S}^{-1}(t)\right)$$

$$= \hat{S}(t) \left(\tilde{q}(t)\bar{q}(t)^{T}\right) \hat{S}^{-1}(t)$$

$$= \hat{S}(t)\rho'(t)\hat{S}^{-1}(t) \qquad (4.114)$$

We can thus write the time evolution of the diagonal elements, which correspond to the local probabilities:

$$\rho_{\tau\tau}'(t+\epsilon) = \hat{S}_{\tau\rho}(t) \; \rho_{\rho\sigma}'(t) \; \hat{S}_{\sigma\tau}^{-1}(t)$$
(4.115)

From this expression we can clearly see, that the evolution of local probabilities in general also involves the off-diagonal terms of  $\rho'$  in addition to the local probabilities, which are the diagonal elements. This has the consequence, we cannot formulate an evolution law for local probabilities if there are non-zero off-diagonal elements in our density matrix. An example is the Ising chain. In the next chapter we will see a class of systems, the so called probabilistic automata, which yield an evolution law for probabilities. In general classical statistical system do not yield an evolution law for probabilities, only for the classical wave functions and the density matrix.

#### Evolution needs density matrix or wave functions!

Equation 4.113 has the same linear structure as the von Neumann equation in quantum mechanics. Thus the superposition principle holds also for the classical system, meaning that a linear combination of solutions is again a solution of equation 4.113.

This setting generalizes to all local chains with  $\mathscr{K}(m)$  involving only variables at m and m+1. It can also be extended, e.g. for  $\mathscr{K}(m)$  involving variables at m, m+1, m+2 (not only next neighbor interaction).

We have reached a key part of the lecture. There are classical wave functions from which a classical density matrix can be constructed. Further, we obtained the law for (time) evolution and for expectation values. Quantum mechanics are included in this formalism if the conjugated wave function is the complex conjugated of the wave function and the time evolution is unitary (or orthogonal in a real setting). We summarize these insights in the following two statements

The general evolution law of classical statistics is based on the density matrix.

Quantum type laws follow from the concept of evolution in classical statistics.

As a sidenote for those familiar with functional integrals in quantum field theory: The operator formalism follows from the functional integral. The functional integral corresponds to the overall probability distribution. Recall that the configuration sum can be viewed as a discrete version of the path integral (4.1).

### 4.8 Non-commuting operators

In quantum mechanics non-commuting operators are typical quantities. Such pairs or sets of operators describe the inability to measure two observables with non-commuting corresponding operators exactly simultaneously. Well known examples are the position and momentum of a particle, or its spin in different directions. The non-commutativity of operators is also encoded in our probabilistic description. The local probabilistic information in the off-diagonal elements of the density matrix can be used for the computation of expectation values of non-commuting operators (see 1.4).

In classical statistics non-commuting operators are at first not expected intuitively. After all, classical statistics are used to describe classical physical systems. And in classical physics observables can be measured exactly simultaneously. This first intuitive expectation is not true, because non-commuting operators are an inherent property of statistics.

Consider a classical density matrix  $\varrho'(t)$  and a local observable  $\hat{A}(t)$ . One may propose the question: If we know the classical density matrix of a system at local point t and the corresponding local observable, can we compute the expectation value  $\langle A(t+\epsilon) \rangle$ ?

The answer is indeed, yes. For that define an operator:

$$\hat{A}(t+\epsilon): \langle A(t+\epsilon) \rangle = tr\{\varrho'(t)\hat{A}(t+\epsilon;t)\}$$
(4.116)

where:

$$\hat{A}(t;t) = \hat{A}(t)$$
 (4.117)

The operator searched for is:

$$\hat{A}(t+\epsilon;t) = \hat{S}^{-1}(t)\hat{A}(t+\epsilon,t+\epsilon)\hat{S}(t)$$
(4.118)

The proof is rather simple, use the time evolution of the density matrix to obtain:

$$\langle A(t+\epsilon) \rangle = tr\{\hat{A}(t+\epsilon;t+\epsilon)\varrho'(t+\epsilon)\}$$
  
=  $tr\{\hat{A}(t+\epsilon;t+\epsilon)\hat{S}(t)\varrho'(t)\hat{S}^{-1}(t)\}$   
=  $tr\{\hat{S}^{-1}(t)\hat{A}(t+\epsilon;t+\epsilon)\hat{S}(t)\varrho'(t)\}$  (4.119)

In general one will then observe that:

$$[\hat{A}(t+\epsilon;t),\hat{A}(t)] \neq 0, \ \epsilon > 0 \tag{4.120}$$

so that we have found a pair of non-commuting operators in classical statistics.

**Example 12** (Occupation number). Assume the Ising chain with the local observable being the occupation number operator:

$$\hat{A}(t+\epsilon;t+\epsilon) = \hat{N} = \begin{pmatrix} 1 & 0\\ 0 & 0 \end{pmatrix}$$
(4.121)

which is time independent. The step evolution operator of the Ising chain has been previously calculated:

$$\hat{S}(t) = \frac{1}{2\cosh(\beta)} \begin{pmatrix} exp(\beta) & exp(-\beta) \\ exp(-\beta) & exp(\beta) \end{pmatrix}$$
(4.122)

which is constant as well. And so:

$$\hat{N}(t+\epsilon;t) = \hat{S}^{-1}\hat{N}\hat{S} \tag{4.123}$$

$$=\frac{1}{1-4exp(-4\beta)}\begin{pmatrix}1 & exp(-2\beta)\\ exp(-2\beta) & -exp(-4\beta)\end{pmatrix}$$
(4.124)

This operator does not commute with the occupation number operator:

$$[\hat{N}(t+\epsilon;t),\hat{N}] = \frac{-1}{1 - exp(-4\beta)} \begin{pmatrix} 0 & exp(-2\beta) \\ exp(-2\beta) & 0 \end{pmatrix} \neq 0, \ T > 0$$
(4.125)

Only in the case of T=0, so at the critical point, both operator commute. At the critical point all spins are aligned, and so the occupation number operators commute.

# 4.9 Quantum formulation for classical statistics

To close this chapter, let us talk about the generalization of what we have observed so far. The weight distribution of an arbitrary local chain can be written as:

$$w = \tilde{f}_f \prod_m \mathscr{K}(m) \tilde{f}_i \tag{4.126}$$

where  $\mathscr{K}(m)$  denotes an arbitrary local factor, especially there is no restriction to next neighbor interactions like before. Further, the concepts of occupation numbers can be generalized:

$$n(m) \to n_{\gamma}(m) \tag{4.127}$$

where  $\gamma \in \mathbb{N}$  denotes multiple occupation sites at the point m. Then the occupation number basis needs to be generalized. Assume the case  $\gamma = 1, 2$ , define:

$$h_1 = h_{11} = n_1 n_2, \ h_2 = h_{10} = n_1 (1 - n_2), \ h_3 = h_{01} = (1 - n_1) n_2, \ h_4 = h_{00} = (1 - n_1) (1 - n_2)$$
(4.128)

But still one can construct the basis functions such that:

$$h_{\tau} h_{\rho} = h_{\rho} \,\delta_{\tau\rho}$$
: orthonormality (4.129)

$$\sum h_{\tau} = 1 \tag{4.130}$$

$$\sum_{\{n\}} h_{\tau} = 1 \tag{4.131}$$

$$\sum_{\{n\}} h_{\tau} h_{\rho} = \delta_{\tau\rho} : \text{ completeness}$$
(4.132)

We conclude that the quantum formalism applies to all local chains. The evolution law for classical the density matrix reads:

$$\underline{\varrho'(t+\epsilon)} = \hat{S}\varrho'(t)\hat{S}^{-1}(t) \tag{4.133}$$

And the evolution law for local observables is:

$$tr\{\varrho'(t)\hat{A(t)}\}$$
(4.134)

These laws are similar to the von Neumann equation for the quantum mechanical density matrix and the quantum law for expectation values. The difference lies in the in general non-unitary time evolution in classical statistics. It remains to find classical subsystems with unitary (orthogonal in the real case) evolution laws. In the next chapter we will investigate an example of classical systems with orthogonal time evolution, the so called probabilistic automata.

# Chapter 5 Probabilistic cellular automata

We have already found many features of the quantum formalism in classical statistics: wave functions as a type of probability amplitudes, the density matrix, operator of observables and the quantum law for expectation values. What is missing is the unitary evaluation, or orthogonal evaluation for real quantum mechanics. This will also be limited to the issue of the relation between the <u>two classical wave functions</u>. Not every classical statistical ??? a unitary evolution. We have found for the ??? claim that the evolution approaches an equilibrium state and initial information is ??? lost. In this lecture we deal with simple classical statistical systems without information loss, namely probabilistic cellular automata.

# 5.1 Deterministic cellular automata

- At t the automaton is characterized by a specific configuration of bits  $\bar{\rho}(t)$
- The updating rule specifies for each  $\bar{\rho}(t)$  the configuration  $\bar{\tau}(t+\epsilon)$  at the next (time-) step.  $\bar{\rho} \to \overline{\bar{\tau}(\bar{\rho})}$
- <u>cellular</u> automata at given t: bits are grouped into cells, labeled by  $x n_{\alpha}(x)$ , definition of neighboring cells  $y_i(x)$  for each cell x, configuration of bits in cell x after updating at  $t + \epsilon$  only depends on bit-configuration for  $y_i(x)$  in the neighborhood of x at t.
- Example: one dimensional CA neighbors of x:  $y_i = \{x \epsilon, x, x + \epsilon\}$



Figure 5.1: One dimensional CA

updating: 
$$\underbrace{\{n_{\alpha}(x-\epsilon), n_{\alpha}(x), n_{\alpha}(x+\epsilon)\}}_{t} \rightarrow \underbrace{\{n_{\alpha}(x)\}}_{t+\epsilon}$$

We will describe the cellular aspect later, for the moment we stay with general automata.

- <u>General automata</u>: Characterized by a map  $\bar{\rho}(t) \rightarrow \bar{\tau}(t+\epsilon), \ \bar{\tau} = \bar{\tau}(\bar{\rho})$
- <u>Reversible</u> automata: inverse map exists,  $\bar{\rho} = \bar{\rho}(\bar{\tau})$ map:  $\bar{\rho}(t) \rightarrow \bar{\tau}(t+\epsilon), \ \bar{\tau} = \bar{\tau}(\bar{\rho})$
- Characterize state of automaton at t by a simple wave function,

$$q(t) \equiv \tilde{q}(t) - \bar{q}(t) \tag{5.1}$$

with components

$$q_{\tau}(t) = \delta_{\tau,\bar{\tau}(t)}; \qquad \begin{pmatrix} 0\\0\\1\\0\\0 \end{pmatrix} \leftarrow \bar{\tau}(t)$$
(5.2)

The updating amounts to a rotation of the vector q(t) to  $q(t+\epsilon)$ 

• Step evolution operator:

$$q_{\tau}(t+\epsilon) = \hat{S}_{\tau\rho}(t)q_{\rho}(t) \tag{5.3}$$

$$\hat{S}_{\tau\rho} = \delta_{\tau,\bar{\tau}(\rho)} = \delta_{\bar{\rho}(\tau),\rho} \tag{5.4}$$

for reversible automata: updating is invertible,  $\bar{\rho}(t)$  exists,  $\hat{S}(t)$ :unique jump operator in each row and column precisely one element is equal to one, all other elements are zero Example:

$$\hat{S} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \hat{S} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \hat{S} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

 $\hat{S}$  is an orthogonal matrix, e.g.  $\hat{S}^T\hat{S}=1\to$  Orthogonal evolution automatic for automata

• Typical task: input configuration  $\bar{\tau}(t_{in}) \to q(t_{in})$  follow evolution to  $t_f = t_i n + m_t \epsilon$ , What is the final bit-configuration:  $q(t_f)$ ?

#### ??? 2 lines missing script 4 page 4 bottom

# 5.2 Probabilistic (cellular) automaton

- Probability distribution for initial configurations:  $\{p_{\tau}(t_{in})\}$ , Encoded in initial wave function  $q(t_{in})$ , with  $p_{\tau}(t) = q_{\tau}^2(t)$ , The wave function is still a unit vector,  $q_t au(t)q_t au(t) = 1$ , but no longer a vector with a single non-zero element (deterministic limit)
- Probabilities get transported  $p_{\tau}(t+\epsilon) = p_{\bar{\rho}(\tau)}(t), \quad \bar{\rho}(t)$ : bit configuration from which  $\tau$  results by updating rule
- Evolution of wave function is described by the same step evolution operator as for the deterministic automaton,

$$\boxed{q_{\tau}(t+\epsilon) = \hat{S}_{\tau\rho}(t)q_{\rho}(t)} \text{ check:}$$

$$q_{\tau}(t+\epsilon) = \hat{S}_{\tau\rho}(t)p_{\rho}(t)$$
(5.5)

$$=\delta_{\bar{\rho}(\tau),\rho}q_{\rho}(t), \qquad \bar{\rho}(\tau)=\bar{\rho}(\tau;t)$$
(5.6)

$$=q_{\bar{\rho}(\tau)}(t) \tag{5.7}$$

$$p_{\tau}(t+\epsilon) = q_{\tau}^2(t+\epsilon) = q_{\bar{q}(\tau)}^2(t) = p_{\bar{\rho}(\tau)}(t)$$
(5.8)

Remark: For CA an evolution law in terms of probabilities exists, wave function and density not necessary for evolution law

They are convenient for other purposes: correlations at different times, coarse graining

Classical density matrix  $\rho(t)$  $\rho_{\tau\rho} = q_{\tau}(t)q_{\rho}(t), \qquad \rho^{T} = \rho$ 

Expectation values of local observables

$$\langle A(t) \rangle = tr\{\rho(t)\hat{A}(t)\}$$
(5.9)

$$=q_{\tau}(t)\hat{A}_{\tau\rho}(t)q_{\rho}(t) \tag{5.10}$$

for local observables at t,  $\hat{S}_{\tau\rho} = a_{\tau}\delta_{\tau\rho}$   $\langle A(t) \rangle = A_{\tau}p_{\tau}(t)$ Generalized Ising model:  $\hat{S}(t)$  defines local factor  $\mathscr{K}(t) = \mathscr{K}(s(t+\epsilon), s(t) = h_{\tau}(s(t+\epsilon))\hat{S}_{\tau\rho}(t)h_{\rho}(s(t))$  q(0) defines boundary factor choose  $\bar{q}(t_{\rho}) = q(t_{\rho})$  as final boundary factor overall probability  $p\{s(t)\} = \bar{f}(t_{\rho})\mathscr{K}(t_{\rho} - \epsilon)...\mathscr{K}(0)f_{in}(0)$ recall  $\hat{S}_{\tau\rho} = exp(-\mathscr{L}_{\tau\rho})$ "Constrained Ising Model", " forbidden neighboring configurations", e.g. probability ???

Probabilistic automata are classical statistical systems. Quantum formalism can be applied. Particularly: orthogonal evolution,  $\hat{S}^T(t)\hat{S}(t) = 1$ ; single wave fct.  $\tilde{q} = \bar{q} = q$ ;  $\rho^T = \rho$ 

# 5.3 Clock automata

• So now order that  $\tau$  is:

$$\tau = 1 \dots N \tag{5.11}$$

We can of course use the bit representation of nature numbers, if we want to describe for example a system with four bits:

$$4 \ bits: \ \tau = 1 \dots 16$$

• Next, we consider the Step evolution operator which defines the particular automaton:

$$\hat{S}_{\tau\rho} = \delta_{\tau,\rho+1} , \quad periodic \ in \quad \rho + N = \rho$$

$$(5.12)$$

Taking the limit  $N \to \infty$  we obtain the continuous rotation, which obviously leads us to the definition of an angle  $\alpha_{\tau}$  as :

$$\alpha_{\tau} \equiv \frac{2\pi\tau}{N} \tag{5.13}$$

So, evolution equation can be written in terms of  $\alpha_{\tau}$  as:

$$q_{\tau}(t) = q(t, \alpha_{\tau}) \tag{5.14}$$

In the case that  $N \to \infty$  the different discrete steps of  $\alpha_{\tau}$  will vanish and the angle becomes a continuous function, but we need to point out that at the moment we are in the regime of discret time. The next step is to consider the limit of zero time differences:

$$q(t + \epsilon, \alpha_{\tau}) = q\left(t, \alpha_{\tau} - \frac{2\pi}{N}\right)$$
$$= q\left(t, \alpha - \Delta\alpha\right)$$
(5.15)

We are almost at the point where our discreet equation becomes a continuous equation:

$$q(t+\epsilon,\alpha) - q(t,\alpha) = q(t,\alpha - \Delta\alpha) - q(t,\alpha)$$
$$\Rightarrow \frac{1}{\epsilon} \Big[ q(t+\epsilon,\alpha) - q(t,\alpha) \Big] = \frac{\Delta\alpha}{\epsilon} \frac{1}{\Delta\alpha} \Big[ q(t,\alpha - \Delta\alpha) - q(t,\alpha) \Big]$$

Continuum limits:

$$N \to \infty, \ \epsilon \to 0, \ \Delta \alpha \to 0$$

Also the particular limit of the ratio:

$$\frac{\Delta\alpha}{\epsilon} = \omega = const. \tag{5.16}$$

For smooth wave function:

$$\partial_t q(t,\alpha) = \omega \partial_\alpha q(t,\alpha)$$
(5.17)

Notice that multiplying both sides with i, in the left side the operator takes the form:

$$-i\omega \frac{\partial}{\partial \alpha} \sim angular \ momentum$$
 (5.18)

Summarizing, we showed that in principle the discrete time evolution turns in a suitable continuum limit to a continuous time evolution ,which is given by a differential equation and that the information is kept and the differential equation is linear.

• Periodic evolution:

$$q(t,\alpha) = q(0,\alpha_{-}\omega t) \tag{5.19}$$

With a particular point at t = 0:

 $\alpha_0$ : position at t = 0

So at a generic t the point of position is:  $\alpha = \alpha_0 + \omega t$ 

This simple clock automaton in a certain sense introduce the notion of time, the notion of the periodicity and it produces the Schrödinger equation!

• Finite automata are clock systems:

$$\hat{S} = unique \ jump \ matrix, \ N \times N, \ N \ arbitrary \ large \ but \ finite$$

Which will give always a periodic system but with different periods.

# Chapter 6

# Generalised Ising model for fermions

In this section, we disucus probabilistic cellular automata that describe quantum field theories in 1+1 dimensions. These classical statistical systems are fully equivalent to quantum systems. First, we will approach free theories without interaction and subsequently discuss interactions as well.

# 6.1 Diagonal Ising model

We consider the Ising spins  $s(m_1, m_2)$  on a two dimensional lattice with  $M^2$  sites  $0 \le m_1 \le M - 1$ ,  $0 \le m_2 \le M - 1$ .

In a diagonal Ising model, the spin at a specific site only interacts with one of its nearest diagonal neighbours:

$$S = -\beta \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} \{ s(m_1+1, m_2+1)s(m_1m_2) - 1 \}$$
(6.1)

For  $\beta > 0$ , the system tends to align along the diagonal.

We can now select a time hypersurfaces label by  $m_2 = \text{const.}$  and introduce time and position scales:

$$t = \epsilon m_1, \quad x = \epsilon m_2 \tag{6.2}$$

Therefore, the spins can also be written as

$$s(m_1, m_2) = s(t, x) = s_{\gamma}(t)$$
 (6.3)

where  $\gamma$  corresponds to a value of x, or  $m_2$ , likewise. It should be noted that this selection is arbitrary due to a symmetry under exchange of  $m_1$  and  $m_2$ . We now can treat the system as a local chain:

#### • Boundary terms

For local chains, the overall probability distribution is given as

$$p[s] = B[s]\exp(-S[s]) \tag{6.4}$$

The boundary term B[s] can be divided in a boundary in x- direction and in t-direction:

$$B[s] = B_x[s]B_t[s] \tag{6.5}$$

For simplicity, we assume periodicity in x-direction  $s(m_1, M) = s(m_1, 0)$ , which yields

$$B_x[s] = \exp\{-\beta \sum_{m_1} [s(m_1+1,0)s(m_1,M-1)-1]\}$$
(6.6)

For the boundary in t-direction, we again use

$$B_t[s] = f_{in}[s(0, m_2)] \bar{f}_f[s(M-1, m_2)]$$
(6.7)

#### • Local factors

The local factors of the system are given by

$$\mathscr{K}(m_1) = \exp(-\mathscr{L}(m_1)) \tag{6.8}$$

with

$$\mathscr{L}(m_1) = -\beta \sum_{m_2} [s(m_1 + 1, m_2 + 1)s(m_1, m_2) - 1]$$
(6.9)

We can therefore identify it with a generalized Ising model.

• As we have seen, the **Quantum formalism** applies for Ising models.  $f_{in}[s(0, m_2)]$  defines the components  $\tilde{q_{\tau}}(0)$  of the wave function in occupation number basis, likewise  $\bar{f_f}[s(M-1, m_2)]$  relates to  $\bar{q_{\tau}}(M-1)$ . As opposed to the Ising chain,  $\tau$  now refers to the entire bit configuration at a fixed t, or  $m_1$ .

#### • If $\beta \to \infty$ , this system becomes a 'unique jump chain'. Since

$$\mathscr{L} = \begin{cases} 0, & s(m_1 + 1, m_2 + 1) = s(m_1, m_2) \text{ aligned} \\ -2\beta, & s(m_1 + 1, m_2 + 1) \neq s(m_1, m_2) \text{ not aligned} \end{cases}$$
(6.10)

the local factor becomes

$$\mathscr{K}(m_1) \to \begin{cases} 1, & aligned \\ 0, & not aligned \end{cases}$$
(6.11)

Hence, the probability becomes 0 if diagonal spins are of opposite sign. In the particle interpretation, we can say the following:

If  $s(0, m_2) = 1$ , which means that at t=0, there is a particle at x= $\epsilon m_2$ , the particle is at x+ $\epsilon$  at t+ $\epsilon$ . Per time step, every bit configuration simply moves one position to the right. for a system with M=8, we would for instance obtain the following updating:

Our unique jump operator therefore represents a system of fermions moving to the right, which is exactly the behavior we would expect of free fermions:

- Each point x is occupied by at least 1 fermion.
- The system is a many fermion system, with particles at positions with s(x)=1.
- At positions with s(x)=-1, fermions are absent. Analogously, one can interpret this as the presence of holes.
- Since the step evolution matrix of the unique jump chain is a unique jump matrix (the evolution only depends on one neighbour and just copies its configuration) the system is a cellular automaton.

#### • Evolution of the wave function

We denote the configuration of occupation or spin at a given time t with

$$\tau = n(m_2) = s(m_2) \tag{6.12}$$

The components of the wave function after a time step  $\epsilon$  is can be calculated using the unique jump matrix

$$\tilde{q}_{\tau}(t+\epsilon) = \hat{S}_{\tau\rho}\tilde{q}_{\rho}(t) \tag{6.13}$$

$$\bar{q}_{\tau}(t+\epsilon) = \left(\hat{S}^T\right)_{\tau\rho}^{-1} \bar{q}_{\rho}(t) = \hat{S}_{\tau\rho}\bar{q}_{\rho}(t)$$
(6.14)

If now we choose the initial conditions  $\bar{f}_f$  such that  $\bar{q}_{\tau}(M-1)0\tilde{q}_{\tau}(M-1)$ , the components of the wave function and its conjugate are equal and the probability becomes

$$p_{\tau}(t) = \tilde{q}_{\tau}(t)\bar{q}_{\tau}(t) =: q_{\tau}^2(t) \tag{6.15}$$

• In a next step, we want to allow **left and right moving fermions** at once. To allow leftmovers, every site now has two occupation numbers  $n_R(t, x)$ ,  $n_L(t, x)$ . The relevant neighbours for updating cell x are now the diagonally right and left neighbours at  $x+\epsilon$  and  $x-\epsilon$ . However, only one occupation number is copied from the right and left neighbour, respectively. An example is given below:

$\tau(t)$ :	(	0	1	1	0	0	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	R
~ /	ĺ	1	0	T	1	1	0 )	L
$\tau(t+\epsilon):$	(	1	0	1	1	0	0	R
		0	1	1	1	- 0	1 /	L

Since every site now has two occupation numbers which can take values of 0 or 1, every cell at position x of the automaton can have 4 different states.

### 6.2 Particle-hole conjugation and complex structure

#### Particle-hole transformation (or conjugation)

Transforms particles to holes and vice versa:

$$n_{R,L}(x) = 1 \leftrightarrow n_{R,L}(x) = 0$$
  

$$s_{R,L}(x) \leftrightarrow -s_{R,L}(x)$$
(6.16)

The transformation  $n \leftrightarrow 1 - n$  maps  $\tau \leftrightarrow \tau^c$ .

#### Particle-hole transformation (or conjugation)

Conjugation of  $q \leftrightarrow q^c$ 

$$q_{\tau}(t) \leftrightarrow q_{\tau}^{c}(t) \equiv q_{\tau^{c}}(t) \tag{6.17}$$

Later, we will use this map to define complex structure.

#### Divide configurations $\tau$ into two parts:

We can divide all states into two classes that are linked to each other by complex conjugation:

$$\{\tau', \tau^c\}, \tau' \leftrightarrow \tau^c$$
 (6.18)

Each class contains half of the configurations.

Example: For a given configuration  $\tau$ , take the total number of particles:

$$(N_{\rm tot})_{\tau}$$
: Sum over all 1 in  $\tau$  (6.19)

with  $(N_{\text{tot}})_{\text{max}} = 2M$ . Divide the configurations as follows:

$$\begin{split} N_{\rm tot} &> M \to \tau' \\ N_{\rm tot} &< M \to \tau^c \\ N_{\rm tot} &= M, N_R > \frac{M}{2} \to \tau' \\ N_{\rm tot} &= M, N_R < \frac{M}{2} \to \tau^c \end{split}$$

We can combine  $\tau'$  and  $\tau^c$  into a two component notation:

$$q = \begin{pmatrix} q' \\ q^c \end{pmatrix}$$

with

- q: N-component vector,  $N = 4^M$
- $q': \frac{N}{2}$ -component vector
- $q^c$ :  $\frac{N}{2}$ -component vector

In this notation, the conjugation can be achieved by multiplication with the matrix:

$$K = \begin{pmatrix} 0 & \mathbb{1}_{N/2} \\ \mathbb{1}_{N/2} & 0 \end{pmatrix}_{N \times N}$$
(6.20)

The map K has the important property that  $K^2 = \mathbb{1}_N$  and its one of the ingredients we need for a complex structure. Next we will introduce the complex vector.

#### Complex structure:

We define  $\frac{N}{2}$  component complex vector:

$$\phi_{\tau'} = \frac{1+i}{\sqrt{2}}q_{\tau}' + \frac{1-i}{\sqrt{2}}q_{\tau}^c \tag{6.21}$$

The complex conjugation is:

$$\phi_{\tau'}^* = \frac{1-i}{\sqrt{2}}q_{\tau}' + \frac{1+i}{\sqrt{2}}q_{\tau}^c \tag{6.22}$$

Notice that complex conjugation interchanges  $q_{\tau'}$  with  $\tau^c \ (\tau' \leftrightarrow \tau^c)$ .

More formally, we have defined the following map:

$$\phi : \mathbb{R}^N \to \mathbb{C}^{\frac{N}{2}}, \quad q \mapsto \phi(q)$$
 (6.23)

Using the matrix K, we can express the complex conjugation as

$$\phi(Kq) = \phi^*(q) \tag{6.24}$$

#### Multiplication with i:

To complete the complex structure we need to define the multiplication with i in  $\mathbb{R}^N$ .

$$i\phi_{\tau'} = \frac{i-1}{\sqrt{2}}q'_{\tau} + \frac{i+1}{\sqrt{2}}q^c_{\tau} = \frac{i+1}{\sqrt{2}}q^c_{\tau} - \frac{1-i}{\sqrt{2}}q'_{\tau}$$

Notice that

with

$$i\phi_{\tau'} = \phi\left(\begin{pmatrix} q_{\tau}^c \\ -q_{\tau'} \end{pmatrix}\right) = \phi(Iq)$$
$$I = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}_{N \times N}$$
(6.25)

The matrices K, I satisfy

$$K^2 = \mathbb{1}, \quad I^2 = -\mathbb{1}, \quad \{K, I\} = 0$$
 (6.26)

The pair (K, I) defines the complex structure. So far, we have defined the complex structure generally for multiparticle states. Next we look closer at 1-particle states.

#### (time)-Local probabilities:

We define

$$p_{\tau'} = \phi_{\tau'}(t)\phi_{\tau'}(t) = \phi_{\tau'}^*(t)\phi_{\tau'}(t) = q_{\tau'}^2(t) + q_{\tau^c}^2(t)$$
(6.27)

as one typically defines in quantum mechanics  $(\psi \cdot \psi^*)$  with normalization

$$\sum_{\tau'} \phi_{\tau'}(t) \phi_{\tau'}(t) = 1$$
(6.28)

Notice that  $p_{\tau'}$  does not contain all the probabilistic information. It gives you only the probability to have the configuration  $\tau'$  or the configuration  $\tau^c$ . It cannot distinguish particle from holes.

#### Evolution

$$(\tau', \tau^c) \stackrel{\text{evolution}}{\to} (\rho', \rho^c)$$

Generally the evolution doesn't mix the classes except few examples.

$$\hat{S}q = \begin{pmatrix} \hat{S}' & 0\\ 0 & \hat{S}' \end{pmatrix} \begin{pmatrix} q'\\ q^c \end{pmatrix}$$
(6.29)

 $\hat{S}':\frac{N}{2}\times\frac{N}{2}$  real orthogonal matrix. The evolution is particle-hole symmetric.

$$q(t+\varepsilon) = \hat{S}'(t)q(t)$$

Since  $\hat{S}'$  is real and orthogonal, we have unitary time evolution as need in quantum mechanics.

$$\hat{S}'^{\dagger}\hat{S}' = 1$$
, unitary time evolution

Observables

$$\hat{A}(t) = \begin{pmatrix} \hat{A}'(t) & 0\\ 0 & \hat{A}'(t) \end{pmatrix}, \quad (\hat{A}')^T = \hat{A}'$$

Notice that total number of particles  $(N_{\tau})_{\text{tot}}$  is not of this form.

# 6.3 Vacuum

Until now we treated particles as being 1s in our configurations  $\tau$ . We now want to change this perspective and consider particles as excitations of a particular vacuum, which is a general lesson from Quantum Field Theory. The particle depends on the vacuum in the sense that if the vacuum changes so do the properties of the particle.

#### Particle-hole symmetric vacuum

Particle-hole symmetry is an important symmetry of the action. Thus we consider a vacuum state which is invariant under particle-hole conjugation  $\hat{C}$  such that this symmetry is not spontaneously broken.

#### Vacuum configurations

We now look at possible vacuum configurations which allow the construction of a  $\hat{C}$ -invariant vacuum wavefunction.

$$\begin{aligned} \tau_A &: \frac{R \mid 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1}{L \mid 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0} \\ \tau_B &: \frac{R \mid 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0}{L \mid 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1} \end{aligned}$$

$$\tau_B = \tau_A^C, \quad \hat{C} : \tau_A \leftrightarrow \tau_B$$

For both configurations we have  $N_R = N_L = \frac{M}{2}$  i.e. the vacuum is half-filled. In addition to that it is invariant under translations by two slides  $m_2 \to m_2 + 2$ ,  $x \to x + 2\epsilon$ .

#### Vacuum wavefunction

We now want to construct a vacuum wavefunction out of the configurations  $\tau_A$  and  $\tau_B$  which is invariant under  $\hat{C}$ . To this end we define:

$$q^{(A)}: q^{(A)}_{\tau} = \delta_{\tau, \tau_A} \tag{6.30}$$

$$q^{(B)}: q^{(B)}_{\tau} = \delta_{\tau,\tau_B} \tag{6.31}$$

$$q_0 = \frac{1}{\sqrt{2}} \left( q^{(A)} + q^{(B)} \right) \tag{6.32}$$

We make the following observations

- The wavefunction is static i.e.  $q_0(t+\epsilon) = q_0(t)$
- Appart from  $\hat{C}$  there are further transformations which leave  $q_0$  invariant:
  - Space translation  $x \to x + \epsilon$
  - Parity  $\hat{P}$
  - Time reversal  $\hat{T}$
  - $n_R \leftrightarrow n_L$

# 6.4 One-particle wave function

#### One particle configurations

We now want to describe a single particle as excitation of the vacuum defined above. The idea is that we add a single Right-/Leftmover at a particular position  $m_2$ . There we be always already a 1 at position  $m_2$  in either  $\tau_A$  or  $\tau_B$ . We add a 1 in the configuration which does not yet have a 1 at position  $m_2$ .

$$\tau_{R,m_2} : \frac{R \mid 0 \quad 1 \quad 1 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1}{L \mid 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0}, \ m_2 = 2$$
  
$$\tau_{R,m_2} : \frac{R \mid 1 \quad 0 \quad 1 \quad 1 \quad 1 \quad 0 \quad 1 \quad 0}{L \mid 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1}, \ m_2 = 3$$

and similarly for  $\tau_{L,m_2}$ .

#### One hole configurations

We can do the same for the addition of a hole, e.g.

$$\tau_{R,m_2}^C : \frac{R | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |$$

and again similarly for left moving holes  $\tau_{L,m_2}^C$ .

The structure of the vacuum leads to the idea of a **generalized particle** or **quantum particle**. By this we mean either a single additional particle or hole.

#### Basis function in position basis in one-particle sector

We define basis vectors  $g_{m_2}^{(R,L)}, g_{m_2}^{(R,L)}$  which either add (g) or subtract (h) a single fermion/bit.

$$\left(g_{m_2}^{(R)}\right)_{\tau} = \delta_{\tau,\tau_{R,m_2}}, \ \left(h_{m_2}^{(R)}\right)_{\tau} = \delta_{\tau,\tau_{R,m_2}^C} \tag{6.33}$$

$$\left(g_{m_2}^{(L)}\right)_{\tau} = \delta_{\tau, \tau_{L, m_2}}, \ \left(h_{m_2}^{(L)}\right)_{\tau} = \delta_{\tau, \tau_{L, m_2}^C} \tag{6.34}$$

With help of this basis we can write the one-particle wavefunction as

$$q_{m_2}^{(1)} = q_R'(m_2)g_{m_2}^{(R)} + q_R^C(m_2)h_{m_2}^{(R)} + q_L'(m_2)g_{m_2}^{(L)} + q_L^C(m_2)h_{m_2}^{(L)}$$
(6.35)

or

$$q(m_2) = \begin{pmatrix} q'_R(m_2) \\ q^C_R(m_2) \\ q'_L(m_2) \\ q^C_L(m_2) \end{pmatrix}$$
(6.36)

where we can interpret the coefficients as

 $(q_R'(m_2))^2$ : probability to find right-moving particle at  $m_2$  $(q_R^C(m_2))^2$ : probability to find right-moving hole at  $m_2$ 

Also they fulfill the condition

$$\sum_{m_2} q^T(m_2)q(m_2) = 1 \tag{6.37}$$

#### **Complex picture**

In the complex picture the one particle wavefunction becomes a two-component complex vector depending on position:

$$\varphi(m_2) = \begin{pmatrix} \varphi_R(m_2) \\ \varphi_L(m_2) \end{pmatrix}$$
(6.38)

$$\varphi_{R,L}(m_2) = \frac{1+i}{\sqrt{2}} q'_{R,L}(m_2) + \frac{1-i}{\sqrt{2}} q^C_{R,L}(m_2)$$
(6.39)

We see that the components contain both  $q'_{R,L}$  and  $q^C_{R,L}$  so that it describes both particles and antiparticles simultaneously.

Antiparticles emerge naturally for this simple model

# 6.5 Continuum limit

So far we saw that our simple Ising model for fermions naturally includes the notion of particles and antiparticles, which arise with the complex structure. We began in the real case constructing a theory symmetric in the particle-hole-picture. The wave functions for holes were identified by subtracting particles. In the complex notion we arrived at a four-component wave function:

$$\phi(m_2) = \begin{pmatrix} \phi_R(m_2) \\ \phi_L(m_2) \end{pmatrix} \tag{6.40}$$

which contains both non-conjugated and conjugated components. These relate to particles and antiparticles respectively.

By now our model is discrete, a 2d lattice with lattice points denoted  $m_k$ . In this section we want to develop a formulation in the continuum limit, assuming the limit of an infinite number of lattice points, simultaneous to a vanishing lattice constant. Of main interest is the time evolution in the continuum limit, where we expect the 2d Dirac equation, as we are dealing with fermions.

The continuum limit of the wave function and the configuration sums is formally defined by the normalization conditions:

$$\phi(x = \epsilon m_1) = \frac{1}{\sqrt{\epsilon}}\phi(m_2) \tag{6.41}$$

$$\int_{x} \equiv \int dx = \epsilon \sum_{m_2} \tag{6.42}$$

Given these formal definitions the continuum can be taken now.

#### **Operators:**

Let us begin with the operators. Transporting into the continuum limit works with the relation  $x = \epsilon m_2$ :

• For both the right- and left-moving particles we obtain the diagonal occupation number operators:

$$\begin{pmatrix} \hat{N}_R(m_2) & 0\\ 0 & 0 \end{pmatrix} \tag{6.43}$$

$$\begin{pmatrix} 0 & 0\\ 0 & \hat{N}_L(m_2) \end{pmatrix} \tag{6.44}$$

with  $(\hat{N}_{R/L}(m_2))_{mm'} = \delta_{mm_2} \delta_{m'm_2}$ . The expectation value of the particle occupation can be calculated straightforwardly from the statistical law for expectation values:

$$\langle N_{R/L}(m_2) = \phi^{\dagger} \hat{N}_{R/L}(m_2)\phi = \phi^{\dagger}(m_2)\phi(m_2)$$
 (6.45)

This is the probability to find either an right-/left-moving particle or right-/left-moving antiparticle at  $m_2$ . Observe that particle and antiparticle can in this context not be distinguished and both contribute. Without any other quantum numbers, like charge, this makes sense.

Of course the concept of occupation numbers remains the same going into the continuum limit. The lattice points as occupation sites are replaced by the positions in the continuous space, connected by the relation  $x = \epsilon m_2$ .

• Next the position operator, which is defined as the sum over all position sites weighted with the occupation. Thus:

$$\hat{X} = \epsilon \sum_{m_2} m_2 \left[ \hat{N}_R(m_2) + \hat{N}_L(m_2) \right]$$
(6.46)

The expectation value of this operator is determined easily to be:

$$\langle X \rangle = \epsilon \sum_{m_2} m_2 \left[ \psi_R^{\dagger}(m_2) \phi_R(m_2) + \psi_L^{\dagger}(m_2) \phi_L(m_2) \right] = \sum_{m_2} m_2 \phi^{\dagger}(m_2) \phi(m_2)$$
(6.47)

With the formal definition of the wave function and configuration sum in the continuum limit, this becomes:

$$\langle X \rangle = \int_{x} x \ \phi^{\dagger}(x) \phi(x)$$
 (6.48)

This the same expression as in quantum mechanics.

• The last operator considered here, is the step-evolution operator. For right-/ left-moving particles the real wave functions evolve like:

$$q_R(t+\epsilon, x) = q_R(t, x-\epsilon) \tag{6.49}$$

$$q_L(t+\epsilon, x) = q_L(t, x+\epsilon) \tag{6.50}$$

so that the step evolution operator is the diagonal operator:

$$\hat{S} = \begin{pmatrix} \hat{S}_R & 0\\ 0 & \hat{S}_L \end{pmatrix} \tag{6.51}$$

where  $(\hat{S}_R)_{m_2,m'_2} = \delta_{m_2,m'_2+1}$  and  $(\hat{S}_L)_{m_2,m'_2} = \delta_{m_2,m'_2-1}$ . The time evolution operator in the continuum case can, in many cases, be related to the Hamilton operator. So that deriving the equation of motion, and from that the Hamilton operator, comes next.

#### Equation of motion:

In the discrete case time evolution implies the relations:

$$\phi_R(t+\epsilon, x+\epsilon) = \phi_R(t, x) \tag{6.52}$$

$$\phi_L(t+\epsilon, x-\epsilon) = \phi_L(t, x) \tag{6.53}$$

or equivalently:

$$\phi_R(t+\epsilon, x+\epsilon) - \phi_R(t, x) = 0 \tag{6.54}$$

$$\phi_L(t+\epsilon, x-\epsilon) - \phi_L(t, x) = 0 \tag{6.55}$$

These expressions can be related to what is called lattice derivatives, which are similarly defined to partial derivatives, but in the discrete case, where a function is only defined at lattice points.

**Definition 4** (Lattice derivative). Let f(x,y) be defined on a 2d lattice with lattice constant  $\epsilon > 0$ , the lattice derivative with respect to x is then defined as:

$$\partial_x f(x,y) := \frac{1}{\epsilon} \left[ f(x+\epsilon,y) - f(x,y) \right]$$

Given the definition of a lattice derivative we can write:

$$0 = \frac{1}{\epsilon} \left[ \phi_R(t+\epsilon, x+\epsilon) - \phi_R(t, x) \right] = (\partial_t + \partial_x) \phi_R(t, x)$$
(6.56)

$$0 = \frac{1}{\epsilon} \left[ \phi_L(t+\epsilon, x-\epsilon) - \phi_L(t, x) \right] = (\partial_t - \partial_x) \phi_L(t, x)$$
(6.57)

The important observation is now that if the wave function, which only need to be defined at the lattice points,  $\phi(x)$  behaves sufficiently smooth, then we obtain nothing else than partial derivatives in the continuum limit. The equation of motion in the continuum limit  $\epsilon \to 0$ , simultaneously the number of lattice points goes to infinity so that the lattice remains of finite size, reads:

$$(\partial_t + \tau_3 \partial_x)\phi(t, x) = 0$$
(6.58)

 $\tau_3$  denotes the third Pauli matrix:

$$\tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which is needed due to the opposite signs in the spatial derivative for right- and left-moving particles in (6.56) and (6.57). The resulting equation of motion is the Dirac equation in 1+1 dimensions, for free and massless particles. It can be explicitly written as:

$$\gamma^{\mu}\partial_{\mu}\phi(t,x) = 0 \tag{6.59}$$

where:

$$\gamma^{1} = -i\tau_{2} = \begin{pmatrix} 0 & 1\\ 1 & 0 \end{pmatrix}$$
$$\gamma^{2} = \tau_{1} = \begin{pmatrix} 0 & 1\\ -1 & 0 \end{pmatrix}$$

The Hamiltonian can then be written off to be:

$$\hat{H} = \begin{pmatrix} \hat{P} & 0\\ 0 & -\hat{P} \end{pmatrix} \tag{6.60}$$

with the, from quantum mechanics, well known momentum operator:

$$\dot{P} = -i\partial_x \tag{6.61}$$

And so the energy-momentum relation reads  $E^2 = p^2$ , which is the relativistic energy-momentum relation for massless particles.

From quantum mechanics it is known that the time evolution, given the Hamilton operator, can be expressed with the unitary time evolution operator:

$$\hat{S}(t,t_0) = T \exp\left\{-i \int_{t_0}^t d\tau \hat{H}(\tau)\right\}$$
(6.62)

where T denotes time ordering. The step evolution operator for infinitesimally steps  $\epsilon > 0$  is then:

$$\hat{S}(t) = \exp\left\{-i\epsilon\hat{H}(t)\right\}$$
(6.63)

Conclude the Hamilton operator of the continuum theory and the step evolution operator in the discrete case are connected to each other, just by this relation.

We summarize that our 2d Ising model for fermions can not only be formulated with a complex structure, but also yields the quantum mechanical Dirac equation in the continuum limit. There are two important remarks:

1. The Dirac equation is a Lorentz covariant equation, it is the fermionic wave equation of relativistic quantum mechanics. It is remarkable that, even though we started with a discrete 2d lattice, we obtain Lorentz covariance without explicit construction. In the end it is a result of the step evolution in the discrete case which is of equal order in time and space. Recall that, like the vacuum constructed in 6.3, our model is CPT symmetric.

2. Introducing the continuum limit is only possible if the wave equations behave sufficiently smooth, so that partial derivatives are obtained. If the microscopical theory on a lattice was deterministic, the continuum limit would not be easy to construct. As deterministic behaviour would intuitively be formulated in delta distributions. The microscopic theory being inherently probabilistic is the key assumption of this lecture. Like Ising spins, we view particles as probabilistic entities in this lecture. This also enters in the wave equations being well behaved in the continuum limit, because information is "smeared" out.

#### Momentum operator

$$\hat{P} = -i\partial_x$$
 : hermitean matrix ;  $\hat{P}(x,y) = -i\delta(x-y)\frac{\partial}{\partial y}$  (6.64)

 $\hat{P}$  does not commute with  $\hat{X} = x$ ;  $\hat{X}(x, y) = \delta(x - y)y$ . The commutator of X and P is:  $[\hat{X}, \hat{P}] = i$ 

Is there a momentum <u>observable</u> to which the operator  $\hat{P}$  is associated?

Measurement prescription such that the possible outcomes of individual measurements are eigenvalues of  $\hat{P}$ ?

• Momentum eigenstates:

$$\hat{P}\varphi_p(x) = p\varphi_p(x) \tag{6.65}$$

$$\varphi_p(x) = \frac{1}{\sqrt{L}} e^{(ipx)} \tag{6.66}$$

where L is the circumference of form m x

How to extract p?

Periodic function with period  $\Delta = \frac{2\pi}{|p|}$ 

For periodic wave function and probability distribution: Determine  $\Delta \Rightarrow$  Determine p: "measurement"

• <u>Statistical observable</u>:

P u "statistical observable"

It's measurement determines properties of the probabilistic information (in our case periodicity)

P has no well defined value for a given bit-configuration.

(Recall: one-particle bit configuration has single additional or missing bit at fixed location  $m_2$ 

Similar to temperature T: Property of probabilistic state, <u>not</u> of given microstate

• Absence of classical correlation function:

Values  $P_{\tau}$  in configuration  $\tau : \underline{\text{not}}$  defined  $\Rightarrow X_{\tau}P_{\tau}: \underline{\text{not}}$  defined  $\langle XP \rangle_{cl} = \sum_{\tau} p_{\tau}X_{\tau}P_{\tau}: \underline{\text{not}}$  defined  $\Rightarrow$  Bell's inequations do not apply to pair of observables (X, P)
• <u>Real formulation</u>: (L = 1)

$$\varphi_p = e^{ipx} = \cos(px) + i\,\sin(px) \tag{6.67}$$

$$Re \ \varphi_p = cos(px) = \frac{1}{\sqrt{2}}(q' + q'')$$
 (6.68)

$$Im \ \varphi_p = \sin(px) = \frac{1}{\sqrt{2}}(q' - q'') \tag{6.69}$$

use:  $\varphi = \frac{1+i}{\sqrt{2}}q' + \frac{1-i}{\sqrt{2}}q''$  with:  $q' = \frac{1}{\sqrt{2}}(\cos(px) + \sin(px)), \ q'' = \frac{1}{\sqrt{2}}(\cos(px) - \sin(px))$ Oscillation between particle and hole

$$q'^2 + q''^2 = \frac{1}{L^2}, \qquad \int_x (q'^2 + q''^2) = 1, \qquad \hat{P} = -I\partial_x, \qquad I = \begin{pmatrix} 0 & 1\\ -1 & 0 \end{pmatrix}$$

• <u>Discrete formulation</u>:

$$\hat{P}(m,n) = -\frac{1}{4\epsilon} I(\delta_{m,n-1} - \delta_{m,n+1})$$
(6.70)

 $\begin{array}{ll} \hat{P}^T = \hat{P}, & p = \frac{2\pi q}{\epsilon M_x}, & \text{q integer}, & L = \epsilon M_x \\ \text{Eigenvalues of } \hat{P}: \ \frac{1}{2\epsilon} sin(\frac{4\pi q}{M_x}) = \frac{1}{2\epsilon} sin(2p\epsilon), & e^{ipx} = e^{\frac{2\pi i q m_2}{M_x}} \\ \text{periodicity } x \to x + \Delta = x + \frac{2pi}{p}, & m_2 \to m_2 + \Delta_m = m_2 + \frac{2\pi}{\epsilon p} = m_2 + \frac{M_x}{q} \end{array}$ 

#### 6.6 Interactions

- Probabilistic automata can be equivalent to quantum field theories for fermions
- Not only free fermions; interactions can be described aswell
- In order to describe the interactions we introduce a Turing automaton, for the reason that its continuum limit is a type of two-dimensional Turing Models. Turing Models are the most simple fermionic Quantum Field Theory and they can be solved partially exactly. So we define the Turing automaton in 1+1 dimensions (1 time, 1 space dimensions)
- The basic setting is that we will still have left-movers and right-movers but they will come in two colours (let us say them red and green) and the basic idea is that we will have sequence of two steps:
  - In the first step we have the propagation (left-movers to the left and right-movers to the right).
  - In the second step they interact, which means they exchange their colours.
- We define  $\hat{S}$  which is a unique matrix and describes automaton. This matrix has the important property of particle-hole symmetry.

#### 6.6.1 Quantum Field Theory for fermions

Finally, we will examine something that we did not extend in the previous lectures and this is an exact map from this type of generalized Ising Models to Grassmann Functional integrals with Grassmann variables.

With the help of Grassman numbers we are able to describe four fermions interaction in the continuum limit as a Turing type model:

$$S = \int_{t,x} \{ -\bar{\psi}_a \gamma^\mu \partial_\mu \psi_a - \frac{1}{2} \bar{\psi}_a \gamma^\mu \psi_a \bar{\psi}_b \gamma_\mu \psi_b + \bar{\psi}_a \gamma^\mu \psi_b \epsilon^{ab} \bar{\psi}_c \gamma_\mu \psi_c \epsilon^{dc} \}$$
(6.71)

The above expression describes a cellular automata that corresponds to a Quantum Field Theory in two dimensions for fermions. It is actually Lorentz invariant, it satisfies particle-hole symmetry, SO(2)-colour symmetry, Parity, Charge conjugation and Time reversal (P,C,T).

Central conclusion: Classical Probabilistic Systems (as the generalized Ising Model that we studied) are equivalent to Quantum Field Theory for interacting fermions. In order to make clear this claim more clear we will present the one-to-one map between these two formulations:

- 1. Same wave function
- 2. Same evolution of the wave function
- 3. Same operators for the observables

A very direct example is Quantum mechanics which are generated from classical statistics, without any addition of new "quantum" axioms. In addition, we get quantum rules from "classical" probabilistic rules. An important consequence of this correspondence is that if we have an example, which can be a no-go theorem that we cannot get Quantum Mechanics from Classical Statistics, it cannot be generally applicable. Finally, since there is not internal contradiction between CS and QM there is no need to modify Quantum Mechanics.

### Chapter 7

## Reduction of the wave function and conditional probabilities

#### 7.1 Reduction of the wave function

Let's consider a single spin quantum system. Its time evolution is given by the Schrödinger equation

$$i\partial_t \phi = H\phi \tag{7.1}$$

with the Hamilton operator

$$H = \mu \vec{B}(t)\vec{\tau} \tag{7.2}$$

To describe the system, we choose a basis of  $\tau_3$  eigenstates.

Let's further imagine that at a certain time t, the measurement of  $s_z(t)$  yields +1 (where we take Spin  $\vec{S} = \frac{\hbar}{2}\vec{s}$  to spare us some prefactors).

Immediately after t, the wave function has to be reset to an eigenfunction of  $s_z$  since a second measurement with vanishing time difference should yield the same result. From there on, the wave function again evolves as given by the Schrödinger equation, and the outcome of a measurement of  $s_x$  after a time intervall  $\Delta$  is given by

$$\langle s_x(t+\Delta) \rangle = \phi^{\dagger}(t+\Delta)\tau_1\phi(t+\Delta)$$
(7.3)

The puzzling part lies in the measuring at time t: Before the measurement, the system is described by its wave function  $\phi(t-\epsilon)$ , which is not necessarily the eigenstate of  $s_z$  with  $s_z = 1$ . Only by the measurement, the system is set to this fixed state, which then again evolves in time:

$$\phi(t-\epsilon) \xrightarrow{Reduction of the wave function} \phi(t+\epsilon) = \begin{pmatrix} 1\\ 0 \end{pmatrix} \text{ for } \epsilon \to 0$$

The reduction of the wave function at time t corresponds to a jump and therefore a discontinuity in the wave function, which hurts unitary time evolution. This is even worse for the case of a mixed state density matrix, which after a measurement is in a pure state:

$$\begin{array}{c} \rho(t-\epsilon) \\ mixed \ state \end{array} \xrightarrow{RWF} \rho(t+\epsilon) \\ pure \ state \end{array}$$

For this reason, many researches proposed that Quantum mechanics would have to be modified. in this chapter, we will show that this puzzle disappears by correct use of some probabilistic concepts.

#### 7.2 Conditional probabilities

Coming back to the two-spin system in 7.1, the expectation value of  $s_x$  after  $\Delta$  can also be expressed as

$$\langle s_x(t+\Delta) \rangle = p_+ - p_-$$
 (7.4)

where  $p_+$  denotes the probability to measure  $s_x(t + \Delta) = 1$  given the condition that  $s_z(t) = 1$ . To make this dependence on the condition a bit more obvious, we introduce the notation

$$p_+ \rightarrow \left( \begin{array}{c} \omega_+ \\ prob. \ for \ s_x=1 \end{array} \right)_{given \ s_z(t)=1} +$$

If at t we would have measured  $s_z = -1$ , we would expect a different value  $\langle s_x(t + \Delta) \rangle$  since  $\phi(t + \epsilon) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  and we would express the corresponding probabilities by  $(\omega_{\pm})_{-}$ . For more general observables with corresponding eigenvalues  $\lambda_{A/B}$  we would write  $(\omega_{\lambda_A}^A)_{\lambda_B}^B$  if the measurement of B is taken first. Conditional probabilities have to be normalized as well:

$$\sum_{\lambda_A} (\omega_{\lambda_A}^A)_{\lambda_B}^B = 1 \tag{7.5}$$

One can very generally say that Physics describe conditional probabilities rather than absolute overall probabilities. For a certain experiment, the initial conditions need to be prepared.

#### 7.3 Measurement correlation

We now define  $\omega_{+-}^{AB}$  to be the overall probability that  $s_z(t) = -1$  and  $s_x(t + \Delta) = +1$ . With this, we can define the measurement correlation

$$_{m} = \omega_{++}^{AB} + \omega_{--}^{AB} - \omega_{+-}^{AB} - \omega_{-+}^{AB}$$
 (7.6)

Since e.g.

$$\omega_{++}^{AB} = (\omega_{+}^{A})_{+}^{B} \cdot p_{+}(t) \tag{7.7}$$

with the probability  $p_+(t)$  to measure  $s_z(t) = 1$ , the measurement correlation involves conditional probabilities:

$$\langle AB \rangle_{m} = (\omega_{+}^{A})_{+}^{B} \cdot p_{+}(t) + (\omega_{-}^{A})_{-}^{B} \cdot p_{-}(t) - (\omega_{+}^{A})_{-}^{B} \cdot p_{-}(t) - (\omega_{-}^{A})_{+}^{B} \cdot p_{+}(t)$$
(7.8)

$$= \left( (\omega_{+}^{A})_{+}^{B} - (\omega_{-}^{A})_{+}^{B} \right) \cdot p_{+}(t) + \left( (\omega_{-}^{A})_{-}^{B} - (\omega_{+}^{A})_{-}^{B} \right) \cdot p_{+}(t)$$
(7.9)

The reduction of the wave function therefore gives a prescription of the initial probabilities and hence also on the measurement correlation.

#### 7.4 Non-uniqueness of conditional probabilities

As an illustrative example, let's imagine a course of physics students all performing the same experiment. Every student is asked to determine his or her measurement correlation with sufficiently high statistics. Comparing their results, the students find out that their results do not coincide. Why is this the case? Some students could have prepared their measurement apparatus in a more precise way than others, and this would effect their measurement correlation. It is worth to mention that the measurement correlation for each single experiment is still a well-defined quantity. But: Conditional probabilities and measurement correlation depend on the setting of the experiment. They involve

- 1. the probabilistic information of the quantum system
- 2. the measurement process

The measurement correlation therefore is no quantity that can be calculated from the information about the quantum two state system alone. It also involves information about the measurement apparatus. The reduction of the wave function seen above describes some kind of 'ideal' measurement.

For many spin quantum systems, the reduction of the wave function is not unique, as the following example shall illustrate: Consider a quantum system with ten spins. At the time t, one measures the value of the second spin  $s_z^{(2)}$  and therefore sets the system to an eigenfunction of  $s_z^{(2)}$  with the measured value. But there are now many eigenfunctions with the corresponding eigenvalue, so it is not clear to which function the system should be reduced to. Uniqueness is only given if one measures a maximal set of commutative ovservables at the time t, which is not the case for most problems.

The reduction of the wave function also appears in *classical statistics*: If one measures an observable B with a spectrum -1, 1 and obtains B=1 at t, one sets all probabilities  $p_{\tau}(t) = 0$  for all states  $\tau$  which correspond to values of B=-1. This resetting needs a new normalisation of the remaining probabilities. As in Quantum Mechanics, this procedure is not unique.

#### Summary:

Conditional probabilities, measurement correlations and reduction of wave function appear both in Quantum mechanics and Classical statistics. They are not universal but rather involve specifications for the measurement process. We thus also have to get rid of the common conception that the reduction of wavefunction possesses absolute conceptual status, which it does not.

#### 7.5 Ideal measurements

We now define the criteria under which we call a measurement an "ideal measurement":

• **Independence of environment:** The outcome of the measurement should only depend on probabilistic information of the system, i.e. the measurement correlation and conditional probabilities should be computable from the density matrix of the system.

Details of the state of the apparatus should not influence the outcome of the measurement. Still, the interaction of the system with the apparatus can play a role for the outcome of a sequence of measurements. • **Reproducibility:** A sequence of two "immediate" measurements of the same observable should yield identical results, i.e.

$$\Delta \to 0: (w_{+}^{A})_{-}^{A} = (w_{-}^{A})_{+}^{A} = 0$$

More precisely, if T is a typical time scale for the time evolution of the quantum system according to the Schrödinger equation,  $\Delta \rightarrow 0$  means  $\delta \ll T$ 

The two criteria do not fix uniquely the notion of an ideal measurement. We will discuss two types of ideal measurements:

- Coherent ideal measurements
- Incoherent ideal measurements

In general they lead to different measurement correlations.

#### 7.6 Unequal time quantum correlations

In Quantum mechanics we can compute the quantum correlation function

$$\langle s_x(t+\Delta)s_z(t)\rangle_Q$$

We know want to know the relation to the measurement correlation according to the reduction of the wavefunction (which both correspond to different types of ideal measurements). In the Heisenberg picture we can write

$$C = \langle s_x(t+\Delta)s_z(t) \rangle = \varphi^{\dagger}(t)U^{\dagger}\tau_1U\tau_3\varphi(t)$$
  
Reference time  $t$ ,  $s_z(t) \to \tau_3$ ,  $s_x(t) \to U^{\dagger}\tau_1U$   
 $U = U(t+\Delta,t)$ ,  $\varphi(t) = \begin{pmatrix} \varphi_+\\ \varphi_- \end{pmatrix}$   
 $\Rightarrow C = (\varphi^*_+, \varphi^*_-)(U^{\dagger}\tau_1U)\begin{pmatrix} \varphi_+\\ -\varphi_- \end{pmatrix}$   
 $\hat{A} = U^{\dagger}\tau_1U = \begin{pmatrix} \hat{A}_{++} & \hat{A}_{+-}\\ \hat{A}_{-+} & \hat{A}_{--} \end{pmatrix}$ 

$$\begin{aligned} \langle s_x(t+\Delta)s_z(t)\rangle_Q &= \hat{A}_{++}|\varphi_+|^2 - \hat{A}_{--}|\varphi_-|^2 + \hat{A}_{-+}\varphi_-^*\varphi_+ - \hat{A}_{+-}\varphi_+^*\varphi_-\\ |\varphi_+|^2 &= p_+(t), \quad |\varphi_-|^2 = p_-(t) \end{aligned}$$

In comparison, the conditional probabilities according to the reduction of the wavefunction are given by

$$\langle s_x |_{s_z=1} \rangle = (w_+)_+ - (w_-)_+ = (1,0)U^{\dagger}\tau_1 U \begin{pmatrix} 1\\0 \end{pmatrix} = \hat{A}_{++} \\ \langle s_x |_{s_z=-1} \rangle = (w_+)_- - (w_-)_- = (0,1)U^{\dagger}\tau_1 U \begin{pmatrix} 0\\1 \end{pmatrix} = \hat{A}_{--}$$

Therefor the measurement correlation is given by

$$\langle s_x(t+\Delta)s_z(t)\rangle_{RWF} = p_+(t)\hat{A}_{++} - p_-(t)\hat{A}_{--}$$

This differs from quantum correlation if

$$\hat{A}_{-+}\varphi_{-}^{*}\varphi_{+} - \hat{A}_{+-}\varphi_{+}^{*}\varphi_{-} \neq 0$$

Which type of correlation we should use, always depends on the type of the measurement.

- Coherent ideal measurement: Interaction with apparatus does not influence the quantum state. ⇒ Measurement correlation = quantum correlation. Here the reduction of wavefunction does not apply.
- Coherent ideal measurements are for example used for test of Bell's inequalities. Here it is especially important that no quantum correlations are affected by the measurement.

# 7.7 Decoherent ideal measurements and reduction of the wave function

The starting point is the density matrix immiadely after the first measurement,

$$\rho(t_+) = \rho(t+\epsilon), \ \epsilon \to 0^+$$

We would like to compute the correlation function

$$\langle s_x(t+\Delta)s_z(t)\rangle = \operatorname{tr}\left(\hat{A}\tau_3\rho(t_+)\right)$$

with

$$\rho(t_{+}) = \begin{pmatrix} p_{+}(t) & c_{+} \\ c_{+}^{*} & p_{-}(t) \end{pmatrix}$$

Recall

$$A = \begin{pmatrix} A_{++} & A_{+-} \\ A_{-+} & A_{--} \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Then

$$\tau_3 \rho(t) = \begin{pmatrix} p_+ & c_+ \\ -c_+^* & -p_- \end{pmatrix}$$

and since we are interested in the trace we will only calculate the diagonal elements,

$$\hat{A}\tau\rho(t_{+}) = \begin{pmatrix} \hat{A}_{++}p_{+} - \hat{A}_{+-}c_{+}^{*} & \dots \\ \dots & \hat{A}_{-+}c_{+} - \hat{A}_{--}p_{-} \end{pmatrix}$$

$$\langle s_{x}(t+\Delta)s_{z}(t)\rangle_{m} = \hat{A}_{++}p_{+} - \hat{A}_{--}p_{-} + \hat{A}_{-+}c_{+} - \hat{A}_{+-}c_{+}^{*}$$

We can now compare that with

$$\langle s_x(t+\Delta)s_z(t)\rangle_{RWF} = p_+(t)\hat{A}_{++} - p_-(t)\hat{A}_{--}$$

For the quantum measurement,  $c_{+} \neq 0$  in contrast to RWF where  $c_{+} = 0$ .

Immediately before the first measurement

$$\rho(t_-) = \begin{pmatrix} p_+(t) & c_- \\ c_-^* & p_-(t) \end{pmatrix}$$

The diagonal elements must be the same in order to have an ideal quantum measurement, but the off diagonal elements are in general different.

In the coherent situation, there is no jump in the wave function,  $c_{+} = c_{-}$ , and the measurement has no influence on the density matrix. Then you get the measurement correlation corresponding to the unequal time correlation.

But if after the measurement we set  $c_+ = 0$  independently of  $c_-$  you get the reduction of the wave function and this is what we mean by decoherence.

By decoherence, one loses the phase information on the quantum correlations. The density matrix starts from a pure state density matrix  $\rho(t_{-})$  and after the first measurement the density matrix is  $\rho(t_{+})$ . If  $c_{+} = 0$  information is lost. This is what happens in an incoherent ideal measurement as we have seen for the Stern-Gerlach when we neglect the other beam.

The reduction of the wave function is a convenient ideal tool for describing these ideal measurements: book keeping of conditional probabilities for decoherent ideal measurements.

#### 7.8 Bell's inequalities

Consider two observables A,B with possible measurement values  $A_i, B_j$ . Then we can define what is a generalized classical correlation function as

$$\langle AB \rangle_{\rm cl} = \sum_{i,j} w_{i,j}^{AB} A_i B_j, \quad w_{i,j}^{AB} \ge 0, \quad \sum_{i,j} w_{i,j}^{AB} = 1$$

 $w_{i,j}^{AB}$  :simultaneous probabilities for A and B

• Classical correlation  $\langle A|B\rangle_{\rm cl}$ 

Measurement correlations for ideal measurements on quantum mechanics violate sometimes Bell's inequalities. • No go theorem for embedding QM in CS.

Measurement correlations in quantum subsystems are not given by classical correlation functions.

Example: Classical correlation functions are not defined.(e.g statistical observables)

• Equivalent observables

B and B' are mapped to the same quantum operator  $\hat{B}$  . . .

$$\begin{array}{ll} A\mapsto \hat{A}\\ \langle A\rangle, & \langle B\rangle=\langle B'\rangle, & \langle AB\rangle\neq \langle A'B'\rangle \end{array}$$

Incomplete statistics for subsystems

<u>Conclusion</u> Quantum mechanics can be embedded in classical statistics, but quantum subsystems is characterized by incomplete statistics.