Chapter 2

Generalized measurements

Figure 2.1: Stern-Gerlach experiment

2.1 Stern-Gerlach experiment

In order to understand the need of introducing the concept of generalized measurement, let us start with a detailed analysis of a model of the Stern-Gerlach experiment, see Fig. 2.2. Consider a spin $\frac{1}{2}$ particle, initial state of which is given by:

$$|\psi\rangle = |s\rangle \otimes |\phi\rangle_z, \tag{2.1}$$

where $|s\rangle$ represents a general spin state of the particle, while $|\phi\rangle_z$ is its spatial wave function, which for the purpose of this model we restrict to represent solely the z coordinate degree of freedom. The particle travels in the x direction (the movement in the x direction we treat classically) and enters a nonuniform magnetic field at the origin of the coordinate frame, which can be approximated as [1]:

$$\vec{B} \approx (B_0 + kz)\hat{e}_z. \tag{2.2}$$

The Hamiltonian that describes the interaction of the particle with the magnetic field is given by

$$H = -\mu\vec{\sigma} \cdot \vec{B}, \tag{2.3}$$

where $\vec{\sigma} = [\sigma_x, \sigma_y, \sigma_z]^T$ is a vector of Pauli matrices. We assume that the interaction lasts for a short time δt during which the particle goes through the magnetic field. After this time the interaction is not present anymore. In order to further simplify the formulas we will set $B_0 = 0$. Before the interaction takes place, the initial state of the particle corresponds to

$$|s\rangle = c_+|+\tfrac{1}{2}\rangle_z + c_-|-\tfrac{1}{2}\rangle_z, \tag{2.4}$$

$$|\phi\rangle_z = \frac{1}{\sqrt[4]{2\pi\sigma^2}} \int \mathrm{d}z\, e^{-\frac{z^2}{4\sigma^2}} |z\rangle, \tag{2.5}$$

where the general spin state is written in terms of states with a definite spin projection on the z axis, while the initial spatial wave function is assumed to be gaussian with a mean deviation σ. In what follows we will for simplicity denote $|\pm\rangle := |\pm\frac{1}{2}\rangle_z$. Let us evolve the state for time δt under the action of Hamiltonian H (we ignore particle free evolution, on the grounds that δt is sufficiently small):

$$|\psi(\delta t)\rangle = e^{-\frac{iH\delta t}{\hbar}}|\psi\rangle = c_+|+\rangle|\phi_+\rangle + c_-|-\rangle|\phi_-\rangle, \tag{2.6}$$

where

$$|\phi_\pm\rangle = \tfrac{1}{\sqrt[4]{2\pi\sigma^2}} \int \mathrm{d}z\, e^{-\frac{z^2}{4\sigma^2} \pm \frac{i\mu k\delta t}{\hbar} z}|z\rangle. \tag{2.7}$$

[1]This is just an approximation, since according to Maxwell's equations magnetic field has to satisfy $\vec{\nabla}\vec{B} = 0$, while these field clearly does not satisfy this condition. In reality we would need to take into account other magnetic field components to satisfy the divergence-free requirement

In order to interpret the above states, let us write them in the momentum representation:

$$|\phi_\pm\rangle = \frac{1}{\sqrt[4]{2\pi\sigma_p^2}} \int \mathrm{d}p\, e^{-\frac{(p\mp\delta p)^2}{4\sigma_p^2}} |p\rangle, \qquad (2.8)$$

where $\sigma_p = \frac{\hbar}{2\sigma}$ is the width of the gaussian wave packet in the momentum representation, while $\delta p = \mu k \delta t$ represents the momentum kick experienced by the particle.

We see that as a result of the evolution, the spin and spatial degrees of freedom become entangled and the particle experiences a momentum kick that depends on the projection of its spin on the z axis. If we now measure the momentum of the particle we will learn some information on the spin state

Let us assume we have performed an ideal projective measurement of the momentum of the particle [2]. The probability of obtaining result p can be calculated using the following formula (watch out for a slight abuse of notation...):

$$p(p) = \langle\psi(\delta t)|\mathbb{1} \otimes |p\rangle\langle p||\psi(\delta t)\rangle, \qquad (2.9)$$

where the identity reminds us that we do not measure the spin states of the particle directly. Explicitly it reads:

$$p(p) = \frac{1}{\sqrt{2\pi\sigma^2}} \left(|c_+|^2 e^{-\frac{(p-\delta p)^2}{2\sigma_p^2}} + |c_-|^2 e^{-\frac{(p+\delta p)^2}{2\sigma_p^2}} \right) \qquad (2.10)$$

We see that the probability distribution is given solely in terms of the spin degrees of freedom parameters and the measurement result p.

The idea of the generalized measurement formalism is to forget the details of the whole interaction between the measured system (here the spin degree of freedom) and the measuring device (here the spatial degree of freedom), and express the resulting probability distribution in the form:

$$p(p) = \langle s|\Pi_p|s\rangle, \qquad (2.11)$$

where Π_p are the respective measurement operators, which need not in general be projective operators. Still they need to be positive (in order for the probaility distribution to be positive) and sum up to identity $\int \mathrm{d}p\, \Pi_p = \mathbb{1}$. In our case the corresponding operators are easy to find and read explicitly:

$$\Pi_p = \frac{1}{\sqrt{2\pi\sigma^2}} \begin{bmatrix} e^{-\frac{(p-\delta p)^2}{2\sigma_p^2}} & 0 \\ 0 & e^{-\frac{(p+\delta p)^2}{2\sigma_p^2}} \end{bmatrix} \qquad (2.12)$$

when written in the $\{|+\rangle, |-\rangle\}$ basis.

Note that in case $\delta p \ll \sigma_p$ the measurement provide us with almost no information, while in the opposite case the measurement results are highly informative regarding the spin information

Figure 2.2: Conceptual scheme of a generalized measurement

2.2 Mathematical formulation

Inspired by the Stern-Gerlach example, we are now ready to present the general formulation of the concept of generalized measurements.

Consider two quantum systems: the system to be measured (S) and the measuring device (M). The general idea of a generalized measurement, is to let the system interact with the measuring device, after which the measurement device state is read out using a standard projective measurement. Initially, the system and the measurement device are uncorrelated and their state is given by, $\rho_{SM} = \rho_S \otimes |0\rangle\langle 0|_M$, where we assumed (without loss of generality, as we may always purify the measuring device system to a larger space) that the M is prepared initially in a pure state. As a result of a unitary interaction of the two

[2] In a real Stern-Gerlach experiment, we let the particle evolve for some time t, and then measure the position when the particle hits the screen. If the time t is long enough this will be equivalent to the measurement of the momentum of the particle—see Problem 2.1

systems the final state reads:

$$\rho'_{SM} = U\rho_{SM}U^\dagger. \qquad (2.13)$$

Finally, a measurement, projecting the state of M onto a $\{|i\rangle_M\}$ basis ($P_i = |i\rangle\langle i|_M$) is performed resulting in the probability distribution:

$$p(i) = \mathrm{Tr}_{SM}\left(\rho'_{SM}\ \mathbb{1}_S \otimes P_i\right) = \\ = \mathrm{Tr}_{SM}\left(U\rho_S \otimes |0\rangle\langle 0|_M U^\dagger\ \mathbb{1} \otimes |i\rangle\langle i|\right) = \\ = \mathrm{Tr}_S(\rho_S \Pi_i), \quad (2.14)$$

where $\Pi_i = {}_M\langle 0|U^\dagger\ \mathbb{1} \otimes P_i\ U|0\rangle_M$. Π_i are generalized measurement operators, which in general need not be projective operators. Still from the above construction if follows that they satisfy positivity ($\Pi_i \geq 0$) and completeness ($\sum_i \Pi_i = \mathbb{1}$) conditions.

Any set of operators, $\{\Pi_i\}$ such that $\Pi_i \geq 0$, $\sum_i \Pi_i = \mathbb{1}$ is called a positive operator valued measure (POVM), since we may regard it as an operator which when traced with the density matrix generates a measure in the space of events i. We have seen above, that the interaction of a quantum system with a measuring system followed by the projective measurement of the latter results in an effective description of the measurement using a POVM.

The question, which is essential for the whole field of quantum estimation theory, is whether for any given POVM there is a physical realization, in the sense of a particular form of interaction between S and M and a particular measurement that results in the effective probability distribution described by this POVM. The answer is yes, and it is known under the name of Naimark dilation theorem. We provide the finite-dimensional case proof below.

Let $\{\Pi_i\}$, $i = 1, \ldots K$ be a POVM, $\Pi_i \in \mathcal{L}(\mathbb{C}_d)$. We will show, that there exist a unitary $U \in \mathcal{L}(\mathbb{C}_{d\cdot K})$ and a projective measurement P_i on $\mathbb{C}_K$ such that,

$$\forall_{\rho_S} \mathrm{Tr}(\rho_S \Pi_i) = \mathrm{Tr}(U\rho_S \otimes |0\rangle\langle 0|U^\dagger\ \mathbb{1} \otimes P_i). \qquad (2.15)$$

Let us define

$$U|\psi\rangle \otimes |0\rangle = \sum_{i=1}^{K} \sqrt{\Pi_i}|\psi\rangle \otimes |i\rangle. \qquad (2.16)$$

Note, that there is no problem in taking a square root from Π_i as it is a positive operator. Note also, that if this U is a legitimate unitary operation, then when accompanied by the projection $P_i = |i\rangle\langle i|$, realizes the required POVM. Consider the above operation U acting on two different input states. U preserves scalar products between the states, as

$$\left(\sum_i \langle\psi'| \otimes \langle i|\sqrt{\Pi_i}\right)\left(\sum_j \sqrt{\Pi_j}|\psi'\rangle \otimes |j\rangle\right) = \\ \langle\psi'|\sum_i \Pi_i|\psi\rangle = \langle\psi'|\psi\rangle. \quad (2.17)$$

The above property is not in itself a sufficient condition for unitarity (it is necessary), since we only defined the map on a subset of state in the whole Hilbert space $\mathbb{C}_{d\cdot K}$ (the subset of states of the form $|\psi\rangle \otimes |0\rangle$). In other words, we have shown, that if we write U as a matrix, then the d columns are legitimate columns take from a unitary matrix—they are orthonormal to each other. As such, we may always perform a completion of the matrix to a full unitary matrix by adding additional orhtogonal vectors until we get all the columns which constitute the the whole orthonormal basis. This ends the proof ■.

Is U in the above construction unique? No. We could as well take:

$$U|\psi\rangle \otimes |0\rangle = \sum_{i=1}^{K} V_i\sqrt{\Pi_i}|\psi\rangle \otimes |i\rangle, \qquad (2.18)$$

where V_i are unitaries. This unitaries, may be understood as rotations of the post-measurement state—they do not affect the probabilities of obtaining different measurement results.

2.3 Post-measurement state

If instead of just calculating the probabilites, we wanted to write down the post-measurement state itself, we need to go back to the representation of the generalized measurement as a subsequent interaction with the measuring device and a projective measurement on it. Given measurement result i the joined output state of the S and M subsystems reads:

$$\rho'^{(i)}_{SM} = \mathbb{1} \otimes P_i \; \rho'_{SM} \; \mathbb{1} \otimes P_i = \mathbb{1} \otimes |i\rangle\langle i| \; U\rho \otimes |0\rangle\langle 0|U^\dagger \; \mathbb{1} \otimes |i\rangle\langle i|. \quad (2.19)$$

Tracing out the M subsystem we get the conditional state

$$\rho'^{(i)}_{S} = \langle i|U|0\rangle \rho_S \langle 0|U^\dagger|i\rangle = K_i \rho K_i^\dagger, \quad (2.20)$$

where we have introduced the so called Kraus operators $K_i = \langle i|U|0\rangle$ (note that this is a partial scalar product, that leaves an operator acting on the S system). This state is subnormalized and its trace gives the probability of obtaining result i:

$$p(i) = \mathrm{Tr}(\rho'^{(i)}_S) = \mathrm{Tr}(K_i \rho_S K_i^\dagger) = = \mathrm{Tr}(\rho_S K_i^\dagger K_i) = \mathrm{Tr}(\rho_S \Pi_i) = p(i), \quad (2.21)$$

where we have used the property that $\Pi_i = K_i^\dagger K_i$. If we insist on writing a normalized conditional state, we should write $\rho'^{(i)}_S / p_i$.

Note, that given Π_i the corresponding K_i are determined only up to a unitary: $K_i = V_i \sqrt{\Pi}$, where V_i can be arbitrary unitary. This represents that fact, that after the measurement result is obtained, one may freely rotate the state depending on the measurement results, and this freedom does not appear in the statistics of the measurement results.

Recalling the example of the Stern-Gerlach experiment, one can see that in the limit $\delta p \ll \sigma_p$ the Π_p operators are practically proportional to identity. This implies that the while we get very little information about the spin, the spin state is also almost not disturbed at all. We refer to such a regime as the *weak measurement* regime. In the opposite case where we obtain a lot of information but at the same time disturb the state we say we deal with a *strong measurement*.

2.4 Decoherence and completely positive maps

Imagine now a situation in which, in the above described protocol, we forget to register the actual measurement result. We can regard this situation in a spirit, that we simply do not have access to the readout of the measurement performed on the subsystem M, or in other words that M should be treated as inaccessible environment with which our system S interacts. In such a situation, the output state of the system S is obtained by simply tracing out the joined state of S and M over subsystem M, and reads:

$$\rho'_S = \mathrm{Tr}_M \left(U\rho_S \otimes |0\rangle\langle 0|U^\dagger \right) = \sum_i K_i \rho_S K_i^\dagger. \quad (2.22)$$

The above formula has a clear intuitional meaning. This is a mixture of different conditional states corresponding to different measurement results i, representing the fact that we have no knowledge of the actual value of i and hence we are forced to consider the mixture only.

The above formula represents a general structure of a quantum channel. Kraus operators, K_i, can be arbitrary operators (not necessary unitary, hermitian,...), but in order to guarantee the trace-preservation condition they need to satisfy: $\sum_i K_i^\dagger K_i = \mathbb{1}$. Note that, the condition of preservation the positivity of the density matrix is automatically satisfied, as for any positve operator P, $KUK^\dagger$ is positive as well. Hence the above transformation is a positive linear map, i.e. it transforms positive operators into positive operators. In fact, it is a *completely positive* (CP) map, which means that even when the map is trivially extended to larger space it remains positive.

The evolution of a quantum state described by Eq. (2.22) is in general no longer unitary, and may in particular generate mixed states out of pure input states, and lead to the so called decoherence of quantum states. To see it, let us go back to the Stern-Gerlach example. Immediately, after the interaction the output state is given by Eq. (2.6). Let us calculate the corresponding reduced density matrix of the S (spin) system:

$$\rho_S(\delta t) = \mathrm{Tr}_M \left(|\psi(\delta t)\rangle\langle\psi(\delta t)|\right) = \left[\begin{array}{cc} |c_+|^2 & c_+^* c_- \langle\phi_-|\phi_+\rangle \\ c_-^* c_+ \langle\phi_+|\phi_-\rangle & |c_-|^2 \end{array} \right], \quad (2.23)$$

where M now corresponds to the spatial degree of freedom, and the reduced density matrix is written in the $\{|+\rangle, |-\rangle\}$ basis. Comparing the above formula, with the density matrix of the input spin state:

$$|s\rangle\langle s| = \left[\begin{array}{cc} |c_+|^2 & c_+^* c_- \\ c_-^* c_+ & |c_-|^2 \end{array} \right] \quad (2.24)$$

we see that while the diagonal elements remain unchanged, the off-diagonal elements are being suppressed the more the more orthogonal (distinguishable) are states $|\phi_\pm\rangle$— the process which we call decoherence. In other words, the more the environment (in this case spatial degrees of freedom) get the information on the spin state of the system the stronger the resulting decoherence process. Note, that decoherence process distinguishes a preferred basis, of the so-called pointer states (in this case $|\pm\rangle$ states), which are not affected by the decoherence process, but superposition of these states are affected, and in the extreme case are transformed into mixtures of pointer states.