Chapter 1
The Statistical Operator

We introduce the statistical operator formalism. After defining the statistical operator and its main properties, we reformulate Quantum Mechanics in this new formalism. Next, we present some of the possible representations of the statistical operator in the phase space. Examples follow along the way.

1.1 Statistical Operator and Density Matrix

In describing quantum mechanical systems, often for simplicity one assumes that the state of the system is perfectly known. However, as for classical systems, this is not true in reality. Usually, we have only a partial knowledge of the system, and the best we can know is represented by

$$\{\psi_k, p_k\} \quad \text{with} \quad \sum_k p_k = 1, \tag{1.1}$$

where ψ_k are the possible states of the system, and p_k the probabilities for each of them to be the actual state of the system. This is called a statistical mixture. Note that this mixture reflects a classic ignorance about the system, that is not related to quantum indeterminism.

Clearly each state ψ_k evolves according to the Schrödinger equation and the usual rules of Quantum Mechanics (QM) apply to each of them. In particular, for a system in a state described by the mixture in Eq. (1.1) the probability for an outcome o_n associated to the eigenstate $|o_n\rangle$ of an observable $\hat{O}$ is

$$\mathbb{P}[o_n] = \sum_k p_k |\langle o_n|\psi_k\rangle|^2, \tag{1.2}$$

which shows a mixing of classical ignorance (determined by p_k) and quantum indeterminism (given by $|\langle o_n|\psi_k\rangle|^2$). Proceeding this way is possible although not ideal, since one has to consider each state separately. A more convenient way is provided by the density matrix or statistical operator formalism.

The statistical operator is defined as follows:

$$\hat{\rho} = \sum_k p_k |\psi_k\rangle \langle\psi_k| . \tag{1.3}$$

To understand the physical meaning of $\hat{\rho}$, let $\{|a_i\rangle\}_i$ be a basis of the Hilbert space $\mathbb{H}$, of dimension n, associated to the quantum system. Then, we can represent the statistical operator as a matrix:

$$\hat{\rho} \;\rightarrow\; \rho_{ij} = \begin{pmatrix} \langle a_1|\hat{\rho}|a_1\rangle & \langle a_1|\hat{\rho}|a_2\rangle & \dots & \langle a_1|\hat{\rho}|a_n\rangle \\ \langle a_2|\hat{\rho}|a_1\rangle & \langle a_2|\hat{\rho}|a_2\rangle & \dots & \langle a_2|\hat{\rho}|a_n\rangle \\ \vdots & \vdots & & \vdots \\ \langle a_n|\hat{\rho}|a_1\rangle & \langle a_n|\hat{\rho}|a_2\rangle & \dots & \langle a_n|\hat{\rho}|a_n\rangle \end{pmatrix}, \tag{1.4}$$

which is called the density matrix. In practice, the density matrix and the statistical operator are often used as synonymous, as we will do. However, as a matter of principle, the density matrix is the representation of the statistical operator with respect to a specific basis.

Example 1.1
Let us consider a system with two degrees of freedom. The associated Hilbert space $\mathbb{H}$ is two-dimensional; let $\{|0\rangle, |1\rangle\}$ be a basis of $\mathbb{H}$. Let us consider the state

$$|+\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right).$$

The corresponding statistical operator is

$$\hat{\rho} = |+\rangle\langle+| = \frac{1}{2}\left(|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|\right),$$

while the density matrix with respect to the chosen basis is

$$\rho_{ij} = \begin{pmatrix} \langle 0|\hat{\rho}|0\rangle & \langle 0|\hat{\rho}|1\rangle \\ \langle 1|\hat{\rho}|0\rangle & \langle 1|\hat{\rho}|1\rangle \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

The different elements of the density matrix have specific physical meanings. The diagonal elements are

$$\langle a_k|\hat{\rho}|a_k\rangle = \langle a_k|\psi\rangle\langle\psi|a_k\rangle = |\langle a_k|\psi\rangle|^2, \tag{1.5}$$

which represent the probabilities for the system to be found in state $|a_k\rangle$ upon a measurement of the observable $\hat{A} = \sum_k a_k \, |a_k\rangle \, \langle a_k|$. The off-diagonal elements instead are

$$\langle a_k|\hat{\rho}|a_j\rangle = \langle a_k|\psi\rangle\langle\psi|a_j\rangle, \tag{1.6}$$

which are different from 0 only if the state $|\psi\rangle$ has components both with respect to the basis states $|a_k\rangle$ and $|a_j\rangle$. As such, the off-diagonal elements measure the presence of quantum coherence among the different states of a chosen basis. It is important to remember that the information provided by the density matrix is always relative to the chosen basis. This is made clear by the following example.

Example 1.2
Instead of representing the statistical operator in Example 1.1 with respect to the basis $\{|0\rangle, |1\rangle\}$, we now represent it with respect to the basis $\{|+\rangle, |-\rangle\}$, where

$$|+\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right), \quad |-\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right).$$

Then, the corresponding density matrix reads

$$\tilde{\rho}_{ij} = \begin{pmatrix} \langle +|\hat{\rho}|+\rangle & \langle +|\hat{\rho}|-\rangle \\ \langle -|\hat{\rho}|+\rangle & \langle -|\hat{\rho}|-\rangle \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

The meaning of the matrix elements is that with certainty the system is in state $|+\rangle$ (the corresponding diagonal element is equal to 1), and there is no quantum coherence between the two basis states $|+\rangle$ and $|-\rangle$ (the off-diagonal elements are null).

1.2 The physical meaning of the density matrix elements

From the previous discussion, it is clear that the diagonal elements of the density matrix have a direct physical interpretation as probabilities of outcomes of suitable measurements, see Eq. (1.5). On the other hand, the off-diagonal elements in general are complex numbers and as such they cannot be directly associated to an observable. The following two examples provide a better understanding of the physical meaning of the off-diagonal elements of the density matrix.

Example 1.3
Consider a mixture in which the state of the system is in

$$|0\rangle \text{ with probability } 1/2, \quad \text{and} \quad |1\rangle \text{ with probability } 1/2.$$

The corresponding statistical operator is

$$\hat{\rho} = \tfrac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right),$$

and the associated density matrix with respect to the basis $\{|0\rangle, |1\rangle\}$ is:

$$\rho_{ij} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The diagonal elements indicate that the probability of finding the system in state $|0\rangle$ or $|1\rangle$ is $1/2$, consistently with the mixture associated to the density matrix. The fact that the off diagonal elements are 0 indicates that no quantum coherence among the two basis states can be observed, since the system is never in a quantum superposition of $|0\rangle$ and $|1\rangle$.

Example 1.4
Consider the following statistical mixture, which is physically different from that of Example 1.3:

$$|+\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \text{ with probability } 1/2,$$
$$|-\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right) \text{ with probability } 1/2.$$

Now, a straightforward calculation shows that the statistical operator is the same as that in Example 1.3:

$$\hat{\rho} = \frac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right),$$

and therefore also the associated density matrix with respect to the basis $\{|0\rangle, |1\rangle\}$ is the same

$$\rho_{ij} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Accordingly, the physical interpretation of the diagonal and off-diagonal elements is also the same: no quantum interference is detected in spite of the fact that each state contributing to the statistical mixture is in a superposition.

The physical situations described in the two Examples above are quite different. In Example 1.3, the system is not in a superposition of the basis states $|0\rangle$ and $|1\rangle$, while it is in Example 1.4. Nevertheless, the density matrix is the same in both cases. These two Examples teach us something important: the statistical operator formalism is a many-to-one map between the set of statistical mixtures and the set of statistical operators: different statistical mixtures can be associated to the same statistical operator.

1.3 Propriesties of the Statistical Operator

The statistical operator as defined in Eq. (1.3), i.e. $\hat{\rho} = \sum_k p_k |\psi_k\rangle \langle\psi_k|$, shares three important properties.

1. It is a linear operator; given two vectors $|\psi\rangle, |\phi\rangle \in \mathbb{H}$ and two coefficients $a, b \in \mathbb{C}$, one has:

$$\begin{aligned} \hat{\rho}\left(a|\psi\rangle + b|\phi\rangle\right) &= \sum_k p_k |\psi_k\rangle \langle\psi_k| \left(a|\psi\rangle + b|\phi\rangle\right), \\ &= \left(a \sum_k p_k |\psi_k\rangle \langle\psi_k|\psi\rangle + b \sum_k p_k |\psi_k\rangle \langle\psi_k|\phi\rangle\right), \\ &= a\hat{\rho}|\psi\rangle + b\hat{\rho}|\phi\rangle . \end{aligned} \tag{1.7}$$

2. It is a positive operator:

$$\langle\psi|\hat{\rho}|\psi\rangle = \sum_k p_k \langle\psi|\psi_k\rangle \langle\psi_k|\psi\rangle = \sum_k p_k |\langle\psi|\psi_k\rangle|^2, \tag{1.8}$$

   which is non-negative for every $|\psi\rangle$ since $p_k \geq 0$.
3. The trace is *1*; given a basis $\{|\phi_n\rangle\}$ of the Hilbert space, we have:

$$\begin{aligned} \mathrm{Tr}\left[\hat{\rho}\right] &= \sum_n \langle\phi_n| \left(\sum_k p_k |\psi_k\rangle \langle\psi_k|\right) |\phi_n\rangle, \\ &= \sum_k p_k \left(\sum_n |\langle\phi_n|\psi_k\rangle|^2\right), \\ &= \sum_k p_k ||\psi_k||^2 = \sum_k p_k = 1, \end{aligned} \tag{1.9}$$

   where in the last line we used the fact that the vectors $|\psi_k\rangle$ defining the statistical mixture associated to the statistical operator are normalized, and that the probabilities p_k sum to 1.

These three conditions turn out to be not only necessary but also sufficient conditions to characterize a statistical operator. As a matter of fact, let us consider a linear and positive operator $\hat{\rho}$ over finite dimensional Hilbert space $\mathbb{H}$. Then, $\hat{\rho}$ is self-adjoint (this follows from its positivity), and therefore it admits a spectral decomposition in the form

$$\hat{\rho} = \sum_k \lambda_k |\psi_k\rangle \langle\psi_k| , \tag{1.10}$$

where $\{\lambda_k\}$ are non-negative eigenvalues and $\{|\psi_k\rangle\}$ are orthonormal eigenstates of $\hat{\rho}$. The trace condition implies that $\sum_k \lambda_k = 1$, meaning that the coefficient $\{\lambda_k\}$ can be interpreted as probabilities. This allows to call $\hat{\rho}$ a statistical operator.

Note that, while the decomposition of the statistical operator as an ensemble of state in general in not unique (as discussed in the previous section), the spectral decomposition instead is unique. Moreover, for a generic statistical mixture the states need not to be orthogonal, while for the mixture associated to the spectral decomposition they are.

1.4 Pure states and statistical mixtures

There are two important classes of statistical operators, which are called respectively pure states and statistical mixtures.

- Pure states: A statistical operator represents a pure state when it corresponds to a unique vector in the Hilbert space. In such a case, it can be represented as

$$\hat{\rho} = |\psi\rangle \langle\psi| . \tag{1.11}$$

  Here, one has maximal knowledge about the states of the system, compatibly with the rules of quantum mechanics (the indeterminacy principle still applies).

- Statistical mixture: When the state is not pure, we have a statistical mixture reading

$$\hat{\rho} = \sum_k p_k |\psi_k\rangle \langle\psi_k| , \tag{1.12}$$

  where, of course, more than one probability p_k is different from zero.

Given a density matrix, it is not always straightforward to understand if it corresponds to a pure state or to a statistical mixture. A first criterion to discriminate between the two is based on the purity of the state. For a pure state one has that

$$\hat{\rho}^2 = \hat{\rho}. \tag{1.13}$$

> ***Exercise 1.1***
> *Prove that the relation in Eq. (1.13) holds for pure states.*

This result can be summarized in the following theorem.

Theorem 1.1. *A statistical operator $\hat{\rho}$ is pure if and only if $\hat{\rho}^2 = \hat{\rho}$, otherwise it is a statistical mixture.*

When the dimension of the Hilbert space is large, it is not computationally easy to compare $\hat{\rho}$ with $\hat{\rho}^2$, and a simpler criterion is desirable. The following theorem is easy to prove.

Theorem 1.2. *Let $\hat{\rho}$ be a statistical operator. Then, in general one has that*

$$\mathrm{Tr}\left[\hat{\rho}^2\right] \leq 1, \tag{1.14}$$

where the equality holds if and only if $\hat{\rho}$ is a pure state.

> ***Exercise 1.2***
> *Prove that the relation in Eq. (1.14).*

Computing the trace of a matrix is computationally easier than comparing two matrices, as the first operation scales with the dimension of the Hilbert space and the second with its square. Therefore, Theorem 1.2 is that commonly used to verify the purity of a state.

1.5 The Bloch Sphere

Two dimensional quantum systems are the easiest example of quantum systems, and they are typically used to set the ground for studies of more complex systems. Moreover, they are particularly relevant since they can encode the qubit, the unit of quantum information. Let us then consider a two dimensional Hilbert space $\mathbb{H}$ and its computational basis $\{ |0\rangle, |1\rangle \}$, which is commonly known as the computational basis. A density matrix $\hat{\rho}$ on $\mathbb{H}$ when represented on a basis becomes a 2×2 matrix of the form

$$\rho = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{with } a, b, c, d \in \mathbb{C}. \tag{1.15}$$

In principle, these four coefficients represent 8 degrees of freedom. However, one can show that the density matrix can be written in terms of three real numbers as follows

$$\begin{aligned} \rho &= \frac{1}{2} \begin{pmatrix} 1 + r_z & r_x - i r_y \\ r_x + i r_y & 1 - r_z \end{pmatrix}, \\ &= \frac{1}{2} \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + r_x \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + r_y \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + r_z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right], \\ &= \frac{\mathbb{1} + \mathbf{r} \cdot \boldsymbol{\sigma}}{2}, \end{aligned} \tag{1.16}$$

where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices and $\mathbf{r} = (x, y, z)$ three real coefficients, where $|\mathbf{r}| \leq 1$. We see that the density matrix is fully controlled by the three-dimensional vector $\mathbf{r}$, which is called the Bloch vector.

> ***Exercise 1.3***
> *Prove that, by imposing i) $\hat{\rho}^\dagger = \hat{\rho}$, ii) $\mathrm{Tr}\left[\hat{\rho}\right] = 1$, iii) $\hat{\rho} \geq 0$, one can describe a two dimensional density matrix as in Eq. (1.16).*

Let us consider the square of $\hat{\rho}$

$$\begin{aligned} \hat{\rho}^2 &= \left(\frac{\hat{\mathbb{1}} + \mathbf{r} \cdot \hat{\boldsymbol{\sigma}}}{2} \right)^2, \\ &= \frac{1}{4} \left(\hat{\mathbb{1}} + 2\mathbf{r} \cdot \hat{\boldsymbol{\sigma}} + \sum_{ij} r_i r_j \hat{\sigma}_i \hat{\sigma}_j \right), \\ &= \frac{1}{4} \left(\hat{\mathbb{1}} + 2\mathbf{r} \cdot \hat{\boldsymbol{\sigma}} + ||\mathbf{r}||^2 + i \sum_{ij} r_i r_j \epsilon_{ijk} \hat{\sigma}_k \right), \end{aligned} \tag{1.17}$$

where we used the relation $\hat{\sigma}_i \hat{\sigma}_j = \delta_{ij} \hat{\mathbb{1}} + i \sum_k \epsilon_{ijk} \hat{\sigma}_k$ with ϵ_{ijk} indicating the Levi-Civita symbol (described by having $\epsilon_{123} = 1$ and being odd for any permutation of two indexes). It follows that

$$\mathrm{Tr}\left[\hat{\rho}^2\right] = \frac{1}{2} \left(1 + ||\mathbf{r}||^2 \right). \tag{1.18}$$

Given the results of the previous section, the condition $\mathrm{Tr}\left[\hat{\rho}^2\right] \leq 1$ implies that $||\mathbf{r}||^2 \leq 1$, where the equality holds only for pure states. This leads to a rather natural way of representing the Bloch vector as a point of a sphere of radius 1, which is called the Bloch sphere. Each point inside the sphere or on its surface represents a density matrix. If the point lies on the surface, it represents a pure state. If it is inside, it represents a statistical mixture.

Example 1.5
The maximally mixed state is $\hat{\rho} = \hat{\mathbb{1}}/2$. *In this case* $\mathbf{r} = 0$, *which corresponds to the center of the Bloch sphere.*

A (pure) state vector can be written as $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$, where its normalization implies that $|a|^2 + |b|^2 = 1$. Given such a constraint on the norm, it also admits a Bloch representation

$$|\psi\rangle = \cos\frac{\theta}{2}\,|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}\,|1\rangle\,, \tag{1.19}$$

which is unique, up to an unimportant global phase. The associated density matrix reads

$$\rho = \begin{pmatrix} \cos^2\frac{\theta}{2} & e^{-i\varphi}\cos\frac{\theta}{2}\sin\frac{\theta}{2} \\ e^{-i\varphi}\cos\frac{\theta}{2}\sin\frac{\theta}{2} & \sin^2\frac{\theta}{2} \end{pmatrix}, \tag{1.20}$$

whose Bloch vector is

$$\mathbf{r} = (\sin\theta\,\cos\varphi, \sin\theta\,\sin\varphi, \cos\theta)\,, \tag{1.21}$$

and represents, in spherical coordinates, a point on the surface of the Bloch sphere (indeed, here we have $||\mathbf{r}|| = 1$). We see that the Bloch sphere representation of a pure state is the same whether ones considers the representation of the state vector or that of the corresponding density matrix, as it should be. The computational basis vectors $|0\rangle$ and $|1\rangle$, which are also eigenstates of $\hat{\sigma}_z$, correspond to the intersection between the Bloch sphere and the z axis. Similarly, the eigenstates of $\hat{\sigma}_x$ correspond to the intersection with the x axis, and those of $\hat{\sigma}_y$ with the axis y, see Fig. 1.1.

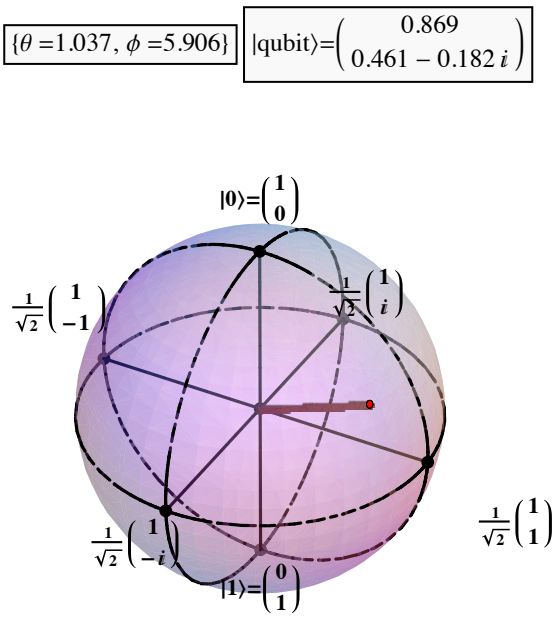

Fig. 1.1: Representation of the state $|\text{qubit}\rangle$ in the Bloch sphere (red arrow). The eigenstates of $\hat{\sigma}_i$, $i = x, y, z$, are also explicitly represented.

Exercise 1.4
Prove that the Bloch vector $\mathbf{r}$ *appearing in Eq. (1.16) can be obtained from* $\mathbf{r} = \mathrm{Tr}\left[\hat{\boldsymbol{\sigma}}\hat{\rho}\right]$.

1.6 Quantum Mechanics in the Density Matrix formalism

It is quite straightforward to rewrite the axioms of QM in density matrix formalism. The rules are the following.

States. To every physical system, an Hilbert space $\mathbb{H}$ is associated. The state of the system is represented by a statistical operator $\hat{\rho}$. We remind that they admit the following decomposition

$$\hat{\rho} = \sum_k p_k \left|\psi_k\right\rangle \left\langle\psi_k\right|, \quad \text{with} \quad \sum_k p_k = 1, \tag{1.22}$$

where the states $\left|\psi_k\right\rangle$ are normalized.

Evolution. Since each state $\left|\psi_k\right\rangle$ evolves according to the Schrödinger equation, it follows that the evolution of $\hat{\rho}$ is given by

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}_t = \sum_k p_k \left[\left(i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left|\psi_k\right\rangle\right)\left\langle\psi_k\right| + \left|\psi_k\right\rangle\left(i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left\langle\psi_k\right|\right)\right], \tag{1.23}$$

which amounts to

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}_t = \left[\hat{H}, \hat{\rho}_t\right], \tag{1.24}$$

to be solved with the initial condition $\hat{\rho}_0 = \hat{\rho}$. This is known as von-Neumann-Liouville equation. It should be clear that its dynamics content is equivalent to that of the Schrödinger equation.

The formal solution of the von-Neumann-Liouville equation is provided by

$$\hat{\rho}_t = \hat{U}_t\hat{\rho}_0\hat{U}_t^\dagger, \quad \text{with} \quad \hat{U}_t = \exp\left[-i\hat{H}t/\hbar\right], \tag{1.25}$$

where, for simplicity, we are considering the case of a time independent Hamiltonian. Due to the cyclic property of the trace ($\mathrm{Tr}\left[\hat{A}\hat{B}\hat{C}\right] = \mathrm{Tr}\left[\hat{C}\hat{A}\hat{B}\right]$), and since $\hat{U}_t$ is unitary, one has

$$\mathrm{Tr}\left[\hat{\rho}_t\right] = \mathrm{Tr}\left[\hat{U}_t\hat{\rho}_0\hat{U}_t^\dagger\right] = \mathrm{Tr}\left[\hat{U}_t^\dagger\hat{U}_t\rho_0\right] = \mathrm{Tr}\left[\hat{\rho}_0\right] = 1, \tag{1.26}$$

indicating that trace is preserved, as it should be. Physically, it means that the probabilities are conserved. This is a property inherited from the Schrödinger equation. Also the trace of the square of the density matrix is conserved:

$$\mathrm{Tr}\left[\hat{\rho}_t^2\right] = \mathrm{Tr}\left[\hat{U}_t\hat{\rho}_0\hat{U}_t^\dagger\hat{U}_t\hat{\rho}_0\hat{U}_t^\dagger\right] = \mathrm{Tr}\left[\hat{U}_t\hat{\rho}_0\hat{\rho}_0\hat{U}_t^\dagger\right] = \mathrm{Tr}\left[\hat{\rho}_0^2\right]. \tag{1.27}$$

This means that pure states are mapped into pure states. Again, this is not a surprise, since a state vector remains a state vector under the Schrödinger evolution. Similarly, statistical mixtures are mapped into statistical mixtures: our ignorance about the state of the system propagates (linearly) during the evolution.

Observables. Observable quantities are represented by self-adjoint operators on $\mathbb{H}$. This axiom does not change with respect to the usual one.

Measurements. Since we are working within the quantum formalism, the possible outcomes of a (ideal) measurement of an observable A correspond to the eigenvalues a_n of the corresponding self-adjoint operator $\hat{A}$, whose spectrum is assumed to be discrete for simplicity. The outcomes are randomly distributed according to the Born

rule. In the language of the density matrix, these statements translate as follows. Given the state

$$\hat{\rho} = \sum_k p_k \left|\psi_k\right\rangle \left\langle\psi_k\right| \tag{1.28}$$

one has that the probability of having a_n as an outcome of the measurement is given by

$$\mathbb{P}[a_n] = \sum_k p_k |\left\langle a_n|\psi_k\right\rangle|^2 = \sum_k p_k \left\langle a_n|\psi_k\right\rangle \left\langle\psi_k|a_n\right\rangle = \left\langle a_n|\hat{\rho}|a_n\right\rangle, \tag{1.29}$$

where the first equality encodes the Born rule ($|\left\langle a_n|\psi_k\right\rangle|^2$) with an average over our ignorance about the state of the system (the sum over k with weights p_k). Let us consider the projection operator $\hat{\mathcal{P}}_n = \left|a_n\right\rangle \left\langle a_n\right|$ associated to the eigenvalue a_n. It is quite simple to see that an equivalent way of writing Eq. (1.29) is

$$\mathbb{P}[a_n] = \mathrm{Tr}\left[\hat{\mathcal{P}}_n \hat{\rho}\right]. \tag{1.30}$$

In a similar way, one can show that the expectation value of an observable is

$$\langle \hat{A} \rangle = \mathrm{Tr}\left[\hat{A}\hat{\rho}\right]. \tag{1.31}$$

State collapse. The density matrix allows to describe two different types of measurements, with associated state collapse. The first type is called selective measurement, and corresponds to that usually described in textbooks. Assuming that the outcome of the measurements of the observable $\hat{A}$ is a_n, then the state collapses to the corresponding eigenstate, whatever the initial state was. In the density matrix formalism, this corresponds to:

$$\hat{\rho}_{\mathrm{before}} \quad \Rightarrow \quad \hat{\rho}_{\mathrm{after}} = \left|a_n\right\rangle \left\langle a_n\right|, \tag{1.32}$$

which can be rewritten as

$$\hat{\rho}_{\mathrm{before}} \quad \Rightarrow \quad \hat{\rho}_{\mathrm{after}} = \frac{\hat{\mathcal{P}}_n \hat{\rho} \hat{\mathcal{P}}_n}{\mathrm{Tr}\left[\hat{\mathcal{P}}_n \hat{\rho}\right]}. \tag{1.33}$$

Note that the effect of the collapse is nonlinear, and it cannot be deduced from the Schrödinger equation, which is linear. Notably, an initially mixed state becomes pure, indeed one has that $(\hat{\rho}_{\mathrm{after}})^2 = \hat{\rho}_{\mathrm{after}}$. This property of the collapse is well known, and it is important as it gives the means to prepare a system in a given state.

The other possibility is a non-selective measurement, where all outcomes are retained and distributed according to the Born rule. Correspondingly, one has

$$\hat{\rho}_{\mathrm{before}} \quad \Rightarrow \quad \hat{\rho}_{\mathrm{after}} = \sum_n p_n \frac{\hat{\mathcal{P}}_n \hat{\rho} \hat{\mathcal{P}}_n}{\mathrm{Tr}\left[\hat{\mathcal{P}}_n \hat{\rho}\right]}, \tag{1.34}$$

with $p_n = \mathbb{P}[a_n] = \mathrm{Tr}\left[\hat{\mathcal{P}}_n \hat{\rho}\right]$. Note that, conversely to the previous case, this operation is linear since the above equation can be trivially expressed as

$$\hat{\rho}_{\mathrm{before}} \quad \Rightarrow \quad \hat{\rho}_{\mathrm{after}} = \sum_n \hat{\mathcal{P}}_n \hat{\rho} \hat{\mathcal{P}}_n. \tag{1.35}$$

Finally, we note that a non-selective measurement can turn an initially pure state into a statistical mixture. Indeed, one has that

$$\begin{aligned}\mathrm{Tr}\left[\left(\sum_n \hat{\mathcal{P}}_n\hat{\rho}\hat{\mathcal{P}}_n\right)^2\right] &= \mathrm{Tr}\left[\sum_n \hat{\mathcal{P}}_n\hat{\rho}\hat{\mathcal{P}}_n\hat{\rho}\right] = \sum_n \mathrm{Tr}\left[\hat{\mathcal{P}}_n\hat{\rho}\hat{\mathcal{P}}_n\hat{\rho}\right], \\ &\leq \sum_n \left(\mathrm{Tr}\left[\hat{\mathcal{P}}_n\hat{\rho}\right]\right)^2 \leq \left(\sum_n \mathrm{Tr}\left[\hat{\mathcal{P}}_n\hat{\rho}\right]\right)^2 = 1.\end{aligned} \tag{1.36}$$

To summarise, selective measurements are nonlinear operations which generate pure states, while non-selective measurements are linear operations which generate statistical mixtures.

Example 1.6
Consider a two dimensional system whose Hamiltonian is $\hat{H} = \hat{\sigma}_z$, where we set $\hbar = 1$ and in the computational basis $\hat{\sigma}_z$ is represented by

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In the Bloch representation, the von Neumann-Liouville equation $\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho} = -i\left[\hat{H}, \hat{\rho}\right]$ reads

$$\frac{\dot{\mathbf{r}}\cdot\hat{\boldsymbol{\sigma}}}{2} = -\frac{i}{2}\left[\hat{\sigma}_z, \hat{\mathbb{1}}\right] - \frac{i}{2}\sum_k r_k\left[\hat{\sigma}_z, \hat{\sigma}_k\right].$$

Given that $\left[\hat{\sigma}_z, \hat{\mathbb{1}}\right] = 0$ and $[\hat{\sigma}_z, \hat{\sigma}_k] = 2i\sum_j \epsilon_{zkj}\hat{\sigma}_j$, and that the set of matrices $\left\{\hat{\mathbb{1}}, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\right\}$ is a basis of space of 2×2 matrices, we obtain three equations for the coefficients of the Bloch vector

$$\dot{r}_x = -2r_y, \quad \dot{r}_y = 2r_x, \quad \dot{r}_z = 0,$$

whose solution is

$$r_x = \cos 2t, \quad r_y = \sin 2t, \quad r_z = const.$$

The Hamiltonian $\hat{\sigma}_z$ makes the Bloch vector rotate around the z axis of the Bloch sphere, both for pure states and for statistical mixtures. Similarly, $\hat{\sigma}_x$ makes the Bloch vector rotate around the x axis and $\hat{\sigma}_y$ around the y axis.