1 The Language of Quantum Information Theory

1.1 Density Operators

1.1.1 Postulates of Quantum Mechanics

We begin by stating the a common axiomatization of Quantum Mechanics (QM) based on Hilbert Spaces following [2], we choose it for mathematical simplicity; for alternatives see e.g. [6].

1. To each physical system $\mathcal{S}$ there corresponds a separable Hilbert Space $\mathcal{H}$ such that states of the system are described by positive and unit trace operators on it. The Hilbert Space of a composite system made up of $\mathcal{S}$ and $\mathcal{S}'$ is given by the tensor product of the Hilbert spaces $\mathcal{H} \otimes \mathcal{H}'$.

2. To each dynamical variable there corresponds a self-adjoint operator on $\mathcal{H}$, called an observable, whose possible values are given by its eigenvalues.

3. Given a system in state ρ and some observable A of it, the probability of measuring A and obtaining the result λ is given by $\mathrm{Tr}[\rho P_\lambda]$ where P_λ is the eigen-projector into the subspace associated with λ. Furthermore the expectation value is $\mathrm{Tr}[\rho A]$.

4. After a measurement with result λ the state of the system becomes $\frac{P_\lambda \rho P_\lambda}{\mathrm{Tr}[P_\lambda \rho P_\lambda]}$.

5. The time evolution of the system in a time interval $(0, t)$ in which no measurement is done is given by some unitary operator U_t according to $\rho_t = U_t \rho U_t^\dagger$ where ρ is the state of the system at time $t = 0$.

Operators satisfying the properties required for a state are called ***Density Operators*** and in contrast to frameworks whose treatment of quantum states is merely as rays in $\mathcal{H}$, they describe statistical mixtures so imperfect state preparation can be handled. To see this consider the spectral resolution of some density operator:

$$\rho = \sum_n p_n \, |\psi_n\rangle\langle\psi_n| \tag{1.1}$$

by definition we have $p_n \geq 0, \sum_n p_n = 1$ any density operator can be seen as a convex sum of rays in $\mathcal{H}$ (provided we identify each one with its associated projector $|\psi\rangle\langle\psi|$) and from it an interpretation of ρ as an statistical mixture of rays is suggested: given a preparation process, there is a probability p_n for the system to be in the state $|\psi_n\rangle\langle\psi_n|$ after it, for this reason states of the form $|\psi\rangle\langle\psi|$ are called ***Pure*** while those who are not we refer to as ***Mixed***. This ***Ensemble Interpretation*** has serious conceptual challenges when one tries to use it outside a fixed preparation procedure due to the non-uniqueness of the decomposition into pure states [5], but is good enough for the porposes of the present work, for a comprehensive discussion of this topic the reader is refered to [7].

1.1.2 Time Evolution

Assuming the evolution to be differentiable in time, we have that there exists a self-adjoint operator H such that $U_t = \exp(-itH)$ [1], called the ***Hamiltonian*** of the systems and which acts as the generator of the dynamics. It is straightforward now to construct a differential equation for ρ_t by taking the derivative of it:

$$\rho_t = e^{-itH} \rho_t e^{itH} \tag{1.2}$$

$$\partial_t \rho_t = -iH\rho_t + \rho_t iH \tag{1.3}$$

$$\partial_t \rho_t = -i[H, \rho_t]. \tag{1.4}$$

Equation (1.4) is called the Liouville-Von Neumann equation, it generalizes the time-dependent Schrödinger equation to mixed states and can be interpreted as the quantum analog of the Liouville equation in classical mechanics (with the Poisson braket) through the quantization rule $\{\bullet, \bullet\} \to -i[\bullet, \bullet]$. As will be seen in later chapters, this type of evolution is characteristic of closed quantum systems.

1.1.3 Purity

Say we got a particular state production processes whose product ρ we characterize via say tomography [5], it becomes immedeatley important to quantify to which extent we can regard the product as being composed of only one pure states (hopefully the one we wanted to prepare) i.e. we want to define the purity of the state, with this motivation one look for a map $\mathcal{E}$ from the space of density operators to the reals such that:

- $\mathcal{E}(\rho)$ is maximal if and only if ρ is pure.
- it is conserved under unitary evolution.

[1] Unless otherwise stated, from here on we assume $\hbar = 1$

The first one makes this map a figure of merit one can try to maximize and the second one is imposed to assure that it doesn't changes in a closed system unless a measurement is made, as allowing the free evolution of the system should not improve the knowledge of the experimenter about the system. The standard choiche (altough not the only one) is the ***Purity***, defined as [5, 2]:

Definition 1 *The purity γ of a state ρ is:*

$$\gamma = \text{Tr}\left[\rho^2\right]. \tag{1.5}$$

The requirements are quickly checked:

$$\text{Tr}\left[\rho_t^2\right] = \text{Tr}\left[(U_t\rho_0 U_t^\dagger)(U_t\rho_0 U_t^\dagger)\right] = \text{Tr}\left[\rho_0^2\right] \tag{1.6}$$

$$\text{Tr}\left[\rho^2\right] = \sum_n p_n^2 \leq 1 \tag{1.7}$$

in the second line the inequality is saturated if and only if $\rho = |\psi\rangle\langle\psi|$.

1.2 Entanglement

One of the key differences between the structure of the state space of classical and quantum systems is the existance of non-separable states when considering multipartite systems [6, 3, 5] which allows the latter to have new a new type of correlations. Here we define entanglement for mixed states following [3]:

Definition 2 *Given a state ρ in a system composed of two subsystems A and B with total Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B\}$, we say it is an **entangled** or **non-separable** if there doesn't exists a set states $\{\rho_j \otimes \sigma_j\}_j$ and coefficients $\{p_j\}_j, \sum_j p_j = 1,\ p_j \geq 0$ such that:*

$$\rho = \sum_j p_j \rho_j \otimes \sigma_j \tag{1.8}$$

*if it does exists, the state is called **separable**.*

For the case of pure state this definition coincides with the usually given one [5]: say $\rho = |\psi\rangle\langle\psi|$ is pure and separable, then:

$$\text{Tr}\left[\rho^2\right] = \sum_{jk} p_j p_k \text{Tr}\left[\rho_j\rho_k\right] \text{Tr}\left[\sigma_j\sigma_k\right] \tag{1.9}$$

and by the Cauchy-Schwartz inequality with the Frobenious inner product

$$\text{Tr}\left[\rho^2\right] \leq \sum_{jk} p_j p_k \text{Tr}\left[\rho_j^2\right] \text{Tr}\left[\rho_k^2\right] \text{Tr}\left[\sigma_j^2\right] \text{Tr}\left[\sigma_k^2\right] \leq 1 \tag{1.10}$$

the first inequality from right to left saturates if and only if all the ρ_j and σ_k are pure, and the first one if and only if all the ρ_j and σ_k are equal between themselves i.e. $p_j = \delta_{0j}$, hence there are pure states in $\mathcal{H}_A$ and $\mathcal{H}_B$ such that:

$$|\psi\rangle\langle\psi| = |\alpha\rangle\langle\alpha| \otimes |\beta\rangle\langle\beta| \,. \tag{1.11}$$

For a classical system all states are separable thanks to the representation via δ functions of probability densities [3] and in this sense entanglement is a purely non-classical phenomena, in fact for pure states this completely exhaust all the possible non-classical correlations. For mixed states the characterization is considerably richer and allows for bipartite states that despite being separable show non-classical correlations [1].

1.2.1 Marginalization

Consider a bipartite system in state ρ with subsystems A and B, and assume only the former can be accessed experimentally; e.g. B is on the other side of the galaxy, has too many degrees of freedom or is simply not of interest and is desirable to prescend from it. Any observable Γ that we decide to measure must be of the form $\Gamma = \Lambda \otimes I$ so that it describes only actions on $\mathcal{H}_A$; in this sense we say it is ***local***. We want to obtain the probability distribution describing the statistics of Λ, by definition:

$$\wp(\lambda) = \mathrm{Tr}\left[\rho(|\lambda^A\rangle\langle\lambda^A| \otimes I)\right] \tag{1.12}$$

$$\wp(\lambda) = \sum_{\lambda' k} \langle\psi_k^B| \langle\lambda'^A| \rho(|\lambda\rangle\langle\lambda| \otimes I) |\lambda'^A\rangle |\psi_k^B\rangle \tag{1.13}$$

$$\wp(\lambda) = \sum_{\lambda' k} \langle\psi_k^B| \langle\lambda'^A| \rho |\lambda'^A\rangle |\psi_k^B\rangle \delta_{\lambda\lambda'} \tag{1.14}$$

$$\wp(\lambda) = \sum_{k} \langle\psi_k^B| \langle\lambda^A| \rho |\lambda^A\rangle |\psi_k^B\rangle \tag{1.15}$$

$$\wp(\lambda) = \langle\lambda^A| \left(\sum_k \langle\psi_k^B| \rho |\psi_k^B\rangle\right) |\lambda^A\rangle \tag{1.16}$$

$$\wp(\lambda) = \mathrm{Tr}\left[|\lambda^A\rangle\langle\lambda^A| \sum_k \langle\psi_k^B| \rho |\psi_k^B\rangle\right] . \tag{1.17}$$

Equation (1.17) suggest that there exists a state in $\mathcal{H}_A$ whose statistics coincide with those of ρ and that it should be given by the sum in (1.17), in this sense we have a marginalization i.e. an assignement of states of in $\mathcal{H}_A \otimes \mathcal{H}_B$ to states in $\mathcal{H}_A$ such that it has the correct statistics i.e. $\mathrm{Tr}\left[\rho^{AB}(\Lambda \otimes I)\right] = \mathrm{Tr}\left[\rho^A \Lambda\right]$ for any observable Λ of A. It is desirable for this map to be unique for the following: assume an experimentalist has an infinite ensemble of copies of the system in state ρ but can only measure local observables in A, although any and as many time as wanted, i.e. it is possible to fully characterize the statistics of any local observable, which state should the experimentalist asign? If the marginalization is not unique there is an ambiguity, an marginalizations sure must exists as local experiments are always possible, hence uniqueness is important to account properly for this experiment; turns out to assure it suffices to demand linearity. This map is called the ***Partial Trace***:

Definition 3 *Given two vector spaces V and W, for simplicity assumed of finite dimension*[2]*, the partial trace taken over W is the map[5]:*

$$\mathrm{Tr}_W: \ A\otimes B \in \mathcal{L}(V\otimes W) \ \mapsto \ A\mathrm{Tr}\left[B\right] \in \mathcal{L}(V)$$

where the $\mathcal{L}$ denotes the space of operator, and the map is linearly extended to all of $\mathcal{L}(V\otimes W)$.

The linear extension makes the partial trace coincide with the sum in (1.17) and the non-manifestly base invariant definition usually given in sources like [4]. Next we prove this is in fact the only linear map with the correct statistics:

Theorem 1 *The partial trace is the only linear map such that $\mathcal{E} :\rightarrow \mathcal{L}(\mathcal{H}_A\otimes\mathcal{H}_B) \rightarrow \mathcal{L}(\mathcal{H}_A)$ such that $\mathrm{Tr}\left[\rho(\Lambda\otimes I)\right] = \mathrm{Tr}\left[\mathcal{E}(\rho)\Lambda\right]$ for all $\rho \in \mathcal{H}_A\otimes\mathcal{H}_B$ and $\Lambda\in\mathcal{L}(\mathcal{H}_A)$.*

Proof 1 *Assume there exists a map $\mathcal{E} :\rightarrow \mathcal{L}(\mathcal{H}_A\otimes\mathcal{H}_B) \rightarrow \mathcal{L}(\mathcal{H}_A)$ with the correct statistics and introduce an orthogonal product basis in $\mathcal{H}_A\otimes\mathcal{H}_B$, $\{|\psi_j^A\rangle\,|\psi_k^B\rangle\}_{jk}$. By construction we have:*

$$\mathrm{Tr}\left[\mathcal{E}(\rho)\Lambda\right] = \sum_{jk} \langle\psi_j^B|\,\langle\psi_k^A|\,\rho(\Lambda\otimes I)\,|\psi_k^A\rangle\,|\psi_j^B\rangle \tag{1.18}$$

$$\mathrm{Tr}\left[\mathcal{E}(\rho)\Lambda\right] = \sum_{k} \langle\psi_k^A| \sum_j \langle\psi_j^B|\,\rho\,|\psi_j^B\rangle\,(\Lambda\,|\psi_k^A\rangle) \tag{1.19}$$

$$\mathrm{Tr}\left[\mathcal{E}(\rho)\Lambda\right] = \sum_{k} \langle\psi_k^A|\,\mathrm{Tr}_B\left[\rho\right](\Lambda\,|\psi_k^A\rangle) \tag{1.20}$$

$$\mathrm{Tr}\left[\mathcal{E}(\rho)\Lambda\right] = \mathrm{Tr}\left[\mathrm{Tr}_B\left[\rho\right]\Lambda\right] \tag{1.21}$$

$$\tag{1.22}$$

As this holds for any Λ and ρ, we have that $\mathcal{E} = \mathrm{Tr}_B$.

Note that in the above proof we have not used any properties of ρ or Λ unlike in (1.17). A few remarks are in order:

1. The partial traces of an state $\rho\in\mathcal{L}(\mathcal{H}_A\otimes\mathcal{H}_B)$ are always mixed, unless ρ is both pure and separable. In this sense we say the subsystems of a system in an entangled state can not be perfectly known, not even if the state of the complete system is perfectly known [5, 1, 3].

[2] although we will use it too for infinite dimensions without inquiring wheter the operators are even traceclass

2. When ρ is entangled and pure $\mathrm{Tr}_B[\rho]$ is an improper mixture [7], in the sense that for none of the possible ensembles that represent this density operator is possible to say that A is in an unknown state $|\psi\rangle \in \mathcal{H}_A$ with some probability; if this were the case we could perform experiments in both A and B to discover this unknown states and then the total state would have not been entangled in the first place.

3. For identical particles this construction is not valid as the space of physical states is not $\mathcal{H}_A \otimes \mathcal{H}_B$ but its symmetrization, otherwise the symmetrization postulate is violated. For proposals of suitable generalizations see [6].

Despite this conceptual difficulties the partial trace remains a key tool in Quantum Information Theory for studying open systems and local operations. Next we explore one of its most promiment applications.

1.2.2 The Von Neumann Entropy

As can be seen in the study of Bell's inequalities in a two qubit system, not all entangled states show the same amount of violation of the inequalities, and it is not even true that non-separability is sufficient to violate them [2], thus not all entangled states are made equal and classifying them is a relevant task; to achieve this we introduce ***Entanglement Measures***, which quantify how strong this non-classical correlations is in a bipartite state. Following [8] we demand from any proper measure:

- it is a map $\mathcal{E} : \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathbb{R}$.
- $\mathcal{E}(\rho) = 0$ if and only if ρ is separable.
- Local unitary operations preserve $\mathcal{E}(\rho) = \mathcal{E}(U_A \otimes U_B \rho U_A^\dagger \otimes U_B^\dagger)$

Bibliography

[1] Gerardo Adesso, Marco Cianciaruso, and Thomas R Bromley. "An introduction to quantum discord and non-classical correlations beyond entanglement". In: *arXiv preprint arXiv:1611.01959* (2016).

[2] Leslie E. Ballentine. *Quantum Mechanics: A Modern Development (2nd Edition)*. World Scientific Publishing Company, Aug. 28, 2014. 739 pp.

[3] Lajos Diosi. *A Short Course in Quantum Information Theory: An Approach From Theoretical Physics*. Springer, June 2, 2011. 167 pp.

[4] Jonas Maziero. "Computing partial traces and reduced density matrices". In: *International Journal of Modern Physics C* 28.01 (2017), p. 1750005.

[5] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. 2010. 702 pp.

[6] Andrés F. Reyes-Lega. "Some aspects of operator algebras in quantum physics". In: *Geometric, Algebraic and Topological Methods for Quantum Field Theory*. WORLD SCIENTIFIC, Sept. 20, 2015, pp. 1–74.

[7] Maximilian A. Schlosshauer. *Decoherence: and the Quantum-To-Classical Transition*. Springer Science & Business Media, July 28, 2007. 428 pp.

[8] Vlatko Vedral et al. "Quantifying entanglement". In: *Physical Review Letters* 78.12 (1997), p. 2275.