Part II

Evolution Part

Suppose a system prepared in a state ρ undergoes a process. The original preparation procedure along with the process can be considered a new preparation procedure. The equivalence class of the new preparation procedure will define the new quantum state after the process. This state depends on the original preparation procedure and the process. Each process corresponds to a linear map,

$$\alpha : \mathcal{B} \to \mathcal{A}.$$

The algebras of observables $\mathcal{A}$ and $\mathcal{B}$ represent the input and output systems respectively. To an initial state ρ of $\mathcal{A}$ the channel associates the output state $\rho \circ \alpha$ of $\mathcal{B}$.

If density matrices are used to describe quantum states in quantum mechanics, then a process must be some operation that sends density matrices to density matrices. So for finite-dimensional state spaces, a process should be a linear map of vector spaces of matrices. It preserves the trace of matrices and takes hermitian matrices with non-negative eigenvalues to hermitian matrices with non-negative eigenvalues. It must take positive operators to positive operators. A map is called positive if it takes positive operators to positive operators. Suppose the process acts only on some part of the system then it must still be a process on the total system. The map corresponding to a process should be positive for the bigger system as well. Such maps are called completely positive. A general quantum process corresponds to a completely positive unital mapping.

1 | Quantum Measurement

In quantum theory, the description of the system requires two physical objects. The first being the state of the system which contains the information known about the system. Second, the observables, which are objects the information is about. Bayes' theorem says that additional information about a system will alter the probabilities of possible outcomes. The notion of information is closely related to the notion of probability. Probability gives one way to describe information about the events. We are interested in quantifying the amount of information contained in a state relative to another state.

The relative entropy of two states ρ and σ is the informational divergence of ρ from σ. Suppose the state σ contains information only about a subsystem $\mathcal{B}$ of $\mathcal{A}$ and E is a projection of norm one of $\mathcal{A}$ onto $\mathcal{B}$ then the state σ should satisfy $\sigma \circ E = \sigma$. In such a case the informational divergence should have two components. First component is the divergence of ρ from σ on the subalgebra $\mathcal{B}$ which is the divergence between the states $\rho|_{\mathcal{B}}$ and $\sigma|_{\mathcal{B}}$. The other component is the remaining information ρ has and this will be the divergence between

the states ρ and $\rho \circ E$. If $R(\cdot\,,\cdot)$ is such a function then,

$$R(\rho, \sigma) = R(\rho|_{\mathcal{B}}, \sigma|_{\mathcal{B}}) + R(\rho, \rho \circ E).$$

Any automorphism α of the algebra $\mathcal{A}$ should change the information contained in the two states similarly hence the information divergence should be invariant under automorphisms of the algebra,

$$R(\rho, \sigma) = R(\rho \circ \alpha, \sigma \circ \alpha).$$

The informational divergence of a state with respect to itself should be zero $R(\rho, \rho) = 0$. If $R(\cdot\,,\cdot)$ is a real-valued functional satisfying the above conditions then there exists a constant $c \in \mathbb{R}$ such that,

$$R(\rho, \sigma) = c\,\mathrm{Tr}\left(\rho\left(\log\rho - \log\sigma\right)\right).$$

The relative entropy of the state ρ with respect to σ is defined as,

$$J(\rho, \sigma) = \mathrm{Tr}\left(\rho\left(\log\rho - \log\sigma\right)\right).$$

In the classical case, the Bayes' rule has been shown to be a special case of the constrained maximization of relative entropy [**?**]. The quantum version of this result is obtained in [**?**]. We will state the result here.

Suppose an observable A has been subjected to measurement. For simplicity we consider the observable to be a discrete observable. Let A be a discrete observable with effects given by the set $\{A_i\}_{i\in I}$ and the corresponding projection operators $\{E_{A_i}\}_{i\in I}$. If the quantum state of the system after the measurement is σ, it carries information that has to be compatible with the possibility of measuring all eigenvalues of A precisely. Such a situation is given by the condition $[\sigma, A] = 0$. Suppose the result of the measurement is A_k then the probability of measuring A_k again should be $Tr(E_{A_k}\sigma) = 1$. Repeated measurements add no new information. The set of all such states such that $Tr(E_{A_k}\sigma) = 1$ is a convex set. Let $p = \{p_i\}_{i\in I}$ such that $\sum_i p_i = 1$. The set,

$$\mathcal{S}_p = \{\sigma \in \mathcal{S}(\mathcal{H}) \mid [E_{A_i}, \sigma] = 0,\ Tr(\sigma E_{A_i}) = p_i\},$$

encodes the data that the measurement outcome A_i corresponding to the projection E_{A_i} occurs with probability p_i. The commutation condition says that they posses a common eigenbasis and also means that $[\sigma, A] = 0$.

Theorem 1.1. (Hellmann-Kamiński-Kostecki)

$$\arg\inf_{\sigma\in\mathcal{S}_p}\{J(\rho, \sigma)\} = \sum_i p_i E_{A_i}\rho E_{A_i}/Tr(E_{A_i}\rho E_{A_i}). \quad \square$$

The strong collapse or the Lüders-von Neumann rule of collapse is a limiting case of the above projection with all p_i going to zero except one. By taking the limit $p_i \to 0$ for $i \neq j$ we get the Lüders-von Neumann's rule of collapse,

$$\rho \to E_{A_j}\rho E_{A_j}/Tr(E_{A_j}\rho E_{A_j}).$$

This amounts to selecting the quantum state that is least distinguishable from the original state among all the states that satisfy the constraint.

Idea of Proof

Given a convex subset $\mathcal{V}$ of a finite dimensional topological vector space and $f : \mathcal{V} \to \mathbb{R}$ is a convex function then σ is a global minimum of the function f on $\mathcal{V}$ if and only if all directional derivatives of f at σ are non negative.

In our case, $D(\cdot,\cdot) = -J(\cdot,\cdot)$ is a jointly convex function. $D(\rho,\cdot) = -J(\rho,\cdot)$ is a convex function on the state space. Now the problem is a minimization of a convex function. $\mathcal{V} = \mathcal{S}_p \subset \mathcal{S}(\mathcal{H})$. Every element of $\mathcal{S}_p$ can be written as follows,

$$\sigma = U\Lambda U^*,$$

where Λ is a diagonal matrix with positive entries and trace 1 and U is a unitary. Since $[\sigma, P_i] = 0$ for every $\sigma \in \mathcal{S}_p$. Now the idea is to parametrise this and optimise it.

□

For proof and generalization of the result to the algebraic case, the interested reader should read the original papers [?],[?],[?] and the references therein. In general measurement channels are given by positive operator-valued measures, where for a measure space, $(\Omega, \Sigma(\Omega))$, and $\epsilon \in \Sigma(\Omega)$, $E(\epsilon)$ is a positive operator, $E(\Omega) = 1$ and for pairwise disjoint ϵ_i, $\sum_i E(\epsilon_i) = E(\vee_i \epsilon_i)$. It should, however, be noted that the Lüders-von Neumann rule is about calibrating with the experimental result and has no predictivity. We will abuse the notation and denote an event characterized by the effect E_{A_i} by E_{A_i} only.

Postulate. (Lüders-von Neumann Collapse) *If an observable A, with values A_i with corresponding projections E_{A_i}, is measured on the system in a state ρ, then the state transforms to,*

$$E_{A_i}\rho E_{A_i}/Tr(E_{A_i}\rho E_{A_i}),$$

on the condition that the result A_i was obtained.

The advantage of this approach to arriving at the Lüders-von Neumann rule is that the starting point is information theoretic and can be formulated in case of GPTs with suitable available structure.

1.1 | Unitary Evolution

Here we give a brief review of the unitary evolution. The purpose of this subsection is to remind ourselves why unitary evolution is used in quantum theory. When it comes to time evolution, the quantum theory continues on with the received view. The symmetries of classical theories are implemented on objects of quantum theory.

The simplest structure a symmetric map should preserve is the convexity of the space of states, physically corresponding to the fact that a state arises from mixing states with certain statistical weights. Symmetry operations may modify the constituent states but do not change the weights. A bijection $\alpha : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ is a symmetry if it preserves the convex structure of $\mathcal{S}(\mathcal{H})$. For $p_i \in [0,1]$ and $\sum_i p_i = 1$,

$$\alpha(\textstyle\sum_i p_i\rho_i) = \sum_i p_i\alpha(\rho_i).$$

Such a map is called a Kadison automorphism.

Theorem 1.2. (Kadison-Wigner) *If a map α is a Kadison automorphism, then Kadison-Wigner theorem says α is of the form,*

$$\alpha(\rho) = U\rho\, U^{-1},$$

where U is unitary or antiunitary and is determined up to phase.

A unitary operator is a map U such that $\langle Ux, Uy\rangle = \langle x, y\rangle$ and an antiunitary operator is a map U such that $\langle Ux, Uy\rangle = \overline{\langle x, y\rangle}$ where $\langle\cdot,\cdot\rangle$ is the inner product on the Hilbert space. To implement the symmetries of the system, the symmetries must be represented in terms of Kadison automorphisms. We seek maps from some group to the set of Kadison automorphisms. Whether a specific transformation is unitary or antiunitary depends on its physical nature. Transformations that belong to a continuous group, such as translations and rotations, can only be unitary because in that case any finite transformation can be generated by a sequence of infinitesimal steps. Let $\mathcal{G}$ be the group of symmetries of the system, then to each $g \in \mathcal{G}$ there should correspond a Kadison automorphism,

$$\alpha :\ g \mapsto \alpha_g.$$

We get a unitary or antiunitary representative $U(g)$ to each element $g \in \mathcal{G}$. For now we will assume $U(g)$ to be unitary. Given $g, h \in \mathcal{G}$ we know that,

$$\alpha_g \alpha_h = \alpha_{gh}.$$

For compatible representative U we have,

$$U(g)U(h) = \lambda(g,h)U(gh),$$

where $\lambda(g,h)$ is a phase factor. A map $U : g \mapsto U(g)$ satisfying the above relation is called a projective unitary representation. $\lambda(g,h)$s are called multipliers. For $g = e$ we get,

$$U(e) = \lambda(e,e)I.$$

We get some conditions on the multipliers $\lambda(g,h)$. Applying several times to f, g, h we get,

$$\lambda(f,g)\lambda(fg,h) = \lambda(g,h)\lambda(f,gh).$$

We also get,

$$\lambda(e,g) = \lambda(g,e) = \lambda(e,e).$$

A projective unitary representation with $\lambda(e,g) = \lambda(g,e) = \lambda(e,e) = 1$ for every $g \in \mathcal{G}$ is said to be normalized. A map $g \mapsto U(g)$ is called a unitary representation of $\mathcal{G}$ on $\mathcal{H}$ if $U(e) = I$ and satisfies,

$$U(g)U(h) = U(gh).$$

Unitary representations are usually much easier to work with. A theorem of Bargmann says for some groups with nicer properties (connected and simply connected) it's possible to get a unitary representation. One can always consider the universal covering group and get a unitary representation of that anyway.

Given a self-adjoint operator A, one can construct a family of unitary operators, $U(t) = e^{-itA}$. Stone's theorem says the opposite is also true. If $t \mapsto U(t)$ is a strongly continuous one-parameter unitary group in the complex Hilbert space $\mathcal{H}$, there exists a unique self-adjoint operator A called the generator of the group such that,

$$U(t) = e^{-itA}.$$

We can, therefore, by Stone's theorem, associate with every one-parameter subgroup of $\mathcal{G}$ a unique self-adjoint operator A_i. The Lie algebra of the group $\mathcal{G}$ is represented by the self-adjoint operators A_i. From the Lie algebra of the group of symmetries, we can obtain the unitary representatives with a factor of $-i$. If the Lie algebra has the basic structure equation, $[a_i, a_j] = \sum c_{ij}^m a_m$, then the self-adjoint operators A_i corresponding to a_i satisfy the commutator relations, $i[A_i, A_j] = \sum c_{ij}^m A_m$.

We get the Schrödinger equation by implementing Galilean symmetries. When the symmetries are taken to be the Galilean group, the time evolution is generated by the Hamiltonian of the system and corresponds to the time translation symmetry of the system.

$$\rho \mapsto e^{-itH} \rho \, e^{itH}.$$

This is the Schrödinger equation.

Postulate. (Schrödinger) *Time translation is given by,* $\rho \mapsto e^{-itH} \rho e^{itH}$.

This notion of evolution is a direct copy-paste of the classical laws formulated for quantum objects.

References

[1] G Ludwig, Foundations of Quantum Mechanics I. Springer-Verlag, 1983

[2] G Ludwig, Foundations of Quantum Mechanics II. Springer-Verlag, 1985

[3] G Ludwig, An Axiomatic Basis for Quantum Mechanics Volume 1. Springer-Verlag, 1985

[4] K Kraus, States, Effects, and Operations. Lectures Notes in Physics, Springer-Verlag, 1983

[5] J M Jauch, Foundations of Quantum Mechanics. Addison-Wesley Publishing Company, 1968

[6] V Moretti, Fundamental Mathematical Structures of Quantum Theory. Springer, 2018

[7] A Peres, Quantum Theory: Concepts and Methods, Kluwer Academic Publishers, 2002