CHAPTER 1

Introduction

Contents

1.1 Motivations

Elementary quantum mechanics, to be briefly reviewed below, deals with simple quantum systems: an harmonic oscillator, an hydrogen atom,... A characteristic property of these systems, in fact an assumption, is that they are isolated[1]. The rest of the universe has no relevance for their time evolution, which by the postulates of the theory is given by an unitary operator $e^{-itH/\hbar}$ acting on the Hilbert space $\mathcal{H}_s$ of the system[2].

Yet measurement, which in general affects the system (this contrasts with classical mechanics), is a central instance when the outside world comes into play and cannot be neglected. But even without measurements, isolation is never truly exact. It is at best an excellent approximation, and only an abstraction: even in the "vacuum" of outer space, every system is surrounded by the relic photons at 3K. And an isolated hydrogen atom of quantum mechanics is always a idealization of a system described by quantum electrodynamics with its infinite number of degrees

[1] In fact this is also a basic hypothesis of elementary classical mechanics.

[2] In these notes we assume most a the time that a units have been chosen so that numerically $\hbar = 1$.

of freedom, so that the stable Bohr trajectories are not really stable.

One of the deep messages of open quantum systems (at least in the author's view, emphasized in these lectures), is the close relationship between measurement and interaction with an environment. This is often advocated for macroscopic objects, which are never found in quantum superpositions (something like a coin on a table in a quantum superposition of head and tail) if only because collisions with the surrounding air are more than enough to turn macroscopic quantum superpositions into classical objects in no time, with either a deterministic or a classical probabilistic description. But we shall see the relevance of this similarity at the level of first principles and apply it also to microscopic objects, though this will require some special types of measurements.

The theory of open quantum system is a, largely successful, attempt too describe non-isolated quantum systems which can exchange energy, particles,... with the rest of the Universe and may be subject to measurements. Of course if 1) we knew the Hamiltonian of the Universe and could deal with its complexity, and 2) if we could extract from the exact time evolution of the whole Universe the information that is relevant for the smaller systems we are interested in, open quantum systems would be an easy, or even useless, subject. But in our present understanding neither 1) nor 2) holds, and the need for a good description of an open quantum system, taking into account its interactions with the environment while involving only its internal degrees of freedom, is motivated by theoretical and practical considerations.

There is a whole range of situations, from those where neglecting the outside Universe is enough for a most accurate description to those where taking it into account requires to incorporate at least some of its degrees of freedom into the description.

Even when a good description of the system in terms if its internal degrees of freedom only is possible, the price to pay can be quite high, at least technically. For instance the state of the system at time $t + dt$ can be computed only from the knowledge of its states at all times $s \in [0, t]$ or $s \in]-\infty, t]$. A simple solvable example is the famous Caldeira-Legett model, when the system of interest, typically a particle in a potential, is coupled linearly to a large number of oscillators. The particle serves as an external source for the oscillators, so that their (inhomogeneous but linear) equations of motion, can be solved explicitly and the value of the dynamical variables of the oscillators at time t can be expressed via standard Greens function techniques in terms of initial conditions at $t = 0$ and the particle coordinates in the whole $[0, t]$. The result is an integro-differential equation for the dynamical variables of the particle.

But in these lectures, we shall concentrate on situations which in many respects are the easiest to deal with just after isolated systems: the so-called Markovian open quantum systems describing instances when the knowledge of the system at time t (plus possibly the effect of the environment between t and $t + dt$) is enough to determine the state at time $t + dt$. This is of course the case with the Schrödinger equation for isolated systems and the big difference is that Markovian open quantum systems may have a non-unitary and/or non-deterministic time evolution.

1.2 About the lectures

Though we already emphasized its high level of simplification when compared to many real situations, the subject of Markovian open quantum systems is vast and we shall only be able to touch some of its aspects, which are chosen more because of familiarity for the lecturer than for some broadly recognized central importance.

The basic notions are explained in a discrete time setting and for simple systems described by finite dimensional Hilbert spaces (most of the time of dimension 2 or 3 in concrete computations). This allows to keep the mathematics at a reasonable level of sophistication, basically within the realm of algebraic manipulations.

The choice of discrete time is made for pedagogical motivations, and we shall use it to "derive" the general continuous time equations and use continuous time to extract information.

This choice of small systems is not only pedagogical. The last decades have seen spectacular progress in the detailed manipulation of simple systems. This was made possible by parallel progresses in fact electronics and low temperature. Part of the motivation comes from quantum metrology, but the most fashionable excuse is the design of a quantum computer. Even if this goal is most likely still far ahead, daily experiments realize the dreams of Gedanken experiments devised by the fathers of quantum mechanics a century or so in the past. Our short journey into the world of Markovian open quantum systems will in particular give some interesting perspectives on wave function collapse, quantum jumps, and some other amusing creatures inhabiting the microscopic kingdom.

The following view engages only the author: thinking about open quantum system in general, and the approach of these notes in particular, is a good starting point to think about what quantum mechanics "really" is. There are many connections with the interpretation(s) of quantum mechanics. These notes do not propose one or even hint at one. But it seems that certain things which are obvious within one interpretation require some computing/thinking within others. Alas, this criterion does not always

select the same interpretation as the intuitive one.

1.3 Up the mountain and down

As with most scientific theories, two complementary aspects play a prominent role:
– To establish the general equations (or even the general formalism) of the theory
– To extract physical information from the above general formalism, i.e. to solve analytically or numerically for some interesting physical situations, to gain intuition, and to compare with experiments.

The first aspect can be seen as climbing-up a mountain to get a general view of the landscape, and the second to go down-hill and visit some of the accessible and pleasant locations. The derivation of the general equations takes some time and is a bit arid, but their contemplation is easier, so before embarking on the climbing-up, we take the opportunity to write them down.

The easiest case is when only averages or equivalently deterministic aspects are considered. The result is that a Markovian open quantum system is a deterministic dynamical system on density matrices. The evolution for one time step is

$$\rho \mapsto \rho' = \Psi(\rho) := \sum_{a \in S} V_a \rho V_a^\dagger, \tag{1.1}$$

where the V_as are arbitrary operators[3] on the Hilbert space of the system, subject only to the constraint

$$\sum_{a \in S} V_a^\dagger V_a = \mathrm{Id}. \tag{1.2}$$

This last equation ensures that the trace is preserved by the evolution. If there is only one V, the equation becomes $V^\dagger V = \mathrm{Id}$ which if the Hilbert space is finite dimensional ensures that V is unitary, and one is left with the usual situation of quantum mechanics. In the general case, the above equation contains a number of new phenomena like decoherence.

From a purely mathematical viewpoint, the situation is quite standard if not dull: the map in Equation 1.1 is a linear transformation on (the vector space spanned by) density matrices, so the study of the dynamical system is reduced to the study of the iteration of a linear transformation, i.e. to a problem of diagonalization. One can show that the linear transformation is contracting, i.e. that all eigenvalues are of modulus ≤ 1.

[3] A finite family in these lectures, i.e. S is a finite set.

A dual view is to go to the Heisenberg picture, i.e. to write down the evolution of observables

$$\Lambda \mapsto \Psi^\dagger(\Lambda) := \sum_{a\in S} V_a^\dagger \Lambda V_a, \tag{1.3}$$

and Equation 1.3 implies that the identity operator is conserved, showing that $\Psi^\dagger$ (and then Ψ in the finite-dimensional setting) has 1 as an eigenvalue.

Equation 1.1 is often called a (discrete) quantum master equation because it is a generalization to density matrices of standard master equations for the evolution of the probability distribution of Markov chains. Indeed, choosing a basis, the diagonal elements of a density matrix can be seen as a probability vector. It is an interesting exercise to find explicit solutions of Equation 1.3 for whom the diagonal of ρ has a closed time evolution under Equation 1.1, and to show that an arbitrary Markov chain dynamics appears in this way.

But Equation 1.3 has many more solutions leading to a much richer phenomenology. Even if the mathematics looks easy, one should keep in mind that finding the right V_as to describe the physics of a given system (even if amenable to a study via the Markovian open quantum systems approach) is a non-trivial task[4], and then the eigenvalues and eigenvectors of the dynamical map Ψ acquire a physical meaning and can be of great relevance for real phenomena.

When randomness due to contact with the rest of the universe is taken into account, the equations become more complicated. The central role is still played by operators V_a, $a \in S$ subject to Equation 1.3. But now we need to partition S in disjoint subsets as $S = \cup_{r\in R} S_r$ and the dynamical map to be iterated becomes

$$\rho \mapsto \rho' := \frac{\sum_{a\in S_r} V_a \rho V_a^\dagger}{\mathrm{Tr} \sum_{a\in S_r} V_a \rho V_a^\dagger} \text{ with probability } \pi_r := \mathrm{Tr} \sum_{a\in S_r} V_a \rho V_a^\dagger. \tag{1.4}$$

The trivial partition with only one subset takes us back to the deterministic case considered above. In all other case, one has to deal with a random dynamical system whose trajectories are called quantum trajectories. These quantum trajectories have a number of fascinating properties, and the reader is invited to look at simple cases for him/herself. She/he should

[4] A problem already pregnant for the use of Markov chains to model classical non-equilibrium dynamics by the way.

look at 2×2 density matrices, find a few solutions to with only two V_αs and implement Equation 1.4 on a computer.

We shall see that, as is obvious from the explicit formulæ with some experience, the above equations are related to measurements, which in particular explains their non-linearity. In this interpretation α labels the possible outcomes of an measurement experiment, and r, more precisely the fact that $\alpha \in S_r$, is the coarse-grained information that is taken into account to make future predictions. Keeping track of the outcome leads to supplement the evolution for ρ by an evolution for counting processes N_r, one for each possible coarse-grained outcome:

$$N_s \mapsto N'_s = N_s + \delta^r_s \text{ if the outcome is } r, \text{ probability } \pi_r := \mathrm{Tr} \sum_{\alpha \in S_r} V_\alpha \rho_s V_\alpha^\dagger. \tag{1.5}$$

The idea to iterate Equation 1.4 which is an equation for the evolution of the density matrix, whereas the only information we collect from the experiment is the counting process, raises interesting questions about the fundamental meaning of density matrices. In these notes, we mainly mention the statistical interpretation, but this viewpoint is hard to sustain when the number of iterations becomes large (why?) and it is tempting to revert to a more Bayesian viewpoint, where the density matrix is our best guess taken into account our knowledge but also our ignorance. In fact, as we shall show and put to good use later on, Equation 1.4 is in essence a Bayesian update equation, and this the way the results of real experiments are interpreted, see Pict.3.

The previous equations in fact describe classical dynamical systems, with classical noise. But they can be seen as a classical incarnation of truly quantum equations, driven by so-called quantum noises, about which we shall say a few words in the last chapters. A central role is played by the counting processes N_s: in the classical version the increments are either 0 or 1 (classical bits), while the quantum version keeps track of both possibilities and their interferences (qubits). The nuance should remind the reader of the basic idea of quantum computing.

Once we shall have established the above equations, we shall look at their continuous time limit, and use it to describe some simple but striking phenomena, mostly quantum jumps. This is only one of the facets of Markovian open quantum systems, but again the one the author is most familiar with.

1.4 Some paradigmatic systems

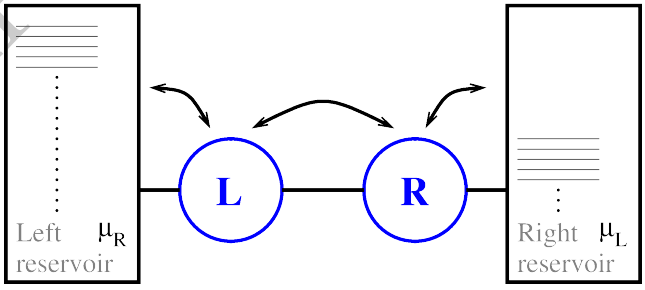
We now say a few words of the systems whose physics plays a central role in what follows.

Contact with reservoirs

– *Equilibrium and approach to equilibrium*: One of the most important example of an open quantum system is a system coupled to a heat bath. One can observe equilibrium properties, but also the approach to equilibrium. A heat bath is though of as a very large (even infinite) object, so that a small system coupled to a heat bath does not perturb it significantly and one can assume that the state of the heat bath is essentially time independent. If moreover all (relevant) relaxation times in the heat bath are much shorter than the time resolution at which the system is observed, the Markovian approximation is relevant.

Picture 1. *The most popular image of a thermal bath on the internet.*

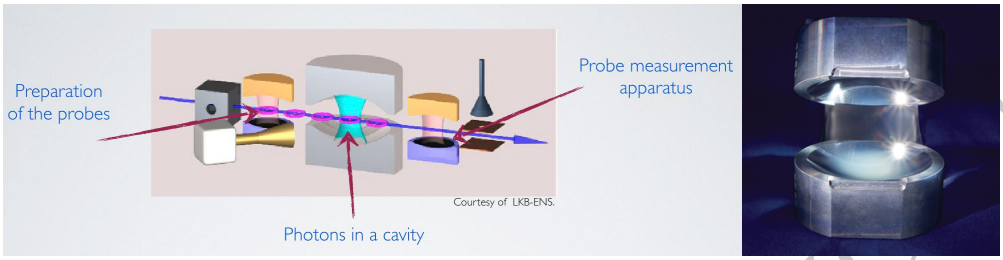
– *Non equilibrium*: This is the case of a system coupled to two heat baths at different temperatures, or to two particle reservoirs at different chemical potentials. The hypotheses of large size and small relaxation times of reservoirs can be used to motivate a Markovian approach.



Picture 2. *The Hamiltonian couples the dots* R *and* L *, resulting in Rabi oscillations. The reservoirs inject or extract electrons with different rates. Coulomb repulsion can be tuned to ensure that at most one electron is in the system at each time.*

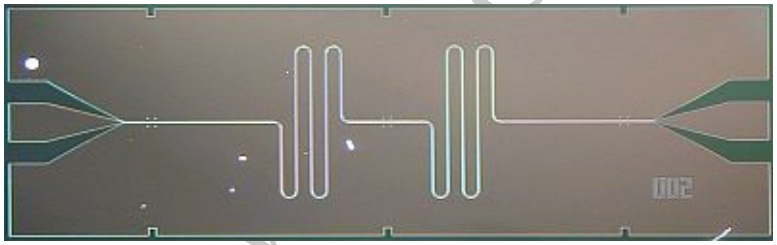
Measurements:

– *Cavity quantum electrodynamics*:

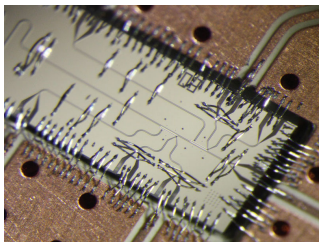


Picture 3. *Left:* $T \simeq 0$. *Rydberg atoms fly through the cavity, interact with the electromagnetic field, and later are measured. This leads to what is described as an indirect measurement of the mode of the cavity. Right: The cavity. This is one of the real experimental settings that we shall model in these notes.*

– *Circuit quantum electrodynamics*:

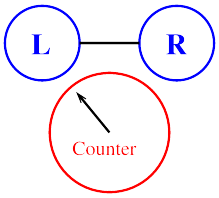


Picture 4. *A resonator in circuit QED.*



Picture 5. *A Toffoli, or "controlled-controlled-not gate", i.e one of the universal quantum gates, precursor of a quantum computer.*

– *Quantum dots*:



Picture 6. *The Hamiltonian couples the dots* R *and* L *, resulting in Rabi oscillations for an electron hopping back and forth between* L *and* R*. Measurements give information on the position of the electron in the system.*

Contact with reservoirs and measurement

– Observation of quantum jumps

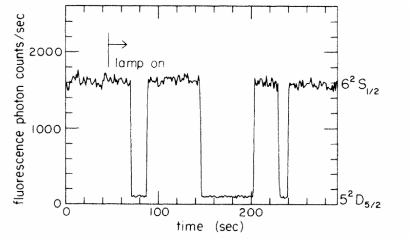
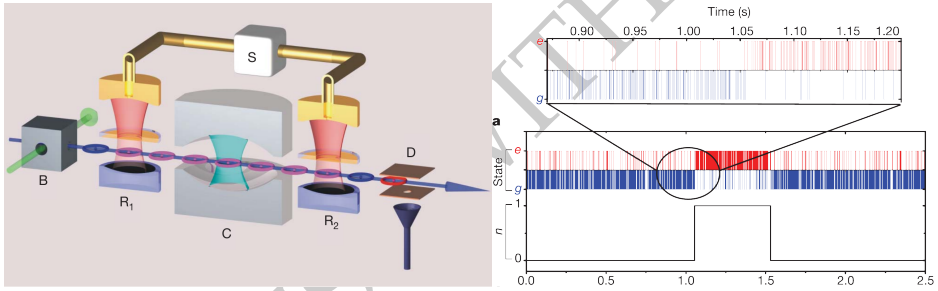
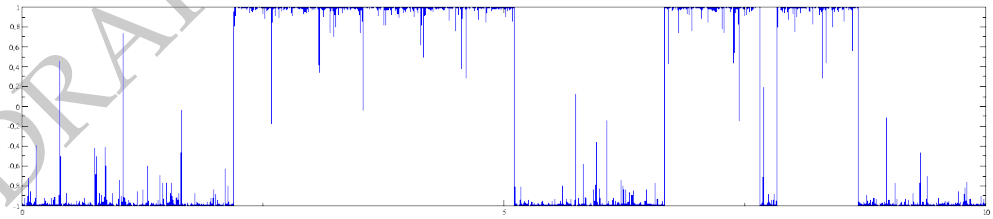


FIG. 2. A typical trace of the 493-nm fluorescence from the $6^2P_{1/2}$ level showing the quantum jumps after the hollow cathode lamp is turned on. The atom is definitely known to be in the shelf level during the low fluorescence periods.

Picture 7. *The very first observation of quantum jumps.*

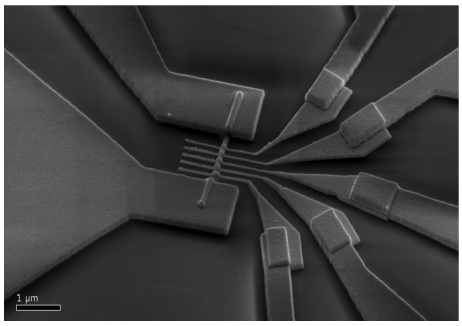


Picture 8. *The cavity at* T $\neq$ 0 *is subject to thermal fluctuations which are observed as quantum jumps. This is one of the real experiments that we shall study in detail in these notes.*

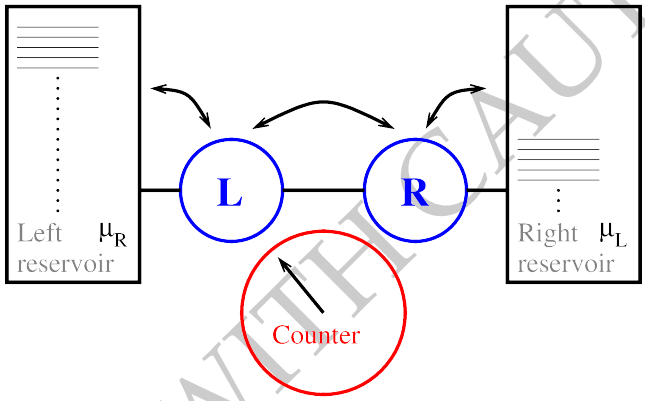


Picture 9. *The theoretical predictions for a typical quantum trajectory in the cavity experiment.*

– Quantum dots again



Picture 10. *A quantum dots device*



Picture 11. *The Hamiltonian couples the dots* R *and* L *, resulting in Rabi oscillations. The reservoirs inject or extract electrons with different rates. Coulomb repulsion can be tuned to ensure that at most one electron is in the system at each time. Indirect measurements give information on the presence and position of an electron in the system.*

CHAPTER 2

Quantum Mechanics in a Nutshell

Contents

2.1 Basic postulates

We summarize very briefly the basic postulates of quantum mechanics, emphasizing some points that will prove relevant in the rest of the lectures.

– Postulate 1:
To every quantum mechanical quantum mechanical system is associated a Hilbert space $\mathcal{H}$. It is important that this Hilbert space is complex, and it is usually assumed that it is separable, so that it has finite or countable orthonormal bases. A standard example is $\mathcal{H} = \mathbb{L}^2(\mathbb{R}^3)$ for a particle in $\mathbb{R}^3$, but many important physical cases involve finite dimensional Hilbert spaces. Finite dimensional Hilbert spaces suffice for most of the material covered in these notes.
– Postulate 1':
(*Configuration space for compound systems*) The Hilbert space $\mathcal{H}_C$ for a compound system $C = A \cup B$ with Hilbert spaces $\mathcal{H}_A$, $\mathcal{H}_B$ is the tensor product

$\mathcal{H}_A \otimes \mathcal{H}_B$ of the elementary subsystems[1]:

$$\mathcal{H}_C = \mathcal{H}_A \otimes \mathcal{H}_B.$$

For instance, for 2 particles in $\mathbb{R}^3$,

$$\mathcal{H} = \mathbb{L}^2(\mathbb{R}^3) \otimes \mathbb{L}^2(\mathbb{R}^3) \equiv \mathbb{L}^2(\mathbb{R}^6).$$

– Postulate 2:
(*First year definition*) The state of the system is described by a unit vector $|\,\varphi\,\rangle$ in $\mathcal{H}$, modulo a phase:

$$|\,\varphi\,\rangle \in \mathcal{H} \quad \||\,\varphi\,\rangle\| = 1 \quad |\,\varphi\,\rangle \sim e^{i\theta}|\,\varphi\,\rangle,\ \theta \in \mathbb{R}$$

(*More advanced definitions*) The state of a system is described by ray in $\mathcal{H}$, i.e. an element in the projective space $\mathbb{P}(\mathcal{H})$:

$$|\,\varphi\,\rangle \in \mathcal{H} \quad |\,\varphi\,\rangle \neq 0 \quad |\,\varphi\,\rangle \sim \lambda|\,\varphi\,\rangle,\ \lambda \in \mathbb{C}^*$$

The correspondence with the initial definition is simply $|\,\varphi\,\rangle \to \||\,\varphi\,\rangle\|^{-1}|\,\varphi\,\rangle$. Equivalently, the state of the system is described by a rank 1 projector in $\mathcal{B}(\mathcal{H})$ (the bounded linear operators on $\mathcal{H}$):

$$P \in \mathcal{B}(\mathcal{H}) \quad P^2 = P \quad \mathrm{Tr}_{\mathcal{H}}\, P = \dim \mathrm{range}(P) = 1 \quad \mathrm{range}(P) \perp \mathrm{kernel}(P)$$

The correspondence with rays is $|\,\varphi\,\rangle \to P^\varphi := \frac{|\,\varphi\,\rangle\langle\,\varphi\,|}{\langle\,\varphi\,|\,\varphi\,\rangle}$
(*A definitive(?) definition*) The previous definitions cover the notion of pure states. We shall need the notion of statistical mixtures in the sequel. The state of a system is described by a density matrix in $\mathcal{B}(\mathcal{H})$, i.e. a positive operator of trace[2] 1 :

$$\rho \in \mathcal{B}(\mathcal{H}) \quad \langle\,\psi\,|\,\rho\,|\,\psi\,\rangle \geq 0\ \forall\,|\,\psi\,\rangle \in \mathcal{H} \quad \mathrm{Tr}_{\mathcal{H}}\,\rho = 1$$

The density matrix is automatically self-adjoint with a discrete spectrum, i.e. has a spectral resolution

$$\rho = \sum_k p_k P_k \quad P_k P_l = \delta_{kl} P_k\ \forall\, k, l \quad \mathrm{range}(P_k) \perp \mathrm{range}(P_l)\ \forall\, k \neq l$$

[1]This principle is possibly to be supplemented by symmetry considerations/conditions when A and B are two copies of the same system. Grossly, for two copies of A, $\mathcal{H}_C = S^2(\mathcal{H}_A)$ (symmetrized tensor product) or $\mathcal{H}_C = A^2(\mathcal{H}_A)$ (antisymmetrized tensor product) depending whether A describes a boson or a fermion.

[2]A bit more on the trace in the next section.

$$p_k > 0 \ \forall k \quad \mathrm{Tr}_{\mathcal{H}} P_k = \dim \mathrm{range}(P_k) < +\infty \ \forall k \quad \sum_k p_k \mathrm{Tr}_{\mathcal{H}} P_k = 1$$

where the sums over k is discrete, i.e. k runs over a finite or countable set[3]. To each pure state corresponds a density matrix which is nothing but the projector associated to the pure state. Conversely, $\rho^2 = \sum_k p_k^2 P_k$ so $\mathrm{Tr}_{\mathcal{H}} \rho^2 = \sum_k p_k^2 \mathrm{Tr}_{\mathcal{H}} P_k \leq 1$ with strict inequality unless $p_k^2 = p_k$ for every k, i.e. $p_k = \delta_{k,k_0}$ for some k_0 and P_{k_0} must have rank 1. Thus the density matrix ρ^2 describes a pure state if and only if it is a projector. In other cases, ρ can be interpreted as a statistical mixture of pure states, but there is some arbitrariness that we shall emphasize later on.

– **Postulate 3**:

The time evolution of a quantum mechanical system is governed by a family of unitary operators. Recall that $U \in \mathcal{B}(\mathcal{H})$ is unitary of U maps $\mathcal{H}$ onto $\mathcal{H}$ and preserves the norm, i.e. $\| |\varphi\rangle \| = \|U|\varphi\rangle\|$ for every ket $|\varphi\rangle$. Equivalently, U is unitary if $UU^\dagger = \mathrm{Id}_{\mathcal{H}}$. If $|\varphi\rangle_t$ or ρ_t is a description of the system at time t, then

$$|\varphi\rangle_t = U(t,t')|\varphi\rangle_{t'} \text{ or } \rho_t = U(t,t')\rho_{t'}U(t,t')^\dagger.$$

Consistency requires

$$U(t,t')U(t',t'') = U(t,t'') \text{ or at least } U(t,t')U(t',t'') = U(t,t'')z(t,t',t'').$$

where $z(t,t',t'')$ is a phase satisfying co-cycle conditions and related to the fact that only rays matter. Under regularity hypotheses, there is an operator $H(t)$ on $\mathcal{H}$, the Hamiltonian, such that

$$i\hbar \frac{\partial}{\partial t} U(t,t') = H(t)U(t,t').$$

In the simple but frequent case when $H(t) = H$ is time-independent

$$U(t,t') = e^{-i(t-t')H/\hbar}.$$

The physics of a system is encoded in its associated Hilbert space $\mathcal{H}$ and its Hamiltonian $H(t)$ (often simply H). Up to the replacement of t by $-it$, there is a fruitful, but only partial, formal analogy with Markov processes, where $U(t,t')$ is analogous to a transition kernel and $H(t)$ is an infinitesimal generator.

[3] A canonical choice is to take integers for the index set and order $p_1 > p_2 > \cdots$. Another convenient, but in general ambiguous, choice is to assume that each P_k has rank 1 and put the multiplicities by allowing repetitions in the list of p_ks.

– Postulate 4:
Observables are what one can measure in experiments. An observable of a quantum system is a (any?) linear self-adjoint operator Λ on $\mathcal{H}$:

$$\Lambda \in \mathcal{L}(\mathcal{H}) \quad \Lambda = \Lambda^\dagger.$$

This is not really a definition, because general self-adjoint operators are unbounded (they do not belong to $\mathcal{B}(\mathcal{H})$) and can only be defined on a dense subspace of $\mathcal{H}$. Not necessarily bounded but densely defined operators on $\mathcal{H}$ do not form a vector space, and the relation $\Lambda = \Lambda^\dagger$ needs to be taken with a grain of salt. But once properly defined, self-adjoint operators have a spectral resolution. The Hamiltonian H is an important example of a self-adjoint operator and its spectral resolution (grossly, its eigenvalues and eigenvectors) is of utmost importance. When $\mathcal{H}$ is finite dimensional an observable Λ is (modulo a choice of orthonormal basis)) a hermitian matrix. There is a finite set $S_\Lambda \subset \mathbb{R}$ and a family of projectors $P_\lambda \neq 0$, $\lambda \in S_\Lambda$ such that

$$\Lambda = \sum_{\lambda \in S_\Lambda} \lambda P_\lambda \quad P_\lambda^\dagger = P_\lambda \quad P_\lambda P_\mu = \delta_{\lambda,\mu} P_\mu \quad \sum_{\lambda \in S_\Lambda} P_\lambda = \mathrm{Id}_{\mathcal{H}}.$$

In the simplest case, Λ is non-degenerate, which means that the projectors have all rank 1, i.e. there is an orthonormal basis $\{|\, i \,\rangle,\ i \in E\}$ of $\mathcal{H}$ such that

$$\Lambda = \sum_{i \in E} \lambda_i |\, i \,\rangle\langle\, i \,| \quad \text{all } \lambda_i \text{ distinct},$$

where $|\, i \,\rangle\langle\, i \,|$ is the orthogonal projection along $|\, i \,\rangle$. Let us say it again: observables are what one can measure in experiments (!). More precisely, the outcome of an (idealized) experiment measuring the observable Λ is an element of the spectrum of Λ. But measurements deserve a special chapter on their own. Before that, we make some comments on density matrices.

2.2 Traces

The trace of an observable is an important quantity. It can be defined for every observable if $\mathcal{H}$ is finite dimensional, and for certain observables if $\mathcal{H}$ is infinite dimensional but separable (which we always assume).

If O is a positive operator in $\mathcal{B}(\mathcal{H})$ and $|\, i \,\rangle$, $i \in I$ is an orthonormal basis of $\mathcal{H}$, the number $\sum_{i \in I} \langle\, i \,|\, O \,|\, i \,\rangle \in [0, +\infty]$ is independent of the choice of orthonormal basis. If this number is finite, Λ is said to be of trace class and its trace is defined to be

$$\mathrm{Tr}_{\mathcal{H}}\, O := \sum_{i \in I} \langle\, i \,|\, O \,|\, i \,\rangle.$$

If O in$\mathcal{B}(\mathcal{H})$, the operator $O^\dagger O \in \mathcal{B}(\mathcal{H})$ is positive (it is simply O^2 is O is an observable), and has a unique positive square root $(O^\dagger O)^{1/2}$ which belongs to $\mathcal{B}(\mathcal{H})$. The operator O is said to be of trace class if $(O^\dagger O)^{1/2}$ is of trace class, i.e. if $\sum_{i\in I}\langle i\,|\,(O^\dagger O)^{1/2}\,|\,i\,\rangle < +\infty$ for some (and then every) orthonormal basis $|\,i\,\rangle$, $i \in I$ of $\mathcal{H}$. Then

$$\mathrm{Tr}_{\mathcal{H}}\, O := \sum_{i\in I}\langle\, i\,|\,O\,|\,i\,\rangle$$

is an absolutely convergent series whose value does not depend on the choice of orthonormal basis. This is of course the standard definition if $\mathcal{H}$ is finite dimensional[4].

As a trivial example, suppose that Λ is an observable (not necessarily bounded) and $0 \neq |\,\varphi\,\rangle \in \mathcal{H}$ with associated rank 1 projector P^φ. Assume that $\Lambda|\,\varphi\,\rangle$ is well defined. Then we claim that $O := \Lambda P^\varphi$ is trace class. Indeed if $|\,\psi\,\rangle \in H$, $P^\varphi|\,\psi\,\rangle = \frac{\langle\,\varphi\,|\,\psi\,\rangle}{\langle\,\varphi\,|\,\varphi\,\rangle}|\,\varphi\,\rangle$ so $O|\,\psi\,\rangle = \frac{\langle\,\varphi\,|\,\psi\,\rangle}{\langle\,\varphi\,|\,\varphi\,\rangle}\Lambda|\,\varphi\,\rangle$ is well-defined and $\|O|\,\psi\,\rangle\| = \frac{|\langle\,\varphi\,|\,\psi\,\rangle|}{\langle\,\varphi\,|\,\varphi\,\rangle}\|\Lambda|\,\varphi\,\rangle\|$ which by the Cauchy-Schwarz inequality is $\leq \||\,\psi\,\rangle\|\frac{\|\Lambda|\,\varphi\,\rangle\|}{\||\,\varphi\,\rangle\|}$ so indeed $O = \Lambda P^\varphi$ belongs to $\mathcal{B}(\mathcal{H})$. Then $O^\dagger O = \frac{\|\Lambda|\,\varphi\,\rangle\|^2}{\||\,\varphi\,\rangle\|^2}P^\varphi$ has positive square root $(O^\dagger O)^{1/2} = \frac{\|\Lambda|\,\varphi\,\rangle\|}{\||\,\varphi\,\rangle\|}P^\varphi$ which belongs to the trace class and has trace equal to $\frac{\|\Lambda|\,\varphi\,\rangle\|}{\||\,\varphi\,\rangle\|}$. Thus $O = \Lambda P^\varphi$ is trace class and $\mathrm{Tr}_{\mathcal{H}}\,\Lambda P^\varphi = \frac{\langle\,\varphi\,|\,\Lambda\,|\,\varphi\,\rangle}{\langle\,\varphi\,|\,\varphi\,\rangle}$.

The case when the system is in a pure state being settled, we turn to mixed states. That a density matrix is trace class is part of the definition. This leaves opened the question whether $\Lambda\rho$ is trace class for a given (possibly unbounded) observable Λ. At least we need that the image $\rho(\mathcal{H})$ of $\mathcal{H}$ to be included in the domain of Λ but this is not enough in general. It it enough if ρ is a *finite* combination of finite rank projectors, as shown by adapting straightforwardly the argument above.

Anyway, if the system is described by the density matrix ρ and Λ is an observable, we define the average of Λ in state ρ as

$$\langle\Lambda\rangle := \mathrm{Tr}_{\mathcal{H}}\,\Lambda\rho$$

whenever the trace makes sense. We shall see in the chapter on measurement that the average of an observable as defined by this formula is indeed a physical quantity.

After this brief incursion in infinite dimensional spaces, let us return to the case when $\mathcal{H}$ is finite dimensional.

[4]Moreover in that case any basis, even if not orthonormal, can be used to compute the trace.

In that case, the formula $\mathrm{Tr}_{\mathcal{H}}\, O^\dagger O'$ defined for every pair O, O' of linear operators on $\mathcal{H}$ endows $\mathcal{B}(\mathcal{H})$ itself with a Hilbert space structure. The trace is fully characterized by the formula $\mathrm{Tr}_{\mathcal{H}}\, |\, \varphi\, \rangle\langle\, \psi\, | = \langle\, \psi\, |\, \varphi\, \rangle$ plus additivity. Use it for example to show that $\mathrm{Tr}_{\mathcal{H}_A\otimes\mathcal{H}_B}\, O_A\otimes O_B = (\mathrm{Tr}_{\mathcal{H}_A}\, O_A)(\mathrm{Tr}_{\mathcal{H}_B}\, O_B)$ whenever O_A and O_B are operators on $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. We shall use very often the commutativity of the trace, $\mathrm{Tr}_{\mathcal{H}}\, O_1O_2 = \mathrm{Tr}_{\mathcal{H}}\, O_2O_1$ from which results the cyclicity of the trace $\mathrm{Tr}_{\mathcal{H}}\, O_1O_2\cdots O_n = \mathrm{Tr}_{\mathcal{H}}\, O_2\cdots O_nO1$.

If $\Lambda = \sum_{\lambda\in S_\Lambda} \lambda P_\lambda$ is the spectral resolution of the observable Λ, we find that

$$\langle\Lambda\rangle = \mathrm{Tr}_{\mathcal{H}}\, \Lambda\rho = \sum_{\lambda\in S_\Lambda} \lambda \mathrm{Tr}_{\mathcal{H}}\, P_\lambda\rho.$$

Observe that $0 \leq \mathrm{Tr}_{\mathcal{H}}\, P_\lambda\rho P_\lambda = \mathrm{Tr}_{\mathcal{H}}\, P_\lambda^2\rho = \mathrm{Tr}_{\mathcal{H}}\, P_\lambda\rho$ where the first equality uses that ρ is a positive operator, the second uses cyclicity and the third the definition of a projector. Moreover $\sum_{\lambda\in S_\Lambda} \mathrm{Tr}_{\mathcal{H}}\, P_\lambda\rho = \mathrm{Tr}_{\mathcal{H}}\, \left(\sum_{\lambda\in S_\Lambda} P_\lambda\right)\rho = \mathrm{Tr}_{\mathcal{H}}\, \mathrm{Id}_{\mathcal{H}}\rho = \mathrm{Tr}_{\mathcal{H}}\, \rho = 1$. So the collection $p_\lambda := \mathrm{Tr}_{\mathcal{H}}\, P_\lambda\rho$, $\lambda \in S_\Lambda$ defines a probability measure on the spectrum of Λ and what we denote by $\langle\Lambda\rangle$ is always a standard probabilistic expectation that could also be denoted by $\mathbb{E}(\Lambda)$ for some probability measure . This works well for a single observable Λ, or for families of commuting observables. But beware of this classical interpretation if non-commuting observables are considered.

We end this discussion by a very important topic when dealing with compound systems:

– Partial trace

Suppose that ρ_C describes the state of a compound system $C = A \cup B$. Suppose that we are interested only in the subsystem A, i.e. we consider only observables on A, i.e. observables Λ of the form $\Lambda = \Lambda_A \otimes \mathrm{Id}_B$ on $\mathcal{H}_C = \mathcal{H}_A \otimes \mathcal{H}_B$ where Λ_A is an observable acting only on $\mathcal{H}_A$. The average formula $\langle\Lambda\rangle = \mathrm{Tr}_{\mathcal{H}_C}\, (\Lambda_A \otimes \mathrm{Id}_B)\rho_C$ defines a linear functional on the space of observables on A, so we know by duality (in finite dimensional spaces) that there is an unique linear operator ρ_A on $\mathcal{H}_A$ such that

$$\mathrm{Tr}_{\mathcal{H}_C}\, (\Lambda_A \otimes \mathrm{Id}_B)\rho_C = \mathrm{Tr}_{\mathcal{H}_A}\, \Lambda_A\rho_A.$$

It is readily checked, as suggested (but not proved) by our choice of notation, that ρ_A is a density matrix on $\mathcal{H}_A$. This density matrix is denoted by $\mathrm{Tr}_{\mathcal{H}_B}\, \rho_C$ (don't forget that this is still an operator, acting on $\mathcal{H}_A$) and is called the partial trace of ρ_C over $\mathcal{H}_B$. The notation and name become clear with an explicit formula. The reader is urged to write this formula, but we shall only give the formula later on because we want to emphasize that the characterization is enough to carry interesting computations.

We propose four such computations to the reader.

– The first one is trivial (but useful). Check that $\mathrm{Tr}_{\mathcal{H}_A}\, \mathrm{Tr}_{\mathcal{H}_B}\, = \mathrm{Tr}_{\mathcal{H}_A\otimes\mathcal{H}_B}$.

– The second is almost trivial. Suppose that ρ_C factorizes as $\rho_C = \rho_A \otimes \rho_B$. Then $\mathrm{Tr}_{\mathcal{H}_B} \rho_C = \rho_A$.
– The third is related to the fact that the partial trace $\mathrm{Tr}_{\mathcal{H}_B}$ acts only on $\mathcal{H}_B$. In fact one can view the partial trace as an extension of the trace. The original trace is a linear map $\mathrm{Tr}_{\mathcal{H}_B} : \mathcal{B}(\mathcal{H}_B) \to \mathbb{C}$ and the extension is as a map $\mathrm{Tr}_{\mathcal{H}_B} \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)) \cong \mathcal{B}(\mathcal{H}_A) \otimes \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_A) \otimes \mathbb{C} \cong \mathcal{B}(\mathcal{H}_A)$. Then if $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{H}$ is a linear map, check that its canonical extension $\Phi : \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)) \to \mathcal{H} \otimes \mathcal{B}(\mathcal{H}_B)$ commutes with $\mathrm{Tr}_{\mathcal{H}_B}$.
– The fourth one is of some importance. It answers the question of how general $\mathrm{Tr}_{\mathcal{H}_B} \rho_C$ is when ρ_C is a pure state. Let ρ_A be a density matrix on $\mathcal{H}_A$. There is a Hilbert space $\mathcal{H}_B$ and a pure state $|\psi\rangle\langle\psi|$ on $\mathcal{H}_C = \mathcal{H}_A \otimes \mathcal{H}_B$ such that $\mathrm{Tr}_{\mathcal{H}_B} |\psi\rangle\langle\psi| = \rho_A$. In particular, even if the compound system is described by a pure state, the effective description at the level of a subsystem is generically a mixture. The reader can find a proof in chapter 11 section 11.2. *Hint*: On can take $\dim \mathcal{H}_B = \dim \mathcal{H}_A$ i.e. take for $\mathcal{H}_B$ a copy of $\mathcal{H}_A$.

2.3 Elementary geometry of density matrices

Density matrices form a convex set (obvious from the definition as positive unit trace operators)... but not a simplex in general.

We shall often work with the simplest non-trivial case, i.e. $\dim \mathcal{H} = 2$ (e.g a spin 1/2), and we take $\mathcal{H} = \mathbb{C}^2$. A general 2×2 complex matrix can be written as a linear combination of the Pauli matrices

$$\frac{1}{2}(W\mathrm{Id} + X\sigma_x + Y\sigma_y + Z\sigma_z)$$

where

$$\sigma_x := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_+ := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \sigma_- := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

and W, X, Y, Z are arbitrary complex numbers. We recall the identities $\sigma_z^2 = 1$, $\sigma_x\sigma_y = -\sigma_y\sigma_x = i\sigma_z$ and the ones obtained by circular permutations of the indices.

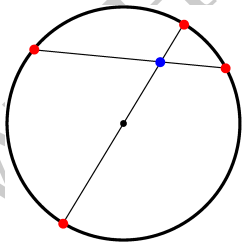
Such a matrix is positive if and only if W, X, Y, Z are real with $X^2 + Y^2 + Z^2 \leq W^2$. It has trace 1 if and only if $W = 1$. Thus it is a density matrix if and only if $W = 1$ and X, Y, Z are real with $X^2 + Y^2 + Z^2 \leq 1$, i.e. a general 2×2 density matrix reads

$$\rho := \frac{1}{2}(\mathrm{Id} + X\sigma_x + Y\sigma_y + Z\sigma_z) \quad X^2 + Y^2 + Z^2 \leq 1,$$

forming a set that can be identified with the unit ball in $\mathbb{R}^3$ (the Bloch ball?).

The characterization of pure states by $\rho^2 = \rho$ coincides for 2×2 matrices with the condition $\mathrm{Det}\,\rho = 0$ i.e. $X^2 + Y^2 + Z^2 = 1$. Thus the set of pure states can be identified with the unit sphere in $\mathbb{R}^3$, commonly called the Bloch sphere in the physics literature. Orthogonal pure states correspond to diametrically opposed points on the Bloch sphere.

Any point of the Bloch ball lays on at least one diameter, and the intersection of this diameter with the Bloch sphere gives a decomposition of the corresponding density matrix as a convex combination of two orthogonal pure states. The origin lays on any diameter and admits a continuum of such decompositions. But note that a point of the Bloch ball lays on a continuum of chords, and the intersection of each such chord with the Bloch sphere gives a decomposition of the corresponding density matrix as a convex combination of two pure –but in general not orthogonal– states. This is a fundamental arbitrariness that we alluded to before : there are in general many ways to build a statistical ensemble of pure states corresponding to a given density matrix.



Picture 12. *Geometry of the Bloch disk.*

When $\dim \mathcal{H} = d > 2$ the situation is much more complicated. In particular, pure states are extremal points of the space of density matrices, but are far from exhausting the boundary. The space of density matrices has (real) dimension $d(d+1)/2$ and its boundary, the singular density matrices ($\mathrm{Det}_{\mathcal{H}}\,\rho = 0$) has (real) dimension $d(d+1)/2 - 1$ while the pure states form a space of complex dimension $d-1$ i.e. of real dimension $2d-2$.

CHAPTER 3

Measurements in a Nutshell

Contents

It is an understatement that measurement is one of the most debated issues in quantum mechanics. On the other hand, there is a whole elaborated theory of measurement from a purely technical and practical viewpoint. This is because most real life measurements are quite far from what is presented in elementary courses. Here we shall content with slight variants of the elementary theory of ideal von Neumann measurements because properly interpreted they are enough for the special aspects of open quantum systems that we shall touch upon in these lectures.

3.1 Ideal von Neumann measurements

Ideal Von Neumann measurements are at the heart of the concept of measurement in quantum mechanics, and it is no surprise that they form the basis of our presentation.

–**Measurement postulate** (1)
(*pure states*) Suppose that the state of the system is described by the pure state $|\varphi\rangle$ just *before* the (ideal von Neumann) measurement of the observable Λ:

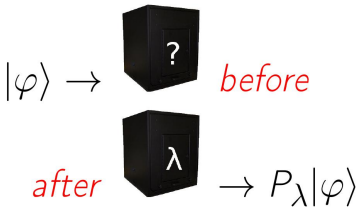
The state of the system just *after* the (ideal von Neumann) measurement of the observable Λ is

$$\text{(Unnormalized)} \qquad P_\lambda|\,\varphi\,\rangle \quad \text{with probability } \frac{\|P_\lambda|\,\varphi\,\rangle\|^2}{\||\,\varphi\,\rangle\|^2},$$

$$\text{(Normalized)} \qquad \frac{P_\lambda|\,\varphi\,\rangle}{\|P_\lambda|\,\varphi\,\rangle\|} \quad \text{with probability } \|P_\lambda|\,\varphi\,\rangle\|^2.$$

This is the famous and mysterious phenomenon of the "collapse of the wave function": measurement is not unitary and is not described by the standard unitary evolution postulate. The measurement apparatus displays λ.

Thus the measurement apparatus is a black box :



The output of the apparatus is λ with probability

$$\frac{\|P_\lambda|\,\varphi\,\rangle\|^2}{\||\,\varphi\,\rangle\|^2} = \frac{\langle\,\varphi\,|\,P_\lambda\,|\,\varphi\,\rangle}{\langle\,\varphi\,|\,\varphi\,\rangle} = \mathrm{Tr}_{\mathcal{H}}\,P_\lambda\frac{|\,\varphi\,\rangle\langle\,\varphi\,|}{\langle\,\varphi\,|\,\varphi\,\rangle} = \mathrm{Tr}_{\mathcal{H}}\,P_\lambda P^\varphi.$$

Ideal measurements lead to $\mathbb{L}^2$ projections (like conditional expectations). The state of the system after measurement depends not on Λ, but only on the family of projectors P_λ. The observable Λ plays the role of a random variable, the algebra generated by the family of projectors P_λ plays the role of the σ-algebra associated to a random variable.

The generalization to density matrices is straightforward
–**Measurement postulate** (2)
(*Density matrices*) Suppose that the state of the system is described by the density matrix ρ just *before* the (ideal von Neumann) measurement of the observable Λ:
The state of the system just *after* the (ideal von Neumann) measurement of the observable Λ is

$$\frac{P_\lambda\rho P_\lambda}{\mathrm{Tr}_{\mathcal{H}}\,P_\lambda\rho P_\lambda} \quad \text{with probability } \mathrm{Tr}_{\mathcal{H}}\,P_\lambda\rho P_\lambda = \mathrm{Tr}_{\mathcal{H}}\,P_\lambda\rho$$

Note that $\mathrm{Tr}_{\mathcal{H}} P_\lambda \rho P_\lambda \geq 0$ and $\sum_{\lambda \in S_\Lambda} \mathrm{Tr}_{\mathcal{H}} P_\lambda \rho = 1$ as it should be for the probabilistic interpretation. Note also that even if ρ describes a mixture, if the observable is non-degenerate i.e. if each P_λ has rank 1 (or if the specific λ we are interested in is such that P_λ has rank 1), the state after the measurement is pure.

The above definition of measurement assumes that the process is instantaneous, which must be an idealization. In general, there must be a piece of unitary evolution that superimposes to those formulæ. We shall put this possibility to good use later on.

3.2 Averages

The probabilities defined above could/should be interpreted in a statistical sense

Suppose that the observable Λ is measured on many independent systems $\Sigma_1, \cdots, \Sigma_N$, all described by the same $|\,\varphi\,\rangle$. Let N_λ be the number of times the result $\lambda \in S_\Lambda$ has been recorded.
By the law of large numbers,

$$\frac{N_\lambda}{N} \simeq \frac{\||P_\lambda|\,\varphi\,\rangle\|^2}{\||\,\varphi\,\rangle\|^2} = \frac{\langle\,\varphi\,|\,P_\lambda\,|\,\varphi\,\rangle}{\langle\,\varphi\,|\,\varphi\,\rangle}.$$

The empirical average of the measure is

$$\sum_{\lambda \in S_\Lambda} \frac{N_\lambda}{N} \lambda \simeq \langle \Lambda \rangle := \frac{\langle\,\varphi\,|\,\Lambda\,|\,\varphi\,\rangle}{\langle\,\varphi\,|\,\varphi\,\rangle} = \mathrm{Tr}_{\mathcal{H}}\, \Lambda \frac{|\,\varphi\,\rangle\langle\,\varphi\,|}{\langle\,\varphi\,|\,\varphi\,\rangle} = \mathrm{Tr}_{\mathcal{H}}\, \Lambda P^\varphi.$$

The generalization to density matrices is again straightforward. Suppose now that on many independent systems $\Sigma_1, \cdots, \Sigma_N$, all described by the same density matrix ρ and Λ is measured on each of them. With the same notations as above, by the law of large numbers,

$$\frac{N_\lambda}{N} \simeq \mathrm{Tr}_{\mathcal{H}}\, P_\lambda \rho$$

The empirical average of the measure is

$$\sum_{\lambda \in S_\Lambda} \frac{N_\lambda}{N} \lambda \simeq \langle \Lambda \rangle := \mathrm{Tr}_{\mathcal{H}}\, \Lambda \rho.$$

3.3 More on measurement

One of the ways to motivate the introduction of density matrices is by considering the succession of two measurements.

3.3.1 Succession of two ideal von Neumann measurements

Suppose that the observable $\Lambda = \sum_{\lambda \in S_\Lambda} \lambda P_\lambda$ is measured on a system Σ described by $|\,\varphi\,\rangle$. Suppose that after the measurement of Λ the system is sent for another measurement of the observable $\Lambda' = \sum_{\lambda' \in S_{\Lambda'}} \lambda P_{\lambda'}$.

Then the probability to measure λ in the first experiment and $\lambda' \in S_{\Lambda'}$ in the second experiment is

$$\frac{\langle\,\varphi\,|\,P_\lambda\,|\,\varphi\,\rangle}{\langle\,\varphi\,|\,\varphi\,\rangle}\frac{\langle\,\varphi\,|\,P_\lambda P_{\lambda'} P_\lambda\,|\,\varphi\,\rangle}{\langle\,\varphi\,|\,P_\lambda\,|\,\varphi\,\rangle} = \frac{\langle\,\varphi\,|\,P_\lambda P_{\lambda'} P_\lambda\,|\,\varphi\,\rangle}{\langle\,\varphi\,|\,\varphi\,\rangle}$$

The probability to measure $\lambda' \in S_{\Lambda'}$ in the second experiment, (the result of the first measurement having been recorded or not) is

$$\sum_{\lambda \in S_\Lambda} \frac{\langle\,\varphi\,|\,P_\lambda P_{\lambda'} P_\lambda\,|\,\varphi\,\rangle}{\langle\,\varphi\,|\,\varphi\,\rangle} = \mathrm{Tr}_{\mathcal{H}} = \mathrm{Tr}_{\mathcal{H}}\, P_{\lambda'} \left(\sum_{\lambda \in S_\Lambda} P_\lambda P^\varphi P_\lambda \right)$$

In general, $\sum_{\lambda \in S_\Lambda} P_\lambda P^\varphi P_\lambda$ is not of the form P^ψ for any $|\,\psi\,\rangle$, in particular it is not equal to P^φ : forgetting about the result of the first measurement and not making the first measurement are radically different procedures. However, we note that $\rho = \sum_{\lambda \in S_\Lambda} P_\lambda P^\varphi P_\lambda$ is a *bona fide* density matrix, whose associated expectation leads to a correct definition for the average of Λ

We take this opportunity to have a glance at conditional probabilities.

In the setting of the succession of two ideal von Neumann measurements, first Λ then λ' we find/define:
– The probability that the second measurement yields λ' knowing that the first measurement yielded λ:

$$P(\Lambda' = \lambda' | \Lambda = \lambda) := \frac{\langle\,\varphi\,|\,P_\lambda P_{\lambda'} P_\lambda\,|\,\varphi\,\rangle}{\langle\,\varphi\,|\,\varphi\,\rangle} \Big/ \frac{\langle\,\varphi\,|\,P_\lambda\,|\,\varphi\,\rangle}{\langle\,\varphi\,|\,\varphi\,\rangle} = \frac{\langle\,\varphi\,|\,P_\lambda P_{\lambda'} P_\lambda\,|\,\varphi\,\rangle}{\langle\,\varphi\,|\,P_\lambda\,|\,\varphi\,\rangle}$$

– The probability that the first measurement yielded λ knowing that the second measurement yielded λ':

$$P(\Lambda = \lambda | \Lambda' = \lambda') := \frac{\langle\,\varphi\,|\,P_\lambda P_{\lambda'} P_\lambda\,|\,\varphi\,\rangle}{\sum_{\tilde{\lambda} \in S_\Lambda} \langle\,\varphi\,|\,P_{\tilde{\lambda}} P_{\lambda'} P_{\tilde{\lambda}}\,|\,\varphi\,\rangle}$$

Thus the temporal order of measurements induces an asymmetry.

3.3.2 Measurements with partial reading of the outcome

This is motivated by the previous discussion when we explored what happened when the first measurement was performed but we simply forgot about. There is however no need to be so extreme and the procedure has some realistic aspects. Suppose for example that our apparatus to measure positions does it with a precision of 1 micrometer but the result is printed with a precision of 1 millimeter. More generally, suppose that Λ is measured, but only a (deterministic) many-to-one function of the result is recorded. This means that S_Λ is partitioned as $S_\Lambda = \cup_r S_{\Lambda,r}$ (which we abbreviate as $S = \cup_r S_r$ as only a single observable is involved in the discussion) and the only information we extract is within which S_r the eigenvalue was found. Then the statistical interpretation leads to the formula:
(*partial reading*) Suppose that the state of the system is described by the density matrix ρ just *before* the (ideal von Neumann) measurement of the observable Λ:
The state of the system just *after* the (ideal von Neumann) measurement with partial reading of the observable Λ as belonging to the subset S_r of the spectrum is

$$\frac{\sum_{\lambda\in S_r} P_\lambda \rho P_\lambda}{\mathrm{Tr}_{\mathcal{H}} \sum_{\lambda\in S_r} P_\lambda \rho P_\lambda} \quad \text{with probability } \mathrm{Tr}_{\mathcal{H}} \sum_{\lambda\in S_r} P_\lambda \rho P_\lambda = \mathrm{Tr}_{\mathcal{H}} \sum_{\lambda\in S_r} P_\lambda \rho$$

Let us insist again that even if ρ describes a pure state, the state after measurement with partial reading does not have to be pure anymore.

CHAPTER 4

Indirect measurements

Contents

The general idea of indirect measurements is central in the study of open systems in general. We work with a compound system $C = A \cup B$, wiht Hilbert space $\mathcal{H}_C = \mathcal{H}_A \otimes \mathcal{H}_B$. Our interest is in A *but* we measure an observable on B.

4.1 Indirect measurements

An observable Λ on C that acts only on B is of the form $\Lambda = \mathrm{Id}_A \otimes \Lambda_B$. If the spectral resolution of Λ_B reads $\Lambda_B = \sum_{\lambda \in S_{\Lambda_B}} \lambda P_{B,\lambda}$,the spectral resolution of Λ reads $\Lambda = \sum_{\lambda \in S_{\Lambda_B}} \lambda \mathrm{Id}_A \otimes \lambda P_{B,\lambda}$. If the initial state of C is described by ρ_C, the eigenvalue λ is observed with probability $\mathrm{Tr}_{\mathcal{H}_C} \left(\mathrm{Id}_A \otimes P_{B,\lambda}\right) \rho_C \left(\mathrm{Id}_A \otimes P_{B,\lambda}\right)$, and then the state of the compound system C becomes

$$\frac{\left(\mathrm{Id}_A \otimes P_{B,\lambda}\right) \rho_C \left(\mathrm{Id}_A \otimes P_{B,\lambda}\right)}{\mathrm{Tr}_{\mathcal{H}_C} \left(\mathrm{Id}_A \otimes P_{B,\lambda}\right) \rho_C \left(\mathrm{Id}_A \otimes P_{B,\lambda}\right)}.$$

If the initial state describes independent systems A and B so that $\rho_C = \rho_A \otimes \rho_B$ the final state is factorized again as $\rho_A \otimes \rho'_B$ where ρ'_B is, as should be expected, the state of B after the measurement of the value λ for Λ_B on B, and the probability is also the expected one.

Thus, the interest is when ρ_C is not factorized. As we shall explain later, an important special case is the following. Before some initial time, A and B are independent: typically two objects that are kept separated and the

state of the compound system is factorized. Then they are put in contact for a while, and they interact, so that the compound system is subject to a unitary evolution. This is a rough approximation of the situation when A and B were two particles that experience a collision. More relevant to our discussion are experiments of the type of the one presented in Pict.3. Then they are separated again and a measurement of Λ_B is performed on B. In this view, A is considered as the system of interest and B as a tool to probe A. Thus A and B play fundamentally asymmetric roles in the discussion. In equations, one starts from an initial density matrix $\rho_C = \rho_A \otimes \rho_B$ and applies a unitary evolution, so that just before the measurement the density matrix is $U\rho_C U^{-1}$. Note that U describes the full evolution, taking into account interactions between A and B but also their "free" evolution (which acts separately on A and B). Then a measurement of Λ_B is performed on B. Adapting the previous formula we find that the eigenvalue λ is observed with probability $\mathrm{Tr}_{\mathcal{H}_C}\left(\mathrm{Id}_A \otimes P_{B,\lambda}\right) U \left(\rho_A \otimes \rho_B\right) U^{-1} \left(\mathrm{Id}_A \otimes P_{B,\lambda}\right)$, and then the state of the compound system C becomes

$$\frac{\left(\mathrm{Id}_A \otimes P_{B,\lambda}\right) U \left(\rho_A \otimes \rho_B\right) U^{-1} \left(\mathrm{Id}_A \otimes P_{B,\lambda}\right)}{\mathrm{Tr}_{\mathcal{H}_C}\left(\mathrm{Id}_A \otimes P_{B,\lambda}\right) U \left(\rho_A \otimes \rho_B\right) U^{-1} \left(\mathrm{Id}_A \otimes P_{B,\lambda}\right)}. \tag{4.1}$$

If our real interest is in A, and B is just used to probe A, it makes sense to take the partial trace over B and view the complete process as a (random) transformation

$$\rho_A \to \rho'_A := \frac{\mathrm{Tr}_{\mathcal{H}_B}\left(\mathrm{Id}_A \otimes P_{B,\lambda}\right) U \left(\rho_A \otimes \rho_B\right) U^{-1} \left(\mathrm{Id}_A \otimes P_{B,\lambda}\right)}{\mathrm{Tr}_{\mathcal{H}_C}\left(\mathrm{Id}_A \otimes P_{B,\lambda}\right) U \left(\rho_A \otimes \rho_B\right) U^{-1} \left(\mathrm{Id}_A \otimes P_{B,\lambda}\right)} \tag{4.2}$$

with probability $\pi_\lambda := \mathrm{Tr}_{\mathcal{H}_C}\left(\mathrm{Id}_A \otimes P_{B,\lambda}\right) U \left(\rho_A \otimes \rho_B\right) U^{-1} \left(\mathrm{Id}_A \otimes P_{B,\lambda}\right)$.

Let us observe that ρ'_A is given as a ratio, and the numerator and denominator can be viewed as linear operations on linear operators on $\mathcal{H}_A$. Quantum mechanics as made us familiar with linear operators on $\mathcal{H}_A$, but not so much with linear operations on linear operators on $\mathcal{H}_A$ (with the expection of conjugation by the evolution operator of course). Such operators are usually called super-operators in the literature. The name is does not look very good to the author, but it is good to have a name and so we shall conform to standard practice.

The numerator defines a map

$$O \to O' := \mathrm{Tr}_{\mathcal{H}_B}\left(\mathrm{Id}_A \otimes P_{B,\lambda}\right) U \left(O \otimes \rho_B\right) U^{-1} \left(\mathrm{Id}_A \otimes P_{B,\lambda}\right)$$

from operators to operators. It maps positive operators to positive operators but does not preserve the trace (that's why there is a denominator!). It

is easily seen that this denominator which maps operators to scalars, can be expressed as

$$\mathrm{Tr}_{\mathcal{H}_C}\ (\mathrm{Id}_A \otimes P_{B,\lambda})\, U\, (O \otimes \rho_B)\, U^{-1}\, (\mathrm{Id}_A \otimes P_{B,\lambda}) = \mathrm{Tr}_{\mathcal{H}_A}\, O'.$$

The denominator is exactly the probability of the observation: this fact has deep consequences on conserved quantities as we shall amply emphasize all along these lectures. Here is the main fact. Suppose that instead of carrying the measurement, we simply take the effective dynamics on A after the evolution by U, i.e. we define $\Psi(\rho_A) := \mathrm{Tr}_{\mathcal{H}_B}\, U\, (\rho_A \otimes \rho_B)\, U^{-1}$, yet another super-operation. Comparing with the result of doing the measurement without reading the result, i.e. with the average of ρ'_A with respect to measurement outcomes, we see see the compensation between denominators and probabilities at work and get[1]:

$$\begin{aligned}
\mathbb{E}(\rho'_A) &:= \sum_{\lambda\in S} \frac{\mathrm{Tr}_{\mathcal{H}_B}\ (\mathrm{Id}_A \otimes P_{B,\lambda})\, U\, (\rho_A \otimes \rho_B)\, U^{-1}\, (\mathrm{Id}_A \otimes P_{B,\lambda})}{\pi_\lambda}\, \pi_\lambda \\
&= \sum_{\lambda\in S} \mathrm{Tr}_{\mathcal{H}_B}\ (\mathrm{Id}_A \otimes P_{B,\lambda})\, U\, (\rho_A \otimes \rho_B)\, U^{-1}\, (\mathrm{Id}_A \otimes P_{B,\lambda}) \\
&= \sum_{\lambda\in S} \mathrm{Tr}_{\mathcal{H}_B}\ (\mathrm{Id}_A \otimes P_{B,\lambda})\, U\, (\rho_A \otimes \rho_B)\, U^{-1} \\
&= \mathrm{Tr}_{\mathcal{H}_B} \sum_{\lambda\in S} (\mathrm{Id}_A \otimes P_{B,\lambda})\, U\, (\rho_A \otimes \rho_B)\, U^{-1} \\
&= \mathrm{Tr}_{\mathcal{H}_B}\, U\, (\rho_A \otimes \rho_B)\, U^{-1} = \Psi(\rho_A).
\end{aligned}$$

To summarize this computation, the effective time evolution of A without measurement coincides with the average under measurement outcomes of the effective time evolution of A followed by measurement. Trivial as it may seem, this observation has deep consequences.

The formulæ can be simplified a bit by assuming that the measurement on B is non degenerate i.e. $\Lambda_B = \sum_\alpha \lambda_I |\, I\,\rangle\langle\, I\,|$ where $\{|\, I\,\rangle\}$ is an orthonormal basis of $\mathcal{H}_B$ and the λ_Is are distinct so $P_{B,\lambda_I} = |\, I\,\rangle\langle\, I\,|$. Then Equation 4.1 becomes

$$\rho_C \to \rho'_C := \frac{\langle\, I\,|\, U\, (\rho_A \otimes \rho_B)\, U^{-1}\,|\, I\,\rangle}{\mathrm{Tr}_{\mathcal{H}_A}\, \langle\, I\,|\, U\, (\rho_A \otimes \rho_B)\, U^{-1}\,|\, I\,\rangle} \otimes |\, I\,\rangle\langle\, I\,|$$

with probability $\mathrm{Tr}_{\mathcal{H}_A}\, \langle\, I\,|\, U\, (\rho_A \otimes \rho_B)\, U^{-1}\,|\, I\,\rangle$. Notice that $\langle\, I\,|\, U\, (\rho_A \otimes \rho_B)\, U^{-1}\,|\, I\,\rangle$

[1] The crucial step is the middle one, which uses the cyclicity of the partial trace on $\mathcal{H}_B$ when operators acting only on $\mathcal{H}_B$ are moved around. The reader is invited to check the validity by using the "conceptual" definition of the partial trace.

is to be interpreted as an operator on $\mathcal{H}_A$ in an obvious way[2]. Observe that when the measurement on B is non degenerate the density matrix of C is again a pure tensor product, i.e. A has a definite density matrix after the measurement and the partial trace on $\mathcal{H}_B$ becomes inessential/tautological. Thus to summarize:

$$\rho_A \to \rho'_A \quad := \quad \frac{\langle I | U (\rho_A \otimes \rho_B) U^{-1} | I \rangle}{\mathrm{Tr}_{\mathcal{H}_A} \langle I | U (\rho_A \otimes \rho_B) U^{-1} | I \rangle} \quad \text{with probability } \mathrm{Tr}_{\mathcal{H}_A} \langle I | U (\rho_A \otimes \rho_B) U^{-1} | I \rangle. \tag{4.3}$$

To make contact with the general formulæ in Equation 1.4, we also diagonalize ρ_B as $\rho_B = \sum_i p_i |i\rangle\langle i|$, where $|i\rangle$ is another orthonormal basis of $\mathcal{H}_B$ and set $V_{Ij} := \sqrt{p_j}\langle I | U | j \rangle$ so that Equation 4.3 becomes

$$\rho_A \to \rho'_A = \frac{\sum_j V_{Ij} \rho_A V_{Ij}^\dagger}{\mathrm{Tr}_{\mathcal{H}_A} \sum_j V_{Ij} \rho_A V_{Ij}^\dagger} \text{ with probability } \mathrm{Tr}_{\mathcal{H}_A} \sum_j V_{Ij} \rho_A V_{Ij}^\dagger. \tag{4.4}$$

Using that $V_{Ij}^\dagger = \sqrt{p_j}\langle j | U^\dagger | I \rangle$ one checks that $\sum_{I,j} V_{Ij}^\dagger V_{Ij} = \mathrm{Id}_{\mathcal{H}_A}$. Thus we recover Equation 1.4 if we set (symbolically) $S := \{a\} = \{(Ij)\}$ and $\{r\} = \{I\}$ and $S_I = \{(Ij)\}$.

The formulæ in Equations (4.3,4.4) for indirect measurement are the main building blocks for pour approach to Markovian open quantum systems: they summarize all the subject. Under the assumption that the time evolution is given by the iteration of this dynamical map, the rest is exploiting consequences: derivation of continuous time evolution, study of examples,... This will occupy us for the second half of the lectures, but we shall immediately start with an illustration in the next chapter.

4.2 A remark on measurement

It is important to stress that measurement plays (at least) two roles in quantum mechanics. The first, and most obvious, is to learn something about the state of a system, and we have reviewed why the information gained in a quantum measurement is statistical in essence. The second, but not less important, is to prepare a system in some state of interest, either a pure state –this is always the case modulo experimental errors if the

[2] A intrinsic definition is possible, but one can simply think in terms of matrix coefficients : an operator O on $\mathcal{H}_C$ is an object on four free indices, two pertaining to $\mathcal{H}_A$ and two pertaining to $\mathcal{H}_B$, so a sandwich of O between any bra and any ket of $\mathcal{H}_B$ gives an object where the $\mathcal{H}_B$ indices are frozen and two free indices on $\mathcal{H}_A$ remain. In the case at hand, and intrinsically, for $|\varphi\rangle, |\psi\rangle \in \mathcal{H}_A$, $\langle \varphi | (\langle I | O | I \rangle) | \psi \rangle := (\langle \varphi | \otimes \langle I |) O (|\varphi\rangle \otimes |I\rangle)$. Of course other sandwiches with others interpretations do exist.

eigenvalue selected in the measurement has multiplicity 1– or a mixture –for instance via partial reading.

An important idea for the rest of these lectures is that indirect measurements blow up two cornerstones of the elementary presentation of idealized measurements in quantum mechanics:
– That the measurement takes places instantaneously. Indeed, we have seen that from the point of view of A is is very natural to consider the interaction with the probe B by incorporating an unitary evolution as part of the measurement process.
– That measurement is a somewhat brutal process that arises when a quantum system is put in contact with a large/classical system. For indirect measurements, the effect is brutal on B, with standard collapse. But if the interaction is small or exists only for a short time, i.e. if U is close to the identity, the effect on A can be small, and shows no apparent sign of collapse. Instead of thinking of U as close to the identity, one could look at another extreme and take U to have such a drastic effect that collapse on B induces a brutal effect (a standard or non-standard collapse) on A as well. However, it seems that to get collapse on A we would moreover need to take a very large B and do partial reading, i.e. have B play the role of a large/classical apparatus.

Thus indirect measurements give us an image of a measurement on A which takes some time, and is obtained by putting A in contact with a small/quantum apparatus. However it is possible to make several several indirect measurements with different/independent probes in succession, with a large spectrum of possible asymptotic behaviors. One of them is that an appropriate succession of indirect measurements leads to a progressive collapse of the wave function of A, just as if the compound system made of a large number of probes played the role of a large/classical apparatus.This is the subject of the next section.

These observations do not intend to mean that our presentation is free of idealizations. In particular, the measurement on B is really thought of as an idealized one. But at least we have gained a lot of flexibility.

CHAPTER 5

A first application: progressive collapse

Contents

Our aim is this chapter is to give a theoretical modeling applying to a class of experiments including the cavity QED experiment, whose principle we recall.

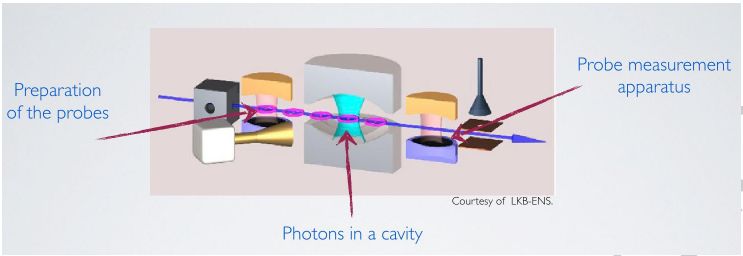
5.1 The cavity QED setup

An high quality factor cavity[1] is exited in its fundamental mode, and can be modeled by a single harmonic harmonic oscillator whose energy levels are proportional to the number of photons. A Rydberg atom[2] is prepared in a given state (left) and sent through the cavity. There, it interacts with the cavity electromagnetic field. Essential for the rest of the analysis is the fact that this interaction does not change the photon number. After some time, the atom leaves the cavity and the coupling of the two ceases. Later, a standard von Neumann measurement of the energy is performed on the atom (right). Things are tuned in such a way that only two energy levels of

[1] A typical order of magnitude is that the size of the cavity is a few centimeters and the number of reflexions before absorption is of order 10^9, allowing the photon to make about 10^4 to 10^5 kilometers.

[2] (An atom in a highly excited state close to ionization, one of whose electrons has a very large orbital number, such atoms show a very large response to electromagnetic fields.

the atom are significantly coupled by the fundamental mode of the cavity, to the atom can be treated to a good approximation as a two level system.

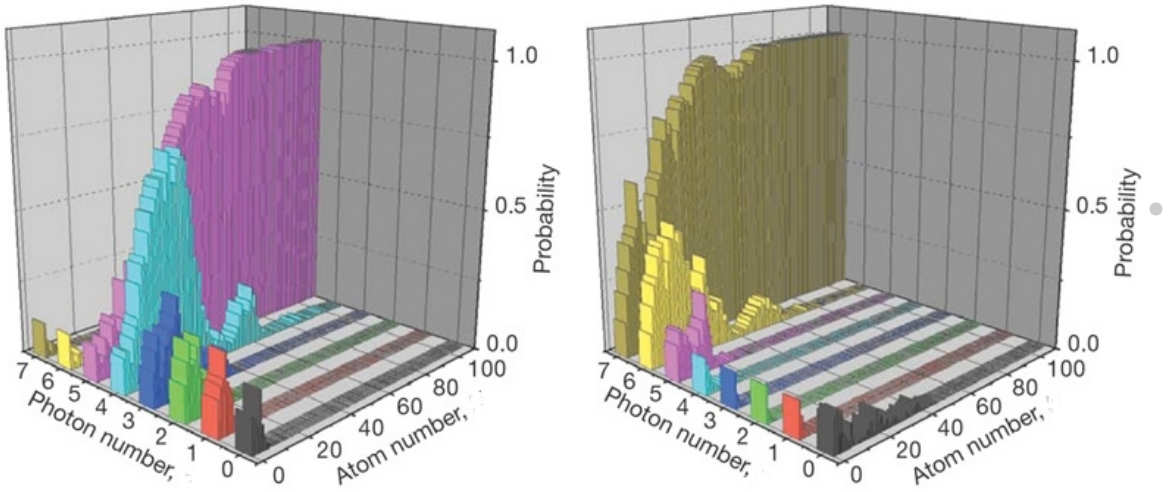


Picture 13. *The setup of a cavity QED experiment. The Rydberg atoms are prepared on the left, sent to the cavity in the middle and measured on the left.*

Before entering the cavity, the Rydberg atom is not intricated with the cavity field, but after the interaction period intrication is present, so that the energy measurement of the atom affects the cavity (or, maybe more properly, affects the knowledge we have of the field in the cavity).

The principle of the experiment is to iterate this procedure, i.e. implement the same protocol for a large number of Rydberg atoms. Each measurement on an atom gives new information on the field content of the cavity. On can say that the Rydberg atoms probe the field in the cavity, and the generic name *probe* is used as a synonym of Rydberg atom in the sequel. And the term *experiment* is used for a complete run, involving a large number of probes. Let us stress again that the direct information collected in the experiment is a listing of the energy level in which each probe is found.

Different fragrances of statistical analysis can be applied, all of them leading to the result that when the number of probes gets large the information collected is compatible only with the presence of a definite number of photons in the cavity, but this number varies from one experiment to another.



Picture 14. *Progressive collapse of the photon content of the cavity. The initial state of the cavity is the same in both experiments, but the fate is different.*

5.2 Modeling

We start with modeling the passage of a single probe[3]. The system is thus A and the probe B. Before the interaction, the density matrix is factorized as $\rho_C = \rho_A \otimes \rho_B$. To generalize the property of the cavity QED experiment, namely that the interaction between the Rydberg atom and the cavity does not change the photon number, we assume that there is an orthonormal basis $\{|\alpha\rangle\}$ of $\mathcal{H}_A$ such that the evolution operator U describing the interaction of the system with the probe decomposes as $U = \sum_\alpha |\alpha\rangle\langle\alpha| \otimes U_\alpha$, where each U_α is unitary on $\mathcal{H}_B$. The basis vector for $\mathcal{H}_A$ with this property are also called pointer states in the literature. The fact that, bu construction, the interaction between A and B does not change the pointer state has been coined by the name *quantum non-demolition measurement*. In the case of the cavity QED experiment, the basis $\{|\alpha\rangle\}$ is just the cavity photon number basis. It is immediate that U is a unitary operator on $\mathcal{H}_C = \mathcal{H}_A \otimes \mathcal{H}_B$. Thus the state of C just before the measurement on the probe is

$$\sum_{\alpha,\beta} (|\alpha\rangle\langle\alpha| \otimes U_\alpha)(\rho_A \otimes \rho_B)\left(|\beta\rangle\langle\beta| \otimes U_\beta^\dagger\right) = \sum_{\alpha,\beta} \rho_{\alpha,\beta} |\alpha\rangle\langle\beta| \otimes U_\alpha \rho_B U_\beta^\dagger$$

where we have set $\rho_{\alpha,\beta} := \langle\alpha|\rho_A|\beta\rangle$. Though this is not really fundamental for the interpretation, one can assume that the measurement on the

[3]The material in this section and the next is taken from a collaboration with Denis Bernard.

probe is non degenerate (this is certainly the case in the cavity QED experiment because the probe has only two levels). Thus we measure $\Lambda := \mathrm{Id}_{\mathcal{H}_A} \otimes \Lambda_B$ and assume a spectral resolution of the form $\Lambda_B = \sum_a \lambda_I |\, I \,\rangle\langle\, I \,|$ where $\{|\, I \,\rangle\}$ is an orthonormal basis of $\mathcal{H}_B$. Thus just after the measurement the density matrix of C is

$$\rho'_C = \rho'_A \otimes |\, I \,\rangle\langle\, I \,| \text{ with probability } \pi_I := \sum_\gamma \rho_{\gamma,\gamma} \langle\, I \,| U_\gamma \rho_B U_\gamma^\dagger |\, I \,\rangle$$

where

$$\rho'_A := \frac{\sum_{\alpha,\beta} \rho_{\alpha,\beta} |\, \alpha \,\rangle\langle\, \beta \,| \langle\, I \,| U_\alpha \rho_B U_\beta^\dagger |\, I \,\rangle}{\sum_\gamma \rho_{\gamma,\gamma} \langle\, I \,| U_\gamma \rho_B U_\gamma^\dagger |\, I \,\rangle}.$$

As observed in general, we see again that, due to the fact that the measurement on B is non degenerate, the state of C which exhibited intrication between A and B just before the measurement, is again factorized after the measurement (and of course trivially $\rho'_A = \mathrm{Tr}_{\mathcal{H}_B} \rho'_C$). Thus we are typically in a Markovian situation: repeating the operation with another probe is simply the iteration of the random dynamical map:

$$\begin{aligned} \rho_{\alpha,\beta} \quad \mapsto \quad \rho'_{\alpha,\beta} &:= \frac{\rho_{\alpha,\beta} \langle\, I \,| U_\alpha \rho_B U_\beta^\dagger |\, I \,\rangle}{\sum_\gamma \rho_{\gamma,\gamma} \langle\, I \,| U_\gamma \rho_B U_\gamma^\dagger |\, I \,\rangle} \\ &\text{with probability } \pi_I := \sum_\gamma \rho_{\gamma,\gamma} \langle\, I \,| U_\gamma \rho_B U_\gamma^\dagger |\, I \,\rangle. \end{aligned} \tag{5.1}$$

We stress again that the density matrix of A is not directly measured and that there is some conceptual issue with a purely statistical interpretation of this density matrix because as one iterates, the number of possible evolutions is exponentially large, so that even if we could do independent experiments (i.e. runs of many probes) starting with a large number L of identical initial conditions for A after about $\log L$ probes no two copies would be in the same state... or so it seems, see the remarks below. What is observed is the outcome of the measurement on each probe, and if N_J is the number of times the eigenvalue λ_I has shown up, we should supplement Equation 5.1 with the evolution of N_J. Thus, we summarize the equations describing the evolution:

$$\rho_{\alpha,\beta} \quad \mapsto \quad \rho'_{\alpha,\beta} = \rho_{\alpha,\beta} \frac{c(I|\alpha,\beta)}{\sum_\gamma \rho_{\gamma,\gamma} p(I|\gamma)} \tag{5.2}$$

$$N_J \quad \mapsto \quad N'_J = N_J + \delta_{IJ} \tag{5.3}$$

$$\text{with probability } \pi_I := \sum_\gamma \rho_{\gamma,\gamma} p(I|\gamma).$$

where we have set $c(I|\alpha,\beta) := \langle I | U_\alpha \rho_B U_\beta^\dagger | I \rangle$ and $p(I|\alpha) := c(I|\alpha,\alpha) = \langle I | U_\alpha \rho_B U_\alpha^\dagger | I \rangle$.

5.3 Solving the model

The analysis of these equations is straightforward. It relies heavily on two simple properties:
– By the positivity of ρ_B we have $p(I|\alpha) \geq 0$, and by the cyclicity of the trace $\sum_I p(I|\alpha) = \mathrm{Tr}_{\mathcal{H}_B} U_\alpha \rho_B U_\alpha^\dagger = 1$. Hence, for each α, $p(\cdot|\alpha)$ defines a probability on the set of outcomes of the probe measurement.
– The Cauchy-Schwarz inequality yields first that $|c(I|\alpha,\beta)| \leq p(I|\alpha)^{1/2} p(I|\beta)^{1/2}$ and second that $\left(\sum_I p(I|\alpha)^{1/2} p(I|\beta)^{1/2}\right)^2 \leq \left(\sum_I p(I|\alpha)\right)\left(\sum_I p(I|\beta)\right) = 1$ (the last equality being from the first property) with equality if and only if the two probability measures $p(\cdot|\alpha)$ and $p(\cdot|\beta)$ coincide.

To simplify the discussion, we make the the assumption for the rest of this section: the probability measures $p(\cdot|\alpha)$ are distinct for different $alphas$, i.e if $\alpha \neq \beta$ then $p(I|\alpha) \neq p(I|\beta)$ for some I. This entails that $|\sum_I c(I|\alpha,\beta)| < 1$ for $\alpha \neq \beta$.

The inequalities obtained so far are enough to deal with averaged equations.
They read:

$$\mathbb{E}\left(\rho'_{\alpha,\beta}\right) := \sum_I \rho_{\alpha,\beta} \frac{c(I|\alpha,\beta)}{\sum_\gamma \rho_{\gamma,\gamma} p(I|\gamma)} \sum_\gamma \rho_{\gamma,\gamma} p(I|\gamma) = \rho_{\alpha,\beta} \sum_I c(I|\alpha,\beta).$$

Thus keeping the notations of the previous chapter, $\Psi(\rho) := \mathbb{E}(\rho')$ we have

$$\mathbb{E}\left(\rho'_{\alpha,\beta}\right) = \Psi(\rho)_{\alpha,\beta} = \rho_{\alpha,\beta} c(\alpha,\beta)$$

where $c(\alpha,\beta) := \sum_I c(I|\alpha,\beta)$ so that in the α basis the action of Ψ is diagonal, with eigenvalues $c(\alpha,\beta)$, which are 1 for $\alpha = \beta$ and of modulus < 1 for $\alpha \neq \beta$. The averaged evolution is almost trivial: the diagonal elements of ρ remain constant, while the off-diagonal elements of ρ decay to 0 exponentially fast. We have rapid decoherence in the α basis,i.e. in the pointer state basis.

The formula for the average has deep probabilistic consequences on the (random) evolution of the density matrix corresponding to a succession of indirect measurements of probes. Let ρ^0 be the density matrix of A at the initial time, and ρ^n be the density matrix of A at time n, i.e. after n probes have been in contact and measured. Initialize the counters by $N_J^0 = 0\ \forall J$

and let N_J^n be the state of the counters at time n. Then:

$$\rho^{n+1}_{\alpha,\beta} = \rho^n_{\alpha,\beta} \frac{c(I|\alpha,\beta)}{\sum_\gamma \rho^n_{\gamma,\gamma} p(I|\gamma)} \tag{5.4}$$

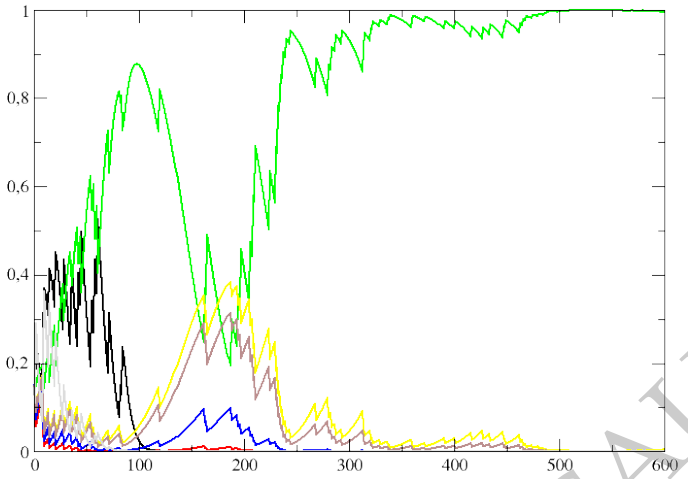
$$N_J^{n+1} = N_J^n + \delta_{IJ} \tag{5.5}$$
$$\text{with probability } \pi_I^n := \sum_\gamma \rho^n_{\gamma,\gamma} p(I|\gamma).$$

This is, as already noted, a Markov process. Thus the expectation computed above, which becomes in the new notations the average of ρ^{n+1} knowing ρ^n, is in fact the expectation of ρ^{n+1} knowing $\rho^0, \cdots, \rho^n$, or even knowing $\rho^0, I_1, \cdots, I_n$ where $I_1, \cdots, I_n$ are the first n probe measurement outcomes. A series of measurement outcomes $I_1, I_2, \cdots$ is called a history.

In the probabilistic jargon, one says that for the information collected by histories up to time n, i.e. by the first n probe measurement outcomes, $|\rho^n|$ is a martingale on the diagonal (where it coincides with ρ^n)and a super-martingale off the diagonal. The martingale convergence theorem[4] ensures that, with probability 1, $|\rho^\infty| := \lim_{n\to+\infty} |\rho^n|$ exists. In general ρ^∞ is still a random variable, i.e. it depends on the history. But for given ρ^0, the martingale convergence theorem again says that the average over histories of $|\rho^\infty|$ is equal to ρ^0 on the diagonal and less than $|\rho^0|$ of the diagonal. Intuition suggests that because the off-diagonal elements decrease in average, ρ^∞ is 0 off the diagonal, and by a fixed point point argument that diagonal elements of ρ^∞ are either 0 or 1 so that $\rho^\infty_{\alpha,\beta} = \delta_{\alpha,\beta}\delta_{\alpha,\Gamma^\infty}$ where Γ^∞ is some random variable on the set of pointer states. This in indeed true (but a clean argument is surprisingly tricky). But this means that for large n, i.e. after (infinitely) many measurements the density matrix of A collapses, i.e. iterated indirect measurements leaving pointer states invariant exhibit *progressive wave function collapse*: asymptotically, the system is in the (pure) pointer state Γ^∞. Moreover the average over histories of $\delta_{\alpha,\beta}\delta_{\alpha,\Gamma^\infty}$ is the diagonal of ρ^0. But this says exactly that $\Gamma^\infty = \gamma$ with probability $\rho^0_{\gamma,\gamma}$, which conforms to the standard *Born rule*.

Thus iterated indirect measurements leaving pointer states invariant exhibit *progressive wave function collapse* towards a (random) pointer state, and this pointer state is found with the probability given by the standard *Born rule*.

[4]The needed hypotheses are trivially satisfied because matrix elements of a density matrix are bounded.



Picture 15. *A sample illustration of the theory of progressive collapse. The diagonal elements of the density matrix are plotted. The dimension of $\mathcal{H}_A$ is 7. The martingale property entails that at each iteration the probability to collapse to a certain state is the contribution of that state.*

As mentioned above, a clean formula for the limit knowing only the property of convergence is tricky. But in the case at hand, an explicit formula for the law of the full process is easy to write down, and gives a complete solution. By another "miracle" of compensation of denominators, one has that the probability of the history $I_1, \cdots, I_N$ is

$$\begin{aligned} \text{Prob}(I_1, \cdots, I_n) &= \sum_\gamma \rho^0_{\gamma,\gamma} p(I_1|\gamma) \cdots p(I_n|\gamma) \\ &= \sum_\gamma \rho^0_{\gamma,\gamma} \prod_J p(J|\gamma)^{N^n_J} \end{aligned}$$

and for this history

$$\rho^n_{\alpha,\beta} = \rho^0_{\alpha,\beta} \frac{c(I_1|\alpha,\beta) \cdots c(I_n|\alpha,\beta)}{\sum_\gamma \rho^0_{\gamma,\gamma} p(I_1|\gamma) \cdots p(I_n|\gamma)},$$

where once again the denominator is the probability $\text{Prob}(I_1, \cdots, I_n)$. It is to be noted that the order of measurements is irrelevant, i.e. the state of the system at time n depends only on the collection N^n_J, and not on the

order of $I_1, \cdots, I_n$. Thus, the reduction of statistics [5] is a bit less awful than expected.

From then on, the discussion is easy:
– How does one sample $\mathrm{Prob}(I_1, \cdots, I_n)$? By choosing a pointer state γ with probability $\rho^0_{\gamma,\gamma}$ and then once γ is chosen, by choosing $I_1, \cdots, I_n$ independently with probability $p(\cdot|\gamma)$.
– For given γ what is the asymptotic evolution ? For given γ one can apply the usual law of large numbers: for large n,

$$\prod_J p(J|\alpha)^{N^n_J} \sim e^{n \sum_I p(J|\gamma) \log p(J|\alpha)}$$

so that

$$\frac{\rho^n_{\alpha,\alpha}}{\rho^n_{\gamma,\gamma}} \sim e^{-n \sum_I p(J|\gamma) \log \frac{p(J|\gamma)}{p(J|\alpha)}}.$$

– Note that $S(\gamma|\alpha) := \sum_I p(J|\gamma) \log \frac{p(J|\gamma)}{p(J|\alpha)}$ has a good interpretation, it is nothing but the relative entropy of $p(\cdot|\gamma)$ with respect to $p(\cdot|\gamma)$, which is > 0 if $\alpha \neq \gamma$ because the two probability measures are distinct by assumption.
– What happens off diagonal ? An easy exercise for the reader, using $|c(I|\alpha, \beta)| \leq p(I|\alpha)^{1/2} p(I|\beta)^{1/2}$.
– What is ρ^∞ if γ was the initial choice ? As the ratio $\frac{\rho^n_{\alpha,\alpha}}{\rho^n_{\gamma,\gamma}}$ decreases exponentially if $\alpha \neq \gamma$, we find $\rho^\infty_{\alpha,\beta} = \delta_{\alpha,\beta}\delta_{\alpha,\gamma}$, i.e. $\Gamma^\infty = \gamma$.
– How does one decide the value of Γ^∞ from the history ? There are at least two solutions. First, at large n the diagonal of ρ^n concentrates on one state $\rho^n_{\alpha,\alpha} \sim \delta_{\alpha,\Gamma^\infty}$. Second, at large n, by the law of large numbers, empirical frequencies $\frac{N^n_I}{n} \sim p(I|\gamma)$ for a certain γ and then $\Gamma^\infty = \gamma$ (no risk of confusion because the $p(\cdot|\gamma)$ are different for different γs).

This yields precisely the results on collapse announced above.

We end this discussion with a number of remarks:
– It is amusing, or even puzzling, that the result of the measurement, which seems to be revealed at large times, can in fact be chosen in advance. This is a consequence of De Finetti's theorem.
– It is easily seen from the explicit formula that $\mathrm{Prob}(I_1, \cdots, I_n)$ is invariant under permutations of $I_1, \cdots, I_n$, and this can be seen as a manifestation of non-demolition: in fcat one could let n probes interact with A and then measure them in a arbitrary order without changing anything. One says that the random variables $I_1, I_2, \cdots$ are exchangeable. It could seem

[5] I.e. the answer to the question "How many identical systems does one have to start with to keep a reasonable statistics after n probe measurements ?".

that the explicit formula for $\mathrm{Prob}(I_1, \cdots, I_n)$ is but a special example of exchangeability. But De Finetti's theorem says exactly the opposite[6]: if an infinite sequence of random is exchangeable, it can alway be samples by first choosing (in a certain random way) a law for one variable, and then choose all the variables independently with that law.
– The formula for the evolution of the diagonal of ρ is nothing but the usual Bayes (update)rule from a prior from the observation:

$$P_\alpha^{n+1} = P_\alpha^n \frac{p(I|\alpha)}{\sum_\gamma P_\gamma^n p(I|\gamma)} \text{ if outcome } I_{n+1} \text{ is } I$$

This formula is robust in that the initial condition is immaterial: as long as the initial prior P^0 charges all states, the large n behavior is the same. In Pict.14 the initial prior is uniform, which is known not to be the case in the cavity.

5.4 Non-demolition and commutation

For a while, we return to a general unitary evolution U for the interaction between A and B. Still assuming a non-degenerate measurement of probes, projecting on the states $|I\rangle$, we define a measurement operator $O_I := (\mathrm{Id}_{\mathcal{H}_A} \otimes |I\rangle\langle I|)\, U$ which in accordance with our general guidelines takes into account both the evolution during the interaction time and the projection due to the measurement.

Though we have not insisted on that, the description of the iterated measurement process needs to enlarge the Hilbert space because each probe comes with its own Hilbert space. Thus if n probes are to be taken into account, the Hilbert space must be

$$\mathcal{H}_n := \mathcal{H}_A \otimes \mathcal{H}_B^{\otimes n} = \mathcal{H}_A \otimes \mathcal{H}_B \cdots \otimes \mathcal{H}_B.$$

[6]Here is a more precise description. Suppose that S is a finite set. A probability measure on S is a family of numbers $p_I, I \in S$ such that $p_I \geq 0$ for each I and $\sum_I p_I = 1$. Hence the set $\mathcal{M}(S)$ of probability measures on S is a simplex, and as such, it is possible to consider probability measures on $\mathcal{M}(S)$ itself. For instance, if S is a set with two elements, a probability measure on $\mathcal{M}(S)$ s just a probability measure on the interval $[0, 1]$. The generalization to infinite sets S is possible but a bit more involved. Anyway, suppose that $X_1, X_2, \cdots$ is an infinite sequence of exchangeable random variables taking value S, i.e. such that, for each $n = 2, 3, \cdots$, $\mathrm{Prob}(X_1 = I_1, \cdots, X_n = I_n)$ is invariant under permutations of $I_1, \cdots, I_n$. De Finetti's theorem states that there exists a probability measure μ on $\mathcal{M}(S)$ such that $\mathrm{Prob}(X_1 = I_1, \cdots, X_n = I_n) = \int_{\mathcal{M}(S)} d\mu(p)\; p_{I_1} \cdots p_{I_n}$. This means that the law of exchangeable variables is a convex combination of laws of independent variables: first one choses a law p on S i.e. an element of $\mathcal{M}(S)$ using μ and then $X_1, X_2, \cdots$ are sampled independently according to p. In the special case of the main text, the measure μ is concentrated on a finite number of points, the set of pointer states.

We shall use the notation $\mathcal{H}_{B,m}$ to denote the m^{th} factor $\mathcal{H}_B$ in $\mathcal{H}_B^{\otimes n}$. Dealing with a full history would lead to $\mathcal{H}_B^{\otimes \mathbb{N}}$ but there are some subtleties to define this infinite tensor product.

For $m = 1, \cdots, n$ we define U_m (acting on $\mathcal{H}_n$) to be the evolution operator for the interaction of the m^{th} probe with A, so that U_n acts as U on $\mathcal{H}_A \otimes \mathcal{H}_{B,m}$ and as the identity on the other $\mathcal{H}_B$ factors. In the same spirit, we define the measurement operators $O_{I,m}$ describing the role of the m^{th} probe. If the density matrix of $\mathcal{H}_n$ before the passage of the m^{th} probe is ρ, then its density matrix after the measurement of the m^{th} probe is proportional to $O_{I,m}\rho O_{I,m}$ if the m^{th} probe was found in state $|\, I \,\rangle$.

We claim that the following conditions are equivalent:
i) There is an orthonormal basis $|\, \alpha \,\rangle$ of $\mathcal{H}_A$ and unitary operators $U^{(\alpha)}$ on $\mathcal{H}_B$ such that $U = \sum_\alpha |\, \alpha \,\rangle\langle\, \alpha \,| \otimes U^{(\alpha)}$
ii) The operators $O_{I,m}$ commute for different Is and ms

We define the iterated measurement process to be non-demolition if these conditions hold. Condition i) is the one we used before, and we see its close connection with a commutativity condition which the the operator expression of exchangeability.

To prove the equivalence, it is enough to deal with the case $n = 2$ (why?).
– We first check that i) $\Rightarrow$ ii). It is immediate that $O_{I,1}O_{J,2} = \sum_\gamma |\, \gamma \,\rangle\langle\, \gamma \,| \otimes |\, I \,\rangle\langle\, I \,| U^{(\gamma)} \otimes |\, J \,\rangle\langle\, J \,| U^{(\gamma)} = O_{J,2}O_{I,1}$.
– To check ii) $\Rightarrow$ i) we decompose $U := \sum_{IJ} U_{IJ} \otimes |\, I \,\rangle\langle\, J \,|$ where each U_{IJ} acts on $\mathcal{H}_A$). Then $O_I = \sum_K U_{IK} \otimes |\, I \,\rangle\langle\, K \,|$, and it is immediate to check that $O_{I,1}O_{J,2} = \sum_{KL} U_{IK}U_{JL} \otimes |\, I \,\rangle\langle\, K \,| \otimes |\, J \,\rangle\langle\, L \,|$ so that $O_{I,1}$ and $O_{J,2}$ (more generally the $O_{I,m}$s) commute if and only if the operators U_{IJ} on $\mathcal{H}_A$ commute for different IJs. Then the $U_{IJ}^\dagger$s commute for different IJs as well. Now multiply $U_{IJ}U_{KL} - U_{KL}U_{IJ} = 0$ on the left by $U_{II'}^\dagger$ and on the right by $U_{J'J}^\dagger$ and sum over I and J. Using $U^\dagger = \sum_{IJ} U_{IJ}^\dagger \otimes |\, J \,\rangle\langle\, I \,|$ and $U^\dagger U = UU^\dagger = \mathrm{Id}_{\mathcal{H}_A} \otimes \mathrm{Id}_{\mathcal{H}_B}$ we check $\sum_K U_{IK}U_{JK}^\dagger = \delta_{IJ}\mathrm{Id}_A$ and obtain $U_{KL}U_{J'I'}^\dagger - U_{J'I'}^\dagger U_{KL} = 0$, i.e. the $U_{IJ}^\dagger$s and the U_{KL}s commute. Thus we have a family of commuting normal operators, which can be diagonalized simultaneously in an orthonormal basis: there is an orthonormal basis $|\, \alpha \,\rangle$ of $\mathcal{H}_A$ and scalars $s_{\alpha IJ}$ such that $U_{IJ} = \sum_\alpha s_{\alpha IJ} |\, \alpha \,\rangle\langle\, \alpha \,|$. It is easily checked that each $U^{(\alpha)} := \sum_{IJ} c_{\alpha IJ} |\, I \,\rangle\langle\, J \,|$ is an unitary operator on $\mathcal{H}_B$.

Thus non-demolition measurements are precisely those for which the probes can be measured in an arbitrary order, independently of the order in which they have interacted with the system.

CHAPTER 6

General discrete Markovian evolution

Contents

6.1 Summary of indirect measurements

Our starting point is the general formula for non-degenerate indirect measurements, culminating with Equation 4.4. We summarize the discussion.

Suppose U is an unitary operator on $\mathcal{H}_A \otimes \mathcal{H}_B$ where A is the system of interest and B is the probe on which the measurement is performed. Suppose that the state of $A \cup B$ is $\rho_A \otimes \rho_B$ before the interaction, and the non-degenerate observable $\Lambda_B = \sum_a \lambda_I |\, I \,\rangle\langle\, I \,|$ ($\{|\, I \,\rangle\}$ is an orthonormal basis of $\mathcal{H}_B$ and the λ_Is are distinct) is measured on B just after the evolution described by U. Suppose moreover that ρ_B is diagonalized, $\rho_B = \sum_i p_i |\, i \,\rangle\langle\, i \,|$, where $|\, i \,\rangle$ is another orthonormal basis of $\mathcal{H}_B$ and set $V_{Ij} := \sqrt{p_j}\langle\, I \,|\, U \,|\, j \,\rangle$. Then the state of the system after the measurement resulting in the observation of λ_I is $\rho'_A \otimes |\, I \,\rangle\langle\, I \,|$ where

$$\rho'_A = \frac{\sum_j V_{Ij}\rho_A V_{Ij}^\dagger}{\mathrm{Tr}_{\mathcal{H}_A} \sum_j V_{Ij}\rho_A V_{Ij}^\dagger} \text{ with probability } \mathrm{Tr}_{\mathcal{H}_A} \sum_j V_{Ij}\rho_A V_{Ij}^\dagger. \tag{6.1}$$

One checks that $\sum_{I,j} V_{Ij}^\dagger V_{Ij} = \mathrm{Id}_{\mathcal{H}_A}$.

A priori, the number of Is and is is the same, but certain p_is might equal 0 so in fact the two are not strictly correlated. An important special case is when $\rho_B := |\psi\rangle\langle\psi|$ for some unit vector $|\psi\rangle \in \mathcal{H}_B$, resulting in

$$\rho'_A = \frac{V_I \rho_A V_I^\dagger}{\mathrm{Tr}_{\mathcal{H}_A} V_I \rho_A V_I^\dagger} \text{ with probability } \mathrm{Tr}_{\mathcal{H}_A} V_I \rho_A V_I^\dagger. \tag{6.2}$$

where we have set $V_I := \langle I|U|\psi\rangle$.

6.2 Reinterpretation

Whether or nor ρ_B is pure, we end up with a general structure involving operators V_a, $a \in S$ for a certain set S with the constraint that $\sum_a V_a^\dagger V_a = \mathrm{Id}_{\mathcal{H}_A}$, an identity which we sometimes call the *closure relation*. For a general density matrix ρ_B we may take $S := \{a\} = \{(Ij)\}$, while for a pure state we may either take the same indexing with many V_a taken to be zero or simply take $S := \{a\} = \{(I)\}$. The question now arises whether mathematics puts other constraints on the V_a's. The answer in no! To make a precise statement:
Suppose V_a, $a \in S$ is a (finite) family of operators acting on a (finite dimensional) Hilbert space $\mathcal{H}$ and such that $\sum_a V_a^\dagger V_a = \mathrm{Id}_{\mathcal{H}}$. Take $\mathcal{K}$ to be a Hilbert space of dimension #S with a fixed orthonormal basis $|a\rangle$, $a \in S$. Let $|\psi\rangle$ be any unit vector in $\mathcal{K}$. There is an unitary operator U on $\mathcal{H} \otimes \mathcal{K}$ such that $V_a = \langle I|U|\psi\rangle$.
The proof is given in section 11.2. The fact that mathematics imposes no further constraint means that only physics can guide us to decide what an appropriate choice for the V_a's is in each situation.

Taking $\mathcal{H} = \mathcal{H}_A$ and $\mathcal{H}_B = \mathcal{K}$, with initial density matrix $\rho_A \otimes |\psi\rangle\langle\psi|$, evolution operator U and measurement on $\mathcal{H}_B$ of the operator $\Lambda = \sum_{a\in S} \lambda_a |a\rangle\langle a|$ with distinct λ_as, we get the dynamical map

$$\rho_A \otimes |\psi\rangle\langle\psi| \mapsto \frac{V_a \rho_A V_a^\dagger}{\mathrm{Tr}_{\mathcal{H}_A} V_a \rho_A V_a^\dagger} \otimes |a\rangle\langle a| \text{ with probability } \mathrm{Tr}_{\mathcal{H}_A} V_a \rho_A V_a^\dagger.$$

It is also possible to do a partial reading of the results, splitting $S = \cup_{r\in R} S_r$ with an arbitrary partition, leading to

$$\rho_A \otimes |\psi\rangle\langle\psi| \mapsto \frac{\sum_{a\in S_r} V_a \rho_A V_a^\dagger \otimes |a\rangle\langle a|}{\mathrm{Tr}_{\mathcal{H}_A} \sum_{a\in S_r} V_a \rho V_a^\dagger} \text{ with probability } \pi_r := \mathrm{Tr}_{\mathcal{H}_A} \sum_{a\in S_r} V_a \rho_A V_a^\dagger.$$

This new density matrix is not factorized, but taking the partial trace over $\mathcal{H}_B$ gives an effective density matrix for A which is the general formula announced in Equation 1.4.

The physical meaning of partial reading is quite clear, especially within the statistical interpretation of density matrices. But is it intriguing that the formula in Equation 6.1 for the effective density matrix on A does not allow to distinguish between a coupling to a pure state but partial reading or coupling to a general density matrix. The difference is hidden in correlations between A and the probe space. Note also that the size of the probe space is in general not the same in both cases.

6.3 Completely positive super-operators

In the case when the measurement is not read at all, the transformation of the effective density matrix on A is of the general form announced in Equation 1.1

$$\rho_A \mapsto \rho'_A = \Psi(\rho_A) := \sum_{a \in S} V_a \rho_A V_a^\dagger,$$

The class of super-operators Ψs on $\mathcal{H}_A$ obtained when taking all possible finite families of operators V_a on $\mathcal{H}_A$ (leaving the closure relation aside for the time being) has an interesting abstract characterization: it is the class of all completely positive super-operators on $\mathcal{H}_A$. This statement is a part of Choi's theorem, see section 11.3 for a precise statement and proof. Let us simply state the definition of completely positive:
– A super-operator on $\mathcal{H}_A$ is called positive if it maps positive operators to positive operators. – If $n < +\infty$ is an integer, a super-operator Ψ on $\mathcal{H}_A$ is called n-positive if $\Psi \otimes \mathrm{Id}_{\mathcal{B}(\mathcal{H}_B)}$, seen as a super-operator on $\mathcal{H}_A \otimes \mathcal{H}_B$ is positive whenever $\mathcal{H}_B$ is a Hilbert space of dimension n.
– A super-operator Ψ on $\mathcal{H}_A$ is called completely positive if it is n-positive for $n = 1, 2, \cdots$.

In conclusion, the class of super-operator Ψs that arise by taking a partial trace of an unitary evolution is exactly the class of trace preserving completely positive super-operators.

A word of vocabulary: in the literature, a transformation of the form $\rho_A \mapsto V_a \rho_A V_a^\dagger$ is called an operation and $V_a^\dagger V_a$ is the corresponding effect. Quantum operations are introduced to embody the many facets of measurements, and the above analysis says essentially that indirect measurement provide the most general kind of quantum operations.

6.4 Markovian evolution: succession of probes

It remains to argue why iteration of the dynamical maps is a sensible description of an interesting class of physical phenomena. The case of probe succession is the clearest.

Suppose that after an indirect probe measurement (possibly with partial reading) on B, the system of interest A is exposed to another probe B′ for another similar indirect measurement process. We want to compute the effective evolution for A. So we assume that the initial state for $A \cup B \cup B'$ is $\rho_A \otimes |\psi\rangle\langle\psi| \otimes |\psi'\rangle\langle\psi'|$. After B has been measured, and the eigenvalue has been found in S_r (probability $\mathrm{Tr}_{\mathcal{H}_A} \sum_{a\in S_r} V_a \rho_A V_a^\dagger$, the new state is

$$\frac{\sum_{a\in S_r} V_a \rho_A V_a^\dagger \otimes |a\rangle\langle a|}{\mathrm{Tr}_{\mathcal{H}_A} \sum_{a\in S_r} V_a \rho V_a^\dagger} \otimes |\psi'\rangle\langle\psi'|.$$

and a small computation is needed because this is not a pure tensor product in the $A \cup B$ system. Assume that the non-degenerate observable $\Lambda' := \sum_{a'\in S'} \lambda'_{a'} |a'\rangle\langle a'|$ is to be measured on B′. The evolution U′ coupling A and B′ leaves B untouched[1], and we write $U'|\Psi'\rangle = \sum_{a'\in S'} V'_{a'} \otimes |a'\rangle$ where $U'|\Psi'\rangle$ is to be interpreted as a vector in $\mathcal{H}_{B'}$ whose components are operators on $\mathcal{H}_A$ and $V'_{a'} := \langle a'|U'|\Psi'\rangle$. Thus the state just before the measurement on B′ is

$$\sum_{a',b'\in S'} (V'_{a'} \otimes \mathrm{Id}_{\mathcal{H}_B}) \frac{\sum_{a\in S_r} V_a \rho_A V_a^\dagger \otimes |a\rangle\langle a|}{\mathrm{Tr}_{\mathcal{H}_A} \sum_{a\in S_r} V_a \rho V_a^\dagger} (V'^\dagger_{b'} \otimes \mathrm{Id}_{\mathcal{H}_B}) \otimes |a'\rangle\langle b'|.$$

When Λ′ is measured on B′ the result is will be $\lambda'_{a'}$, for $a' \in S'_r$, with probability

$$\frac{\mathrm{Tr}_{\mathcal{H}_A\otimes\mathcal{H}_B\otimes\mathcal{H}_{B'}} \sum_{a'\in S'_r} \sum_{a\in S_r} V'_{a'} V_a \rho V_a^\dagger V'^\dagger_{a'} \otimes |a\rangle\langle a| \otimes |a'\rangle\langle a'|}{\mathrm{Tr}_{\mathcal{H}_A} \sum_{a\in S_r} V_a \rho V_a^\dagger}.$$

The trace over $\mathcal{H}_B \otimes \mathcal{H}_{B'}$ for fixed a and a′ is trivial, and the formula simplifies to

$$\frac{\mathrm{Tr}_{\mathcal{H}_A} \sum_{a'\in S'_r} \sum_{a\in S_r} V'_{a'} V_a \rho V_a^\dagger V'^\dagger_{a'}}{\mathrm{Tr}_{\mathcal{H}_A} \sum_{a\in S_r} V_a \rho V_a^\dagger},$$

and the resulting density matrix is

$$\frac{\sum_{a'\in S'_r} \sum_{a\in S_r} V'_{a'} V_a \rho V_a^\dagger V'^\dagger_{a'} \otimes |a\rangle\langle a| \otimes |a'\rangle\langle a'|}{\mathrm{Tr}_{\mathcal{H}_A} \sum_{a'\in S'_r} \sum_{a\in S_r} V'_{a'} V_a \rho V_a^\dagger V'^\dagger_{a'}}.$$

Note that the factor $\mathrm{Tr}_{\mathcal{H}_A} \sum_{a\in S_r} V_a \rho V_a^\dagger$ has canceled between the numerator and the denominator. Anyway, at the level of effective density matrices

[1] That is, acts as $\mathrm{Id}_{\mathcal{H}_B}$.

on A, we have indeed an evolution $\rho_A \to \rho'_A \to \rho''_A$ with

$$\rho'_A = \frac{\sum_{a\in S_r} V_a \rho_A V_a^\dagger}{\mathrm{Tr}_{\mathcal{H}_A} \sum_{a\in S_r} V_a \rho V_a^\dagger} \text{ with probability } \mathrm{Tr}_{\mathcal{H}_A} \sum_{a\in S_r} V_a \rho V_a^\dagger$$

and then

$$\rho''_A = \frac{\sum_{a'\in S'_{r'}} V'_a \rho'_A V'^\dagger_{a'}}{\mathrm{Tr}_{\mathcal{H}_A} \sum_{a'\in S'_{r'}} V'_a \rho'_A V'^\dagger_{a'}} \text{ with probability } \mathrm{Tr}_{\mathcal{H}_A} \sum_{a'\in S'_{r'}} V'_a \rho'_A V'^\dagger_{a'}.$$

Again, cancellations between numerators and denominators occur at every corner.

We have considered two probes, but the extension to an arbitrary number of probes is straightforward. In fact it is hidden in our formulæ and revealed by a reinterpretation of independent interest. Let $U'U$ act on $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_{B'}$ in a natural way (U acts only on $\mathcal{H}_A \otimes \mathcal{H}_B$, and U' only on $\mathcal{H}_A \otimes \mathcal{H}_{B'}$. As a consequence, $U'U|\psi\rangle \otimes |\psi'\rangle = \sum_{a\in S, a'\in S'} V'_{a'} V_a |a\rangle \otimes |a'\rangle$ and if $\Lambda \otimes \Lambda'$ is measured on $\mathcal{H}_B \otimes \mathcal{H}_{B'}$ with partial reading in $S_r \times S'_{r'}$, the result is exactly the same as the original procedure : the same density matrices appear as possible results, and they appear with the same probabilities. So even if the order in which the probes interact is important, the order in which they are measured is not! Of course, strictly speaking, this is only a mathematical identity, because while B' interacts with A, B could still experience an internal evolution. So Λ and Λ' should better be conserved by the internal evolution of B and B' respectively to make this interpretation physically meaningful. But the mathematical statement by itself means that two probes in succession can be treated as just one compound probe. From there, the extension to an arbitrary number of probes is immediate.

We conclude that indirect measurements on successions of independent probes result in Markovian dynamics for the effective density matrix of the system, and it is clear that if the probes were slightly correlated (as they probably are in real experiments), the Markovian formula would remain a useful approximation.

Finally, let us observe that if we consider an infinite succession of identical probes, by combining the unitary operator that shift the probes by one unit i.e. brings the next probe close to A and the unitary evolution of that probe with A, we get a time independent unitary evolution operator, and when probe measurements (possibly with partial reading) are performed on top, this results in a time homogeneous infinite Markovian evolution for the effective density matrix describing A. The details of this construction are explained in detail in section 11.6

6.5 Markovian evolution: general discussion

So far we have worked in a context of interactions with successive probes. But if one averages (or does a measurement without reading, this is the same: remember this is one of our initial observations) the one time step evolution for the effective density matrix of A also describes the contact with an arbitrary auxiliary system, a reservoir in particular. As the previous discussion clearly exemplifies, the Markovian character of the evolution is closely tied with the fact that successive probes are independent, or to make a useful approximation at least only weakly correlated. So the question is to what extend this weak correlation assumption holds when contact with a reservoir is analyzed. At this point the discussion becomes at best qualitative. For a large reservoir in contact with the small system A, one can argue that the back-reaction of A on the reservoir is small. The degrees of freedom of the reservoir that interact with the system at a certain time then go away[2], interact with other degrees of freedom in the reservoir and loose their correlations with A in a typical relaxation time τ which should be compared to the typical time T it takes them to get back in contact with A. If $\tau \ll T$, a Markovian approximation makes sense, and it is typically as if the degrees of freedom of the reservoir that interact with A at some time t go away forever and never come back again. This can be made rigorous under certain assumptions.

Anyway, we have now reach the general laws governing the evolution of Markovian open quantum systems in discrete time. Our next task are to study simple examples, and then to describe the possible continuous time limits in general. From now on, only the effective density matrix on A plays a role, and most of the time we suppress the index A : the Hilbert space is simply denoted by $\mathcal{H}$, the density matrix by ρ, etc.

[2]Think for instance of the –classical or quantum– situation when a large reservoir of gas at a given temperature is in contact through a piston with a recipient containing just a few molecules.

CHAPTER 7

Sample illustrations

Contents

7.1 Solutions of the closure relation

In this section we look for some simple solutions to the closure relation

$$\sum_{a\in S} V_a^\dagger V_a = \mathrm{Id}_{\mathcal{H}}. \tag{7.1}$$

We start with generic solutions, and then give a classification in the case $\#S = 2$.

– **Randomization** (or **convex combinations**):
This is a generic construction. Let R be a finite index set, S_r, $r \in R$ a family of finite sets indexed by R and p_r, $r \in R$ be probability measure on R ($p_r \geq 0$ for each $r \in R$ and $\sum_{r\in R} p_r = 1$). Suppose given a solution $\sum_{a_r\in S_r} V_{r,a_r}^\dagger V_{r,a_r} = \mathrm{Id}_{\mathcal{H}}$ of the closure relation be given for each $r \in R$. Set $S = \cup_{r\in R}\{r\} \times S_r$ (the disjoint unions of the S_rs). For $a = (r, a_r) \in S$ set $W_a := \sqrt{p_r} V_{r,a_r}$. Then $\sum_{a\in S} W_a^\dagger W_a = \mathrm{Id}_{\mathcal{H}}$ gives a solution of the closure relation.

– **Hamiltonian evolution**:

When $\#S = 1$ and $\dim \mathcal{H} < \infty$, Equation 7.1 says that V is unitary. Using the randomization tricks above with several unitary operators leads to random Hamiltonian evolution.

– Classical Markovian evolution (0):
Before the next construction of solutions of the closure equation, we note that to every solution V_a, $a \in S$ of Equation 7.1 and every choice $|\alpha\rangle$, $\alpha \in R$ of an orthonormal basis of $\mathcal{H}$ is associated a classical Markov matrix[1] $M_{\alpha,\beta}$, $\alpha, \beta \in R$, namely

$$M_{\alpha,\beta} := \sum_{a\in S} \langle \alpha | V_a | \beta \rangle \langle \beta | V_a^\dagger | \alpha \rangle = \sum_{a\in S} |\langle \alpha | V_a | \beta \rangle|^2.$$

The first expression for $M_{\alpha,\beta}$ shows via the closure relation that $\sum_{\alpha\in R} M_{\alpha,\beta} = 1$ for every $\beta \in R$ and the second that $M_{\alpha,\beta} \geq 0$ for every $\alpha, \beta \in R$. Thus the two conditions in the definition of a Markov matrix are satisfied. We note also that, setting as usual $\Psi(\rho) := \sum_{a\in S} V_a \rho V_a^\dagger$ we have

$$\langle \alpha | \Psi(\rho) | \alpha \rangle = \sum_{\beta\in R} M_{\alpha,\beta} \langle \beta | \rho | \beta \rangle + \text{terms involving non-diagonal coefficients of } \rho.$$

This gives an easy access to certain spectral properties of the operator Ψ, say as an operator on self-adjoint operators. Indeed, suppose that O is a non-zero self-adjoint operator such that $\Psi(O) = lO$. By the standard spectral theorem for self-adjoint operators (in the finite dimensional setting) one can diagonalize O in an orthonormal basis, and let us chose this basis to define the stochastic matrix $M_{\alpha,\beta}$, so that $O = \sum_{\alpha\in R} O_\alpha |\alpha\rangle\langle\alpha|$. Then clearly $\Psi(O) = lO$ translates to $\sum_{\beta\in R} M_{\alpha,\beta} O_\beta = lO_\alpha$ i.e. l is in the spectrum of the matrix $M_{\alpha,\beta}$. In particular, l has modulus ≤ 1 from the standard results on Markov chains[2]. Thus, as was expected, quantum Markovian evolutions are contracting.

Now that we have seen classical Markov matrices arising from quantum Markovian evolution, it is natural to ask whether any classical Markov matrix arises in that way. It is easy to guess that the answer is yes, and we give below to different possibilities (among many others) which have quite different quantum interpretations and consequences.

– Classical Markovian evolution (1):
Suppose that $|\alpha\rangle$, $\alpha \in R$ is an orthonormal basis of $\mathcal{H}$ and let M^α_β, $\alpha, \beta \in R$

[1] I.e. the generator of a discrete time Markov chain, such a generator is also called a stochastic matrix.

[2] This result itself can also essentially be obtained by elementary means –this is Doeblin's theory, which would take us too far away.

be a (classical) stochastic matrix, i.e. the generator of a Markov chain: $M_{\alpha\beta} \geq 0$ for $\alpha, \beta \in R$ and $\sum_{\alpha\in R} M_{\alpha\beta} = 1$ for $\beta \in R$. Take $S = R$, and for $a \in S$, $\alpha, \beta \in R$ set $\langle \beta \,|\, V_a \,|\, \alpha \rangle := \delta_{a,\alpha}\sqrt{M_{\beta\alpha}}$. Then $\langle \beta \,|\, V_a^\dagger \,|\, \alpha \rangle = \delta_{a,\beta}\sqrt{M_{\alpha\beta}}$. One checks that

$$(V_a^\dagger V_a)_{\alpha\beta} = \delta_{a\alpha}\delta_{a\beta},$$

so that $\sum_{a\in S} V_a^\dagger V_a = \mathrm{Id}_{\mathcal{H}}$ and we have a solution to the closure relation. Moreover, If ρ is a density matrix on $\mathcal{H}$ one checks that

$$(V_a \rho V_a^\dagger)_{\alpha\beta} := \langle \alpha \,|\, V_a \rho V_a^\dagger \,|\, \beta \rangle = \sqrt{M_{\alpha a}}\rho_{aa}\sqrt{M_{\beta a}}.$$

In particular $\mathrm{Tr}_{\mathcal{H}} V_a \rho V_a^\dagger = \rho_{aa}$ So the corresponding evolution if "a" is measured is

$$\rho_{\alpha\beta} \to \rho'_{\alpha\beta} = \sqrt{M_{\alpha a}}\sqrt{M_{\beta a}} \text{ with probability } \rho_{aa}.$$

This example shows immediate purification: after one measurement, the density matrix is a pure state!
The average evolution is

$$\rho_{\alpha\beta} \to \rho'_{\alpha\beta} = \sum_{a\in S} \sqrt{M_{\alpha a}}\rho_{aa}\sqrt{M_{\beta a}}$$

and in particular on the diagonal $\rho'_{\alpha\alpha} = \sum_a M_{\alpha a}\rho_{aa}$. Thus, the diagonal of the density matrix evolves in average just like the probability distribution for the Markov chain with transition kernel $M_{\alpha\beta}$.

– **Classical Markovian evolution** (2):
Suppose as above that $|\, \alpha \,\rangle$, $\alpha \in R$ is an orthonormal basis of $\mathcal{H}$ and let M_β^α, $\alpha, \beta \in R$ be a (classical) stochastic matrix. But this time take $S = R \times R$ and set, for $x = (a, b) \in S$ and $\alpha, \beta \in R$, $\langle \alpha \,|\, V_x \,|\, \beta \rangle := \sqrt{M_{\alpha\beta}}\delta_{\alpha a}\delta_{\beta b}$. Then $\langle \alpha \,|\, V_x^\dagger \,|\, \beta \rangle = \sqrt{M_{\beta\alpha}}\delta_{\beta a}\delta_{\alpha b}$. Then

$$(V_x^\dagger V_x)_{\alpha\beta} = M_{ab}\delta_{\alpha b}\delta_{\beta b},$$

so that $\sum_{x\in S}(V_x^\dagger V_x)_{\alpha\beta} = \delta_{\alpha\beta}$ and we have a solution to the closure relation. Moreover

$$(V_x \rho V_x^\dagger)_{\alpha\beta} = M_{ab}\rho_{bb}\delta_{\alpha a}\delta_{\beta a},$$

and $\mathrm{Tr}_{\mathcal{H}} V_x \rho V_x^\dagger = M_{ab}\rho_{bb}$ so that the evolution if "x" is measured is

$$\rho_{\alpha\beta} \to \rho'_{\alpha\beta} = \delta_{\alpha a}\delta_{\beta a} \text{ with probability } M_{ab}\rho_{bb}.$$

This example shows again immediate purification: after one measurement, the density matrix is a pure state, but it is even a diagonal state!
The average evolution is

$$\rho_{\alpha\beta} \to \rho'_{\alpha\beta} = \delta_{\alpha\beta} \sum_{b\in R} M_{\alpha b}\rho_{bb}$$

which exhibits immediate de-coherence: all off diagonal coefficients are 0 after one step of evolution. Notice however that again the diagonal of the density matrix evolves in average just like the probability distribution for the Markov chain with transition kernel $M_{\alpha\beta}$.

These two examples of lifts of a classical Markovian evolution to a quantum Markovian evolution have a very different physics. In the first case, the V_a operators allow the transition from state a to any other state. In the second case the V_x operators allow only transition from state a to state b if $x = (a, b)$. Only physical arguments or constraints can decide in a concrete modeling problem which lift is most appropriate, and maybe none is.

– The simplest case, $\#S = 2$:
Our aim is to classify the solutions of the equation $V^\dagger V + V'^\dagger V' = \mathrm{Id}$ where we write simply Id for $\mathrm{Id}_{\mathcal{H}}$. We use the polar decomposition: there are unitary operators U, U' and positive operators L, L' such that $V = UL^{1/2}$ and $V' = U'L'^{1/2}$. The unitary operators do not contribute to the closure relation, which reduces to $L + L' = \mathrm{Id}_{\mathcal{H}}$. Thus L and L' commute and can be diagonalized simultaneously. Thus, the general solution to the closure relation when $\#S = 2$ is parameterized by two unitary operators and a positive operator L with all eigenvalues in $[0, 1]$, with $V = UL^{1/2}$ and $V' = U'(\mathrm{Id} - L)^{1/2}$.

Using the general results of the previous chapter and the explicit examples of this section, we are in position to study two important examples when $\dim \mathcal{H} = 2$ and $\#S = 2$.

7.2 Rabi oscillations observed via indirected measurements

The first one is a case when the system has an Hamiltonian internal evolution but also interacts with probes on which measurements are performed. We assume that the measurement are of non-demolition type in the orthonormal basis $|\alpha\rangle$, $\alpha = 1, 2$, but that these basis vectors are not eigenstates of the Hamiltonian evolution. To simplify the discussion we assume that we can decouple the periods when the system evolves under its internal Hamiltonian evolution and the periods of interactions with probes.

We also select the simplest matrices allowing to illustrate the phenomena we are interested in. More complicated matrices only make the algebra less transparent. For the internal Hamiltonian evolution par within a time step we take a rotation per time step of the form $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. The two V matrices corresponding to the non-demolition measurement are diagonal and we take them to be $\begin{pmatrix} \sqrt{p} & 0 \\ 0 & \sqrt{q} \end{pmatrix}$ and $\begin{pmatrix} \sqrt{1-p} & 0 \\ 0 & \sqrt{1-q} \end{pmatrix}$ with $p, q \in [0,1]$ (and $p \neq q$ to allow the indirect measurement to discriminate between the two states of the system).

An important feature of the evolution is purification: the basic observation is that in average $\sqrt{\det\rho}$ decreases during the evolution. Indeed $\sqrt{\det\rho}$ is homogenous of degree 1 so that the magic of compensations between probabilities and denominators works! If the one time step evolution is $\rho \to \rho'$, an elementary computation yields $\mathbb{E}\left(\sqrt{\det\rho'}\right) = (\sqrt{pq} + \sqrt{(1-p)(1-q)})\sqrt{\det\rho}$ and it is enough to check that for $p \neq q$ one has $\sqrt{pq} + \sqrt{(1-p)(1-q)} < 1$ (the Cauchy inequality again). Then the martingale convergence theorem implies that after many iterations the determinant of the density matrix is small, which for $\dim\mathcal{H} = 2$ is sufficient to conclude that ρ is close to a pure state.

Using the explicit formulæ for the Markovian evolution, it is easy to do accurate simulations.
The more p differs form q and the more the measurement is efficient. The larger θ is and the more the Hamiltonian evolution is rapid. However, we are interested in a situation when one time step has only a small effect on the system, i.e. we want to be close to the continuous time limit. This means we want p and q close and θ small, and the question of the different behaviors is the question of the relative importance of $p - q$ and θ. We content with pictures for the moment.

7.3 Thermal fluctuations observed via indirected measurements

The second one is when the system is in contact with a thermal bath but also interacts with probes on which measurements are performed. We assume that the measurement are of non-demolition type in the orthonormal basis $|\alpha\rangle$, $\alpha = 1, 2$ which is the basis of energy eigenstates. This allows to simplify the discussion be concentrating on the diagonal of the density matrix: at a fundamental level, the thermal noise induces tran-

sitions between the energy eigenstates, but we assume that in the time between the passage of two probes, the fluctuations of the thermal noise are averaged so that the evolution of the diagonal of ρ is described by a classical master equation for a Markov chain with the right equilibrium distribution. In the case $\dim \mathcal{H} = 2$, this is enough to fix the structure of the chain up to a time scale, leading to a stochastic matrix of the form $\begin{pmatrix} e^{-\tau} + u(1-e^{-\tau}) & u(1-e^{-\tau}) \\ (1-u)(1-e^{-\tau}) & 1-u(1-e^{-\tau}) \end{pmatrix}$ where $\frac{u}{1-u} = e^{-\frac{\Delta E}{kT}}$ with ΔE the gap between the two energy eigenstates. Note that the transition matrix is $\exp \tau \begin{pmatrix} u-1 & u \\ 1-u & -u \end{pmatrix}$. For the measurement, we use the same matrices as for the case of Rabi oscillations.

Note that this is a model for one of the cavity experiments. Again, the system is very easy to simulate.
To be close to the continuous time limit, we assume again that p and q are close, and also that τ is small. The question of the different behaviors is the question of the relative importance of $p - q$ and τ. We content with pictures for the moment.

7.4 Concluding remarks

It turns out that a precise description of the competition is easier when a continuous time limit is considered explicitly, and the derivation of the general continuous time equations is our next task. It the discrete time step ϵ goes to 0 and the number of evolution steps n to infinity keeping $t = n\epsilon$ finite, one of our main results will be that $p - q$ should scale like $\epsilon^{1/2}$. On the other hand, θ (Rabi oscillations) and τ scale like ϵ.

Let us simply stress that the qualitative features of the competition in the case of Rabi oscillations and in the case of thermal noise have deep similarities (as suggested by the pictures above), but one important difference which occurs in the scaling of parameters. This is due to the fact that the case of Rabi oscillation exhibits the quantum Zeno effect. Concretely, for thermal fluctuations, the average time between jumps scales like a constant when $|p - q[\ll \epsilon^{1/2}$, while for Rabi oscillations this average time explodes like the ratio $|p - q[\epsilon^{-1/2}$ and the Rabi frequency must be rescaled to recover a finite average time between jumps.