Chapter 5

Information, correlations, and more

5.1 Quantum information formulations

Quantum information science has experienced enormous developments during the last thirty years. I do not cover this wide and fascinating field in these notes, but shall only discuss briefly some relations with the question of the formalism. Indeed quantum information theory led to new points of view and to new uses and applications of quantum theory. This renewal is considered by some authors as a real change of paradigm, and referred to as "the second quantum revolution".

From what I know, the interest in the relations between Information Theory and Quantum Physics started really in the 1970's, from the confluence of several questions and new results. Let me quote a few.

— This period experienced a better understanding of the relations and of the conflicts between General Relativity and Quantum Mechanics: the theoretical discovery of the Bekenstein-Hawking quantum entropy for black holes, the black hole evaporation (information) paradox, the more general Unruh effect and quantum thermodynamical aspects of gravity and of events horizons (with more recently many developments in quantum gravity and string theories, such as "Holographic gravity", "Entropic Gravity", etc.).

— The ongoing discussions on the various interpretations of the quantum formalism, the meaning of quantum measurement processes, and whether a quantum state represent the "reality", or some "element of reality" on a quantum system, or simply the observer's information on the quantum system, experienced a revival through the development of the concept of decoherence.

— The 1970's saw of course the theoretical and experimental developments of quantum computing. A standard reference is the book by Nielsen and Chuang [NC10]. This field started from the realization that quantum entanglement and quantum correlations can be used as a resource for performing calculations and the transmission of information in a more efficient way than when using classical correlations with classical channels.

— This led for instance to the famous "It from Bit" idea (or aphorism) of J. A. Wheeler (see e.g. in [Zur90]) and others (see for instance the book by Deutsch [Deu97], or talks by Fuchs [Fuc01, Fuc02]). Roughly speaking this amounts to reverse the famous statement of Laudauer "Information is Physics" into "Physics is Information", and to state that Information is the good starting point to understand the nature of the physical world and of the physical laws.

This point of view that Quantum Physics has to be considered from the point of view of Information has been developed and advocated by several authors in the area of quantum gravity and quantum cosmology. Here I shall just mention some old or recent attempts to use this point of view to discuss the formalism of "standard" quantum physics, not taking into account the issues of quantum gravity.

In the quantum-information-inspired approaches a basic concept is that of "device", or "operation", which represents the most general manipulation on a quantum system. In a very oversimplified presentation [1], such a device is a "black box" with both a quantum input system A and quantum output system B, and with a set I of classical settings $i \in I$ and a set O of classical responses $o \in O$. The outputs do not need to be yes/no answers to a set of compatible quantum observables (orthonormal basis in the standard formalism) but may be more general (for instance associated to a POVM). The input and output systems A and B may be different, and may be multipartite systems, e.g. may consist in collections of independent subsystems $A = \underset{\alpha}{\cup} A_\alpha$, $B = \underset{\alpha}{\cup} B_\beta$.

This general concept of device encompass the standard concepts of *state* and of *effect*. A *state* corresponds to the preparation of a quantum system S in a definite state; there is no input $A = \emptyset$, the setting i specify the state, there is no response, and $B = S$ is the system. An *effect*, corresponds to a destructive measurement on a quantum system S; the input $A = S$ is the system, there are no output $B = \emptyset$, no settings i, and the response set O is the set of possible output measurements o. This concept of device contains also the general concept of a *quantum channel*; then $A = B$, there are no settings or responses.

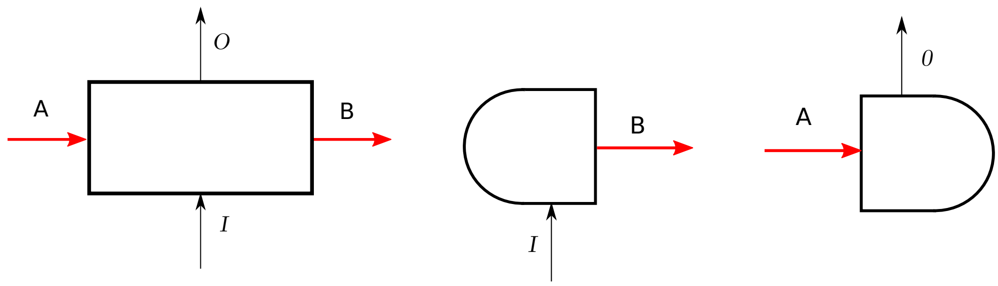

Figure 5.1: A general device, a state and an effect

Probabilities $p(i|o)$ are associated to the combination of a state and an effect, this correspond to the standard concept of probability of observing some outcome o when making a measurement on a quantum state (prepared according to i).

General information processing quantum devices are constructed by building causal circuits out of these devices used as building blocks, thus constructing complicated apparatus out of simple ones. An information theoretic formalism is obtained by choosing axioms on the properties of such devices (states and effects) and operational rules to combine these devices and circuits and the associated probabilities,

1. slightly more general than in some presentations

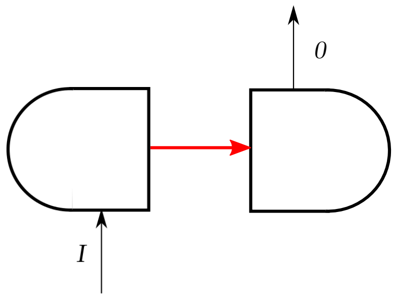

Figure 5.2: Probabilities are associated to a couple state-effect

thus obtaining for instance what is called in [CDP11] an "operational probabilistic theory". This framework is also called "generalized probabilistic theories", where "preparations", "transformations" and "measurements" replace the concepts of state, channel and effect.

This kind of approach is mainly considered for finite dimensional theories (which in the quantum case correspond to finite dimensional Hilbert spaces), both because it seems problematic to formulate them properly in the infinite dimensional case, and because systems with a finite number of distinct orthogonal pure states are those that are usually considered in quantum information science.

This approach leads to a pictorial formulation of quantum information processing. It shares similarities with the "quantum pictorialism" logic formalism, rather based on category theory, and presented for instance in [Coe10].

It can also be viewed as an operational and informational extension of the older "convex set approach" (developed notably by G. Ludwig, see [Lud85] and [Aul01],[BC81] for details). This later approach, which is also related to the quantum logic approach, puts more emphasis on the concept of states than on the concept of observables in quantum mechanics.

I shall not discuss much further these quantum information-theoretic approaches, since I di not know much about information theory and quantum information. Let me just highlight the recent attempts of Hardy [Har01, Har11] and those of Chiribella, D'Ariano & Perinotti [CDP10, CDP11] (see [Bru11] for a short presentation of this last formulation). Another proposal is that of Masanes and Müller [MM11].

In [CDP11] the standard complex Hilbert space formalism of QM is derived from six informational principles: Causality, Perfect Distinguishability, Ideal Compression, Local Distinguishability, Pure Conditioning and Purification Principle. The first two principles are not very different from the principles of other formulations (causality is defined in a standard sense, and distinguishability is related to the concept of differentiating states by measurements). The third one is related to the existence of reversible maximally efficient compression schemes for states. The four and the fifth are about the properties of bipartite states and for instance the possibility to performing local tomography and the effect of separate "atomic" measurements on such states. The last one, "purification principle", distinguishes quantum mechanics from classical mechanics, and states that any mixed state of some system $\mathcal{S}$ may be obtained from a pure state of a larger composite system $\mathcal{S}+\mathcal{S}'$. See [Bru11] for a discussion of the relation of this purification principle with the discussions of the "Heisenberg cut" between the system measured and the measurement device (see Heisenberg's 1935 reply to EPR in [CB]). Of course this is related to previous discussions by von Neumann in [vN32].

In [MM11] a set of five informational axioms on the properties of physics states allow to derive the formalism of quantum (and of classical) mechanics. This approach shares elements with [Har01, Har11], but the fifth axiom is simpler. Again the existence of reversible continuous transformations between states (axiom 4′) is a very important feature to obtain by geometrical argument the standard Hilbert space structure for (finite dimensional) quantum mechanics.

5.2 Quantum correlations

The world of quantum correlations is richer, more subtle and more interesting than the world of classical correlations. Most of the puzzling features and apparent paradoxes of quantum physics come from the properties of these correlations, especially for entangled states. Entanglement is probably *the* distinctive feature of quantum mechanics, and is a consequence of the superposition principle when considering quantum states for composite systems. Here I discuss briefly some basic aspects. Entanglement describes the particular quantum correlations between two quantum systems which (for instance after some interactions) are in a non separable pure state, so that each of them considered separately, is not in a pure state any more. Without going into history, let me remind that if the terminology "entanglement" ("Verschränkung") was introduced in the quantum context by E. Schrödinger in 1935 (when discussing the famous EPR paper). However the mathematical concept is older and goes back to the modern formulation of quantum mechanics. For instance, some peculiar features of entanglement and its consequences have been discussed already around the 30′ in relation with the theory of quantum measurement by Heisenberg, von Neumann, Mott, etc. Examples of interesting entangled many particles states are provided by the Stater determinant for many fermion states, by the famous Bethe ansatz for the ground state of the spin 1/2 quantum chain, etc.

5.2.1 Entropic inequalities

von Neumann entropy: The difference between classical and quantum correlations is already visible when considering the properties of the von Neumann entropy of states of composite systems. Remember that the von Neumann entropy of a mixed state of a system A, given by a density matrix ρ_A, is given by

$$S(\rho_A) = -\mathrm{tr}(\rho_A \log \rho_A) \tag{5.2.1}$$

In quantum statistical physics, the log is usually the natural logarithm

$$\log = \log_{\mathrm{e}} = \ln \tag{5.2.2}$$

while in quantum information, the log is taken to be the binary logarithm

$$\log = \log_2 \tag{5.2.3}$$

The entropy measures the amount of "lack of information" that we have on the state of the system. But in quantum physics, at variance with classical physics, one must be very careful about the meaning of "lack of information", since one cannot speak about the precise state of a system before making measurements. So the entropy could (and should) rather be viewed as a measure of the number of independent measurements we can make on the system before having extracted all the information, i.e. the amount of information we can extract of the system. It can be shown also that the entropy give the maximum information capacity of a quantum channel that we can build out of the system. See [NC10] for a good introduction to quantum information and in particular on entropy viewed from the information theory point of view.

When no ambiguity exists on the state ρ_A of the system A, I shall use the notations

$$S_A = S(A) = S(\rho_A) \tag{5.2.4}$$

The von Neumann entropy shares many properties of the classical entropy. It has the same convexity properties

$$S[\lambda\rho + (1-\lambda)\rho'] \geq \lambda S[\rho] + (1-\lambda)S[\rho'] \quad , \qquad 0 \leq \lambda \leq 1 \tag{5.2.5}$$

It is minimal $S = 0$ for systems in a pure state and maximal for systems in a equipartition state $S = \log(N)$ if $\rho = \frac{1}{N}\mathbf{1}_N$. It is extensive for systems in separate states.

Relative entropy: The relative entropy (of a state ρ w.r.t. another state σ for the same system) is defined as in classical statistics (Kullback-Leibler entropy) as

$$S(\rho\|\sigma) = \mathrm{tr}(\rho\log\rho) - \mathrm{tr}(\rho\log\sigma) \tag{5.2.6}$$

with the same convexity properties.

The differences with the classical entropy arise for composite systems. For such a system AB, composed of two subsystems A and B, a general mixed state is given by a density matrix ρ_{AB} on $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. The reduced density matrices for A and B are

$$\rho_A = \mathrm{tr}_B(\rho_{AB}) \quad , \qquad \rho_B = \mathrm{tr}_A(\rho_{AB}) \tag{5.2.7}$$

This corresponds to the notion of marginal distribution w.r.t. A and B of the general probability distribution of states for AB in classical statistics. Now if one considers

$$S(AB) = -\mathrm{tr}(\rho_{AB}\log\rho_{AB}) \quad , \qquad S(A) = -\mathrm{tr}(\rho_A\log\rho_A) \quad , \qquad S(B) = -\mathrm{tr}(\rho_B\log\rho_B) \tag{5.2.8}$$

one has the following definitions.

Conditional entropy: The *conditional entropy* S(A|B) (the entropy of A conditional to B in the composite system AB) is

$$S(A|B) = S(AB) - S(B) \tag{5.2.9}$$

The conditional entropy $S(A|B)$ corresponds to the remaining uncertainty (lack of information) on A if B is known.

Mutual information: The *mutual information* (shared by A and B in the composite system AB)

$$S(A:B) = S(A) + S(B) - S(AB) \tag{5.2.10}$$

Subadditivity: The entropy satisfies the general inequalities (triangular inequalities)

$$|S(A) - S(B)| \leq S(AB) \leq S(A) + S(B) \tag{5.2.11}$$

The rightmost inequality $S(AB) \leq S(A) + S(B)$ is already valid for classical systems, but the leftmost is quantum. Indeed for classical systems the classical entropy H_{cl} satisfy only the much stronger lower bound

$$\max(H_{\mathrm{cl}}(A), H_{\mathrm{cl}}(B)) \leq H_{\mathrm{cl}}(AB) \tag{5.2.12}$$

Subadditivity implies that if AB is in a pure entangled state, $S(A) = S(B)$. It also implies that the mutual information in a bipartite system is always positive

$$S(A:B) \geq 0 \tag{5.2.13}$$

In the classical case the conditional entropy is always positive $H_{\mathrm{cl}}(A|B) \geq 0$. In the quantum case the conditional entropy may be negative $S(A|B) < 0$ if the entanglement between A and B is large enough. This is a crucial feature of quantum mechanics. If $S(A|B) < 0$ it means that A and B share information resources (through entanglement) which get lost if one gets information on B only (through a measurement on B for instance).

Strong subadditivity: Let us consider a tripartite systems ABC. The entropy satisfies another very interesting inequality

$$S(A) + S(B) \leq S(AC) + S(BC) \tag{5.2.14}$$

It is equivalent to (this is the usual form)

$$S(ABC) + S(C) \leq S(AC) + S(BC) \tag{5.2.15}$$

Note that 5.2.14 is also true for the classical entropy, but then for simple reasons. In the quantum case it is a non trivial inequality.

The strong subadditivity inequality implies the triangle inequality for tripartite systems

$$S(AC) \leq S(AB) + S(BC) \tag{5.2.16}$$

so the entropic inequalities can be represented graphically as in fig. 5.3

The strong subadditivity inequality has important consequences for conditional entropy and mutual information (see [NC10]). Consider a tripartite composite system ABC. It implies for instance

$$S(C|A) + S(C|B) \geq 0 \tag{5.2.17}$$

and

$$S(A|BC) \leq S(A|B) \tag{5.2.18}$$

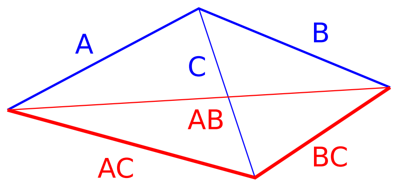

Figure 5.3: Entropic inequalities: the length of the line "X" is the von Neumann entropy $S(X)$. The tetrahedron has to be "oblate", the sum AC+BC (fat red lines) is always ≥ the sum A+B (fat blue lines).

which means that conditioning A to a part of the external subsystem (here C inside BC) increase the information we have on the system (here A). One has also for the mutual information

$$S(A:B) \leq S(A:BC) \tag{5.2.19}$$

This means that discarding a part of a multipartite quantum system (here C) increases the mutual information (here between A and the rest of the system). This last inequality is very important. It implies for instance that if one has a composite system AB, performing some quantum operation on B without touching to A cannot increase the mutual information between A and the rest of the system.

Let us mention other subadditivity inequalities for tri- or quadri-partite systems.

$$S(AB|CD) \leq S(A|C) + S(B|D) \tag{5.2.20}$$

$$S(AB|C) \leq S(A|C) + S(B|C) \tag{5.2.21}$$

$$S(A|BC) \leq S(A|B) + S(A|C) \tag{5.2.22}$$

5.2.2 Bipartite correlations

The specific properties of quantum correlations between two causally separated systems are known to disagree with what one would expect from a "classical picture" of quantum theory, where the quantum probabilistics features come just from some lack of knowledge of underlying "elements of reality". I shall come back later on the very serious problems with the "hidden variables" formulations of quantum mechanics. But let us discuss already some of the properties of these quantum correlations in the simple case of a bipartite system.

I shall discuss briefly one famous and important result: the Tsirelson bound. The general context is that of the discussion of non-locality issues and of Bell's [Bel64] and CHSH inequalities [CHSH69] in bipartite systems. However, since these last inequalities are more of relevance when discussing hidden variables models, I postpone their discussion to the next section 5.3.

This presentation is standard and simply taken from [Lal12].

5.2.2.a - The Tsirelson bound

The two spin system: Consider a simple bipartite system consisting of two spins 1/2, or q-bits 1 and 2. If two observers (Alice $\mathcal{A}$ and Bob $\mathcal{B}$) make independent measurements of respectively the value of the spin 1 along some direction $\vec{n}_1$ (a unit vector in

3D space) and of the spin 2 along $\vec{n}_2$, at each measurement they get results (with a correct normalization) $+1$ or -1. Now let us compare the results of four experiments, depending whether $\mathcal{A}$ choose to measure the spin 1 along a first direction $\vec{a}$ or a second direction $\vec{a}'$, and wether $\mathcal{B}$ chose (independently) to measure the spin 2 along a first direction $\vec{b}$ or a second direction $\vec{b}'$. Let us call the corresponding observables A, A', B, B', and by extension the results of the corresponding measurements in a single experiment A and A' for the first spin, B and B' for the second spin.

$$\text{spin 1 along } \vec{a} \quad \rightarrow \quad A = \pm 1 \quad ; \quad \text{spin 1 along } \vec{a}' \quad \rightarrow \quad A' = \pm 1 \tag{5.2.23}$$

$$\text{spin 2 along } \vec{b} \quad \rightarrow \quad B = \pm 1 \quad ; \quad \text{spin 2 along } \vec{b}' \quad \rightarrow \quad B' = \pm 1 \tag{5.2.24}$$

Now consider the following combination M of products of observables, hence of products of results of experiments

$$M = AB - AB' + A'B + A'B' \tag{5.2.25}$$

and consider the expectation value $\langle M \rangle_\psi$ of M for a given quantum state $|\psi\rangle$ of the two spins system. In practice this means that we prepare the spins in state $|\psi\rangle$, chose randomly (with equal probabilities) one of the four observables, and to test locality $\mathcal{A}$ and $\mathcal{B}$ may be causally deconnected, and choose independently (with equal probabilities) one of their own two observables, i.e. spin directions. Then they make their measurements. The experiment is repeated a large number of time and the right average combination M of the results of the measurements is calculated afterwards.

A simple explicit calculation shows the following inequality, known as the Tsirelson bound [Cir80]

Tsirelson bound: For any state and any choice orientations $\vec{a}$, $\vec{a}'$, $\vec{b}$ and $\vec{b}'$, one has

$$|\langle M \rangle| \leq 2\sqrt{2} \tag{5.2.26}$$

while, as discussed later, "classically", i.e. for theories where the correlations are described by contextually-local hidden variables attached to each subsystem, one has the famous Bell-CHSH bound

$$\langle |M| \rangle_{\text{“classical”}} \leq 2. \tag{5.2.27}$$

The Tsirelson bound is saturated if the state $|\psi\rangle$ for the two spin is the singlet

$$|\psi\rangle = |\text{singlet}\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle \right) \tag{5.2.28}$$

and the directions for $\vec{a}$, $\vec{a}'$, $\vec{b}$ and $\vec{b}'$ are coplanar, and such that $\vec{a} \perp \vec{a}'$, $\vec{b} \perp \vec{b}'$, and the angle between $\vec{a}$ and $\vec{b}$ is $\pi/4$, as depicted on 5.4.

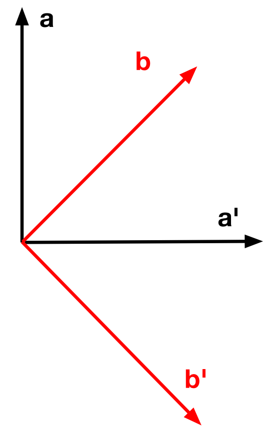

Figure 5.4: Spin directions for saturating the Tsirelson bound and maximal violation of the Bell-CHSH inequality

5.2.2.b - Popescu-Rohrlich boxes

Beyond the Tsirelson bound ? Interesting questions arise when one consider what could happen if there are "super-strong correlations" between the two spins (or in general between two subsystems) that violate the Tsirelson bound. Indeed, the only mathematical bound on M for general correlations is obviously $|\langle M\rangle| \leq 4$. Such hypothetical systems are considers in the theory of quantum information and are denoted Popescu-Rohrlich boxes [PR94] . With the notations of the previously considered 2 spin system, BR-boxes consist in a collection of probabilities $P(A,B|a,b)$ for the outputs A and B of the two subsystems, the input or settings a and b being fixed. The (a,b) correspond to the settings I and the (A,B) to the outputs O of fig. 5.1 of the quantum information section. In our case we can take for the first spin

$$a=1 \ \rightarrow \ \text{chose orientation } \vec{a} \quad , \qquad a=-1 \ \rightarrow \ \text{chose orientation } \vec{a}' \tag{5.2.29}$$

and for the second spin

$$b=1 \ \rightarrow \ \text{chose orientation } \vec{b} \quad , \qquad b=-1 \ \rightarrow \ \text{chose orientation } \vec{b}' \tag{5.2.30}$$

The possible outputs being always $A=\pm 1$ and $B=\pm 1$.

The fact that the $P(A,B|a,b)$ are probabilities means that

$$0 \leq P(A,B|a,b) \leq 1 \quad , \qquad \sum_{A,B} P(A,B|a,b) = 1 \quad \text{for} \quad a,b \text{ fixed} \tag{5.2.31}$$

Non signalling: If the settings a and b and the outputs A and B are relative to two causally separated parts of the system, corresponding to manipulations by two independent agents (Alice and Bob), enforcing causality means that Bob cannot guess which setting (a or a') Alice has chosen from his choice of setting (b and b') and his

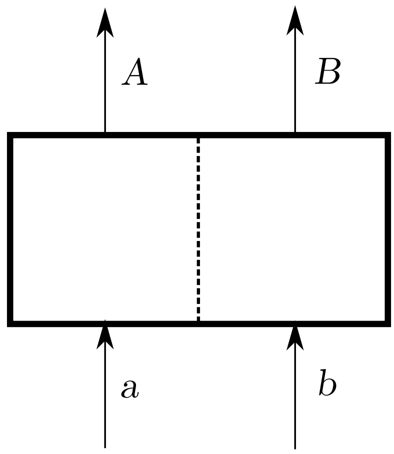

Figure 5.5: a Popescu-Rohrlich box

output (B or B), without knowing Alices′ output A. The same holds for Alice with respect to Bob. This requirement is enforced by the non-signaling conditions

$$\sum_A P(A,B|a,b) = \sum_A P(A,B|a',b) \tag{5.2.32}$$

$$\sum_B P(A,B|a,b) = \sum_B P(A,B|a,b') \tag{5.2.33}$$

A remarkable fact is that there are choices of probabilities which respect the non-signaling condition (hence causality) but violate the Tsirelson bound and even saturate the absolute bound $|\langle M\rangle| = 4$. Such hypothetical devices would allow to use "super-strong correlations" (also dubbed "super-quantum correlations") to manipulate and transmit information in a more efficient way that quantum systems (in the standard way of quantum information protocols, by sharing some initially prepared bipartite quantum system and exchanges of classical information) [BBL+06] [vD05] [PPK+09] . However, besides these very intriguing features of "trivial communication complexity", such devices are problematic. In particular it seems that no interesting dynamics can be defined on such systems [GMCD10].

Before concluding the discussion on the nature of quantum correlations, it is of course important to discuss their "non-local" nature and some of the issues related to the EPR-paradox and the attempts to explain theses correlations by hidden variables models. I discuss some of theses questions, in particular Bell inequalities, in the next section 5.3. A summary will be made afterwards in 5.4.

5.3 Hidden variables, contextuality and local realism

5.3.1 Hidden variables and "elements of reality"

In this section I discuss briefly some features of quantum correlations which are important when discussing the possibility that the quantum probabilities may still have, to some extent, a "classical interpretation" by reflecting our ignorance of inaccessible "sub-quantum" degrees of freedom or "elements of reality" of quantum systems, which could behave in a more classical and deterministic way. In particular a question is: which general constraints on the properties of such degrees of freedom are enforced by quantum mechanics?

This is the general idea of the "hidden variables" program and of the search of explicit hidden variable models. These ideas go back to the birth of quantum mechanics, and were in particular proposed by L. de Broglie in his first "pilot wave model", but they were dismissed by most physicist after the famous 1927 Solvay Congress and the tremendous advances of quantum physics in the 1930's. Despite the discussions that followed the EPR paper of 1935, hidden variables models underwent a revival only in the 1960's, from the works of D. Bohm, and mostly from the work of J. Bell, and the subsequent experimental developments that started in the 1970's.

Let me state first that I am not going to review in details or advocate hidden variable models. This section will illustrate the very strong constraints that the quantum formalism enforces on the basic idea. For more in-depth analysis and some incisive criticism see for instance [Per95] and [STr07].

In most discussions about realism and locality, see for instance [Per95] and [Lal12], the Bell's inequality and the issues of non-locality are discussed before the question of "contextuality" (the concept will be explained below). I choose in these notes to discuss first contextuality, since I think it is a more general issue that puts the other problems in perspective, andI shall discuss the Bell's inequalities and some hidden variable ideas after.

The basic idea of hidden variables is that the quantum states (the $|\psi\rangle$'s) of a quantum system $\mathcal{S}$ could be described by some ensembles of (partially or totally) hidden variables $\mathfrak{v}$. These hidden variables $\mathfrak{v}$ belong to some set (probability space) $\mathfrak{V}$, with some unknown statistics and possibly some unknown dynamics. Each $\mathfrak{v}$ (an element of the set $\mathfrak{V}$) represents an instance of a possibly infinite collection of these more fundamental sub-quantum variables. The specific outcome a (a real number) of a measurement operation of a physical observable A is supposed to be determined by the actual hidden variable $\mathfrak{v}$.

$$\text{observable } A \,+\, h.v.\ \mathfrak{v} \overset{\text{measurement}}{\longrightarrow} \text{outcome } a = \mathrm{function}(A,\mathfrak{v}) \tag{5.3.1}$$

In these schemes, the quantum undeterminism is not fundamental. It is assumed to come from our lack of knowledge about the exact state of the hidden variables. In other word, the pure quantum states $|\psi\rangle$ of the system should correspond to some classical probability distributions $p_\psi(\mathfrak{v})$ on $\mathcal{V}$. Note that the outcome of a measurement may depend not only on the choice of the observable that is measured, but also on the choice and the details of the measurement apparatus. This is usually denoted the "context" of the measurement. One might expect that, if the idea of hidden variable makes sense, the influence of the (partially unknown) context may be taken into account through the hidden variables themselves. This notion of "context of a measurement" and of "contextuality" is a very important one (it is basically the whole point in the discussion of hidden variables) and it will be discussed more in the next sections. Finally the measurement process should be amenable to a description in this framework, taking into effect the possible (deterministic) back reaction of a measurement operation on the the hidden variables $\mathfrak{v}$.

A very simple class of hidden variable models has been discussed by J. von Neumann in his 1932 book [vN32, vN55]. The hidden variables (the element of reality) are assumed to be in one to one correspondence with the possible outcomes a of all the observables A of the system. $\mathfrak{v}$ = {outcomes of observables}. Denoting by $f(A,\mathfrak{v})$ the well determined outcome a of a measurement of the observable A (one speaks of "dispersion free" variables) this function must then satisfy the addition law

$$C = A + B \quad \Longrightarrow \quad c = a + b \quad \text{i.e} \quad f(C,\mathfrak{v}) = f(A,\mathfrak{v}) + f(B,\mathfrak{v}) \tag{5.3.2}$$

As argued in [vN32, vN55], this is clearly inconsistent with quantum mechanics. Indeed if A and B do not commute, the possible outcomes of C (the eigenvalues of the operator C) are not in general sums of outcomes of A and B (sums of eigenvalues of A and B), since A and B do not have common eigenvectors. See [Bub10] for a detailed discussion of the argument and of its historical significance.

5.3.2 Context free hidden variables

More general hidden variable models have been rediscussed a few decades later, especially by and after J. Bell. Let me first present the class of models called (in the modern terminology) "context free" or "noncontextual" hidden variables models. The idea is that one should consider only the correlations between results of measurements on a given system for *sets of commuting observables*. Indeed only such measurements can be performed independently and in any possible order (on a single realization of the system), and without changing the statistics of the outcomes. Any such given set of observables can be thought as a set of classical observables, but of course this classical picture is not consistent from one set to another.

Thus the idea is still that a hidden variable $\mathfrak{v}$ assigns to any observable A a definite outcome $a = f(A,\mathfrak{v})$ as in 5.3.1. This assumption is often called "value definiteness" (VD). To be compatible with quantum mechanics, the outcome a must be one of the eigenvalues of the operator A. To a given pure quantum state ψ corresponds a probability distribution $p_\psi(\mathfrak{v})$ over the probability space $\mathfrak{V}$. It is such that the quantum expectation value of any observable corresponds to the probabilistic expectation of the corresponding outcome over the hidden variable distribution

$$\langle\psi|A|\psi\rangle = \int_{\mathfrak{V}} p_\psi(\mathfrak{v})\; f(A,\mathfrak{v}) \tag{5.3.3}$$

However the too strong constraint 5.3.2 should be replaced by the more realistic constraint, valid only for the set of outcomes $\{f(A,\mathfrak{v})\}$ for compatible observables (commuting operators)

$$\text{If } A_1 \text{ and } A_2 \text{ commute, then} \quad \begin{cases} f(A_1+A_2,\mathfrak{v}) = f(A_1,\mathfrak{v}) + f(A_2,\mathfrak{v}) \\ \qquad\qquad \text{and} \\ f(A_1A_2,\mathfrak{v}) = f(A_1,\mathfrak{v})f(A_2,\mathfrak{v}) \end{cases} \tag{5.3.4}$$

Moreover, these conditions are extended to any family $\mathcal{F} = \{A_i,\, i = 1, 2, \cdots\}$ of commuting operators.

The models are context-free or non contextual. The term "context free" means that the outcome a for the measurement of the first observable A_1 is supposed to be independent of the choice of the other compatible observable A_2. In other word, the outcome of a measurement depends only of the choice of observable and of the state of the hidden variable, but not of the "context" of the measurement, that is of the other quantities measured at the same time.

Here I presented purely deterministic hidden variable models. We shall discuss the possibility that a is a random variable (with a probability law fixed by $\mathfrak{v}$) later.

5.3.3 Gleason's theorem and contextuality

5.3.3.a - Gleason's theorem excludes general context-free models

These kind of models seems more realistic. However, they are immediatly excluded by Gleason's theorem [Gle57]. This was already observed and discussed by J. Bell in [Bel66]. Indeed, if to any $\mathfrak{v}$ is associated a function $f_\mathfrak{v}$, defined over the set of observables by

$$f_\mathfrak{v} \quad ; \qquad A \to f_\mathfrak{v}(A) = f(A, \mathfrak{v}) \tag{5.3.5}$$

which satisfy the consistency conditions 5.3.4, this condition is true in particular for any family of commuting projectors $\{P_i\}$, whose outcome in 0 or 1

$$P \text{ projector such that } P = P^\dagger = P^2 \quad \Longrightarrow \quad f_\mathfrak{v}(P) = 0 \text{ or } 1 \tag{5.3.6}$$

In particular, this is true for the family of projectors $\{P_i\}$ onto the vector of any orthonormal basis $\{\vec{e}_i\}$ of the Hilbert space $\mathcal{H}$ of the system. This means simply that defining the function f on the unit vectors $\vec{e}$ by

$$f(\vec{e}) = f_\mathfrak{v}(P_{\vec{e}}) \quad , \qquad P_{\vec{e}} = |\vec{e}\rangle\langle\vec{e}| \tag{5.3.7}$$

(remember that $\mathfrak{v}$ is considered fixed), this function must satisfy for any orthonormal basis

$$\{\vec{e}_i\} \quad \text{orthonormal basis} \qquad \Longrightarrow \qquad \sum_i f(\vec{e}_i) = 1 \tag{5.3.8}$$

while we have for any unit vector

$$f(\vec{e}) \;=\; 0 \text{ or } 1 \tag{5.3.9}$$

This contradicts strongly Gleason's theorem (see 4.4.2), as soon as the Hilbert space of the system $\mathcal{H}$ has dimension $\dim(\mathcal{H}) \geq 3$. Indeed, 5.3.8 means that the function f is a frame function (in the sense of Gleason), hence is continuous, while 5.3.9 (following from the fact that f is function on the projectors) means that f cannot be a continuous function. So

$$\dim(\mathcal{H}) \geq 3 \quad \Longrightarrow \quad \begin{array}{c}\text{no context-free deterministic hidden variable model}\\ \text{is compatible with quantum mechanics}\end{array} \tag{5.3.10}$$

5.3.3.b - The special case of $n = 2$

Gleason's theorem does not apply to the case $\dim(\mathcal{H}) = n = 2$. It is in fact quite easy to construct a context-free hidden variable model that describes all the observable of the 2-level system (the q-Bit). Here is a variant of a model by J. Bell. Any pure state $|\psi\rangle$ can be represented as a vector $\vec{n} = (\sin(\theta)\cos(\varphi), \sin(\theta)\sin(\varphi), \cos(\theta))$ of the unit 2-sphere in 3 dimensions through the Bloch sphere representation

$$|\vec{n}\rangle = \cos(\theta/2)|\uparrow\rangle + \sin(\theta/2)e^{i\varphi}|\downarrow\rangle \tag{5.3.11}$$

The algebra of observables is generated by the Pauli matrices and any self-adjoint physical operator A can be written as

$$A = \alpha\,\mathbf{1} + \vec{\beta}\cdot\vec{\sigma} \quad , \qquad \vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z) \tag{5.3.12}$$

$\vec{\beta}\cdot\vec{\sigma}$ is the traceless part of the operator A, with $\vec{\beta} = (\beta_x, \beta_y, \beta_z)$ a real vector. One has

$$\langle\vec{n}|\vec{\beta}\cdot\vec{\sigma}|\vec{n}\rangle = \vec{\beta}\cdot\vec{n} \tag{5.3.13}$$

so that the eigenvalues and eigenvectors of A are

$$a_\pm = \alpha \pm |\vec{\beta}| \quad , \qquad |\psi_\pm\rangle = |\vec{n}_\pm\rangle \quad , \qquad \vec{n}_\pm = \pm\vec{\beta}/|\vec{\beta}| \tag{5.3.14}$$

One can take as hidden variables the unit vectors $\vec{v}$

$$\mathfrak{V} = \mathcal{S}_2 \quad , \qquad \mathfrak{v} = \vec{v} \qquad |\vec{v}| = 1 \tag{5.3.15}$$

with the outcomes

$$f(A, \vec{v}) = \begin{cases} \alpha + |\vec{\beta}| & \text{if } \vec{\beta}\cdot\vec{v} \geq 0, \\ \alpha - |\vec{\beta}| & \text{if } \vec{\beta}\cdot\vec{v} < 0. \end{cases} \tag{5.3.16}$$

and for the probability distribution associated to the pure quantum state $|\vec{n}\rangle$ the distribution on the sphere with support on the

$$p_{\vec{n}}(\vec{v}) = \frac{1}{\pi} \begin{cases} \vec{n}\cdot\vec{v} & \text{if } \vec{n}\cdot\vec{v} \geq 0, \\ 0 & \text{if } \vec{n}\cdot\vec{v} < 0. \end{cases} \tag{5.3.17}$$

5.3.3.c - Probabilistic models = quantum mechanics

One may consider also partially deterministic hidden variable models, where to a hidden variable instance $\mathfrak{v}$ is attached a probability law $p(\ ;\mathfrak{v})$ defined as the probability that the outcome of a measurement of the observable A is a

$$\mathfrak{v} \ \rightarrow\ p(a|A;\mathfrak{v}) \ =\ \text{probability that measurement of } A \rightarrow a \tag{5.3.18}$$

so that the expectation value of A with respect to the law $p(\ ;\mathfrak{v})$ and the quantum expectation value are

$$\mathbb{E}[A|\mathfrak{v}] = \int da\ a\ p(a|A;\mathfrak{v}) \quad , \qquad \langle\psi|A|\psi\rangle = \int_{\mathfrak{V}} p_\psi(\mathfrak{v})\ \mathbb{E}[A|\mathfrak{v}] \tag{5.3.19}$$

For such a model to be context-free, it must satisfy $\mathbb{E}[F(A)|\mathfrak{v}] = \int da\;\; F(a)\;\, p(a|A;\mathfrak{v})$. Then using Gleason's theorem again, it is easy to shown that the only consistent and minimal realization is to take for hidden variable the state vector itself (therefore there is no sub-quantum classical indeterminism), and for the probability law the law that is given by the Born rule itself (full quantum indeterminism)

$$\mathfrak{v} = \psi \quad , \qquad p(a|A;\psi) = \langle\psi|\delta(a\mathbf{1} - A)|\psi\rangle \tag{5.3.20}$$

Gleason's theorem is a very serious problem for the idea of hidden variables. It excludes the hypothesis that the values of all the possible observables of a quantum system are unknown to us but preexist the act of observation. Such a concept is (I think) often called the "strong realism hypothesis".

However, some remaining possibilities may still be considered, that correspond to a weaker notion of realism, for instance:

1. There are still context-free hidden variables, but they describe only some specific subset of the quantum correlations, not all of them.
2. There are hidden variables, but they are fully contextual.

I now discuss two famous cases where the first option has been explored, but appears to be still problematic. The second option raises also very serious questions, that will be shortly discussed in 5.4.

5.3.4 The Kochen-Specker theorem

The first option is related to the idea that some subsets of the correlations of a quantum system have a special status, being related to some special explicit "elements of reality" (the "be-ables" or "maybe-ables" in the terminology of J. Bell), in contrast to the ordinary observables which are just "observ-ables". Thus a question is whether for a given quantum system there are some finite families of non commuting observables which can be associated to context-free hidden variables.

In fact the problems with non-contextual hidden variable models have been shown to arise already for very small such subsets of observables, first by S. Kochen and E. Specker [KS67] and by J. Bell in [Bel66]. This is the content of the Kochen-Specker theorem. This theorem provides in fact examples of *finite families* of unit vectors $\mathcal{E} = \{\vec{e}_i\}$ in a Hilbert Space $\mathcal{H}$ (over $\mathbb{R}$ or $\mathbb{C}$) of finite dimension ($\dim(\mathcal{H}) = n$), such that it is impossible to find any frame function such that

$$f(\vec{e}_i) = 0 \text{ or } 1 \quad \text{and} \quad (\vec{e}_{i_1}, \cdots, \vec{e}_{i_n}) \quad \text{orthonormal basis} \quad \Longrightarrow \quad \sum_{a=1}^{n} f(\vec{e}_{i_a}) = 1 \tag{5.3.21}$$

The original example of [KS67] involves a set with 117 projectors in a 3 dimensional Hilbert space, that generates the group of symmetry of a 3d polyhedra and is a very nice example of non-trivial 3d geometry. Simpler examples in dimension $n = 3$ and $n = 4$ with a smaller number of projectors have been provided by several authors (Mermin, Babello, Peres, Penrose). I refer to [Lal12] for details and I shall not discuss more these examples and their significance. But all these examples show that the non–contextual character of quantum correlations is a fundamental feature of quantum mechanics.

5.3.5 The Bell-CHSH inequalities and local realism

5.3.5.a - The local realism hypothesis

Another important example, where the relations between contextuality and locality are discussed, is the situation of the so-called Einstein-Poldovski-Rosen (EPR) paradox. It was first considered by J. Bell in in his famous 1964 paper [Bel64]. Consider a bipartite system $\mathcal{S}$ that consists of two causally independent subsystems $\mathcal{S}_1$ and $\mathcal{S}_2$, for instance a pair of time-like separated photons in a Bell-like experiment. We are interested in the correlations between the result of independent measurements on $\mathcal{S}_1$ and $\mathcal{S}_2$. If A is some observable for the system $\mathcal{S}_1$ and B some observable for the system $\mathcal{S}_2$, this pair of observables is compatible, since A and B, (or more exactly $A\otimes 1$ and $1\otimes B$) commute. Thus in quantum mechanics the result of a measurement on $\mathcal{S}_1$ should depend on the state of the whole system $\mathcal{S}$ and of the choice of the observable A, but it should not depend on the measurement made on $\mathcal{S}_2$. In other word, the result of a measurement on $\mathcal{S}_1$ might depend on the local context of $\mathcal{S}_1$ but it should not depend on the context of $\mathcal{S}_2$.

Following J. Bell, let us assume that some hidden variables model (some "elements of reality") underly the bipartite system $\mathcal{S}$, and that, in the spirit of the argument by EPR, this model is local in the sense that it is $\mathcal{S}_1$-versus-$\mathcal{S}_2$ context free. This means that the "sub-quantum" state of the whole system $\mathcal{S}$ is assumed to be described by some hidden variable $\mathfrak{v}$. The result a of the measurement of A on $\mathcal{S}_1$ is determined (or obeys a probabilistic law) that depends on the hidden variable $\mathfrak{v}$, on the observable A chosen, and possibly of the local context of the measurement on $\mathcal{S}_1$, but *not on the local context of the measurement done on $\mathcal{S}_2$.* Similarly, the result b of the measurement of B on $\mathcal{S}_2$ is depends on the hidden variable $\mathfrak{v}$, on the observable B chosen, and possibly of the local context for $\mathcal{S}_2$, but *not on the local context for $\mathcal{S}_1$.*

For a purely deterministic (dispersion free) hidden variable model with such a constraint of locality, this means that a hidden variable $\mathfrak{v}$ assigns determined outputs a and b to the measurements of A and B (respectively on $\mathcal{S}_1$ and $\mathcal{S}_2$)

$$\mathfrak{v} \qquad \to \qquad a = f_1(A,\mathfrak{v})\,,\ b = f_2(B,\mathfrak{v}) \tag{5.3.22}$$

For a partially deterministic local hidden variable model, a hidden variable $\mathfrak{v}$ assigns *independent probability distributions* for the outcomes a and b of the measurements of A and B (respectively on $\mathcal{S}_1$ and $\mathcal{S}_2$)

$$\mathfrak{v} \qquad \to \qquad p_1(a|A;\mathfrak{v})\,,\ p_2(b|B;\mathfrak{v}) \tag{5.3.23}$$

Since in general we have seen that on a single subsystem no context-free hidden variable model is compatible with quantum mechanics, in general A and B must be understood as

$$A \ = \ \text{measured observable} + \text{local context for } \mathcal{S}_1 \tag{5.3.24}$$

$$B \ = \ \text{measured observable} + \text{local context for } \mathcal{S}_2 \tag{5.3.25}$$

A notable exception (in fact the only one) is the situation where the subsystems $\mathcal{S}_1$ and $\mathcal{S}_2$ are 2-level quantum systems (qBits, spins 1/2, photons with two polarization states. In this case we may forget about the local contexts and take 5.3.22 at face value. They will not change the argument leading to the Bell-CHSH inequalities anyway.

In any case, hidden variables models for a bipartite system that assign outcome to separate measurements on the two subparts according to 5.3.22 or 5.3.23 are denoted usually "local hidden variable models". I tend to find this denomination a bit misleading, since for such models the "hidden variables" are in general still contextual. The denomination "local" indicates that the output of a measurement is assumed to depend on the hidden variable only through the local context of this measurement. This is the hypothesis of "local realism". One might perhaps rather call it the hypothesis of "local contextuality", since it defines "locally-contextual-only hidden variables models", or "hidden variables models that satisfy contextual local realism". Since the denominations "local realism" and "local hidden variable models" are standard I shall use them.

5.3.5.b - The Bell-CSHS inequality

Let me now recall the derivation of the famous Bell-CHSH inequality. In a general local hidden variable model, a quantum state ψ of $\mathcal{S}$ corresponds to some probability distribution $q_\psi(\mathfrak{v})$ over the hidden variables $\mathfrak{v}$. $q_\psi(\mathfrak{v})$ represent our ignorance about the "elements of reality" of the system. If this description is correct, the probability for the pair of outcomes $(A,B)\to(a,b)$ in the state ψ is given by the famous representation

$$p_\psi(a,b|A,B)=\sum_{\mathfrak{v}} q_\psi(\mathfrak{v})\,p_1(a|A;\mathfrak{v})\,p_2(b|B;\mathfrak{v})) \tag{5.3.26}$$

with p_1 and p_2 the outcome probabilities as in 5.3.23 It is this peculiar form which implies the famous Bell and BHSH inequalities on the correlations between observables on the two causally independent subsystems. If we consider for observables for $\mathcal{S}_1$ (respectively $\mathcal{S}_2$) two (not necessarily commuting) projectors P_1 and P'_1 (respectively Q_2 and Q'_2), with outcome 0 or 1, and take for observables

$$A=2P_1-1 \quad A'=2P'_1-1 \quad B=2Q_1-1 \quad B'=2Q'_1-1 \tag{5.3.27}$$

The outcomes a, a', b and b' are -1 or 1. Let us (the experimentalist) perform a series of experiments on an ensemble of independently prepared instances of the bipartite system $\mathcal{S}$, choosing randomly with equal probabilities 1/4 to measure (A,B), (A',B), (A,B') or (A',B'), and combine the results to compute the average

$$\langle M\rangle=\langle AB\rangle-\langle AB'\rangle+\langle A'B\rangle+\langle A'B'\rangle \tag{5.3.28}$$

The same argument than the argument used in 5.2.2 when discussing the Tsirelson's bound, using the general inequality

$$a,a',b,b'\in[-1,1] \quad\Longrightarrow\quad a(b-b')+a'(b+b')\in[-2,2] \tag{5.3.29}$$

and the fact that for the hidden variable model the outcomes a, a', b and b' are a priori well defined for each instance of $\mathfrak{v}$ so that one can use 5.3.26 implies the Bell-CSHS inequality

$$-2\le\langle M\rangle\le 2 \tag{5.3.30}$$

For a bipartite quantum mechanical state that consists of two q-Bits, this inequality is known to be violated for some simple quantum states, in particular the fully entangled single state $|\psi\rangle = (1/\sqrt{2})(|1\rangle\otimes|0\rangle - |0\rangle\otimes|1\rangle)$ and some adequate choice of observables. Indeed $\langle M\rangle$ is known to saturate the Tsirelon's bound $|\langle M\rangle| \leq 2\sqrt{2}$ for such states.

From the quantum mechanics point of view, the reason for the violation of the bound 5.3.30 is simple. Assuming that all quantum states give probabilities of the form expected from the general local hidden variable model 5.3.26 and that the subsystems probabilities $p_1(a|A;\mathfrak{v})$ and $p_2(b|B;\mathfrak{v})$ obey the quantum rules and are representable by density matrices for the subsystems (this is the maximal assumption allowed by local contextuality) means that any quantum state (mixed or pure) ω of the whole system could be represented by a density matrix of the form

$$\rho_\omega = \sum_{\mathfrak{v}} q_\omega(\mathfrak{v})\,(\rho_1(\mathfrak{v})\otimes\rho_2(\mathfrak{v})) \tag{5.3.31}$$

where $q_\omega(\mathfrak{v}) \geq 0$ is a density probability over $\mathfrak{V}$ and $\rho_1(\mathfrak{v})$ and $\rho_2(\mathfrak{v})$ density matrices relative to the subsystems $\mathcal{S}_1$ and $\mathcal{S}_2$. Such mixed states for a bipartite system are called separable states. But its is well known that not all mixed states of a bipartite systems are separable, in particular pure entangled states are not separable. In fact finding a criteria for characterizing separate states of a general bipartite or multipartite quantum system is a very interesting problem in quantum information.

I am not going to discuss the many and very interesting generalizations and variants of Bell inequalities (for instance the spectacular GHZ example for tripartite systems) and the possible consequences and tests of contextuality. I will not review either all the experimental tests of violations of Bell-like inequalities in various contexts, starting from the first experiments by Clauser, and those by Aspect et al., up to the most recent ones, that have basically closed all the possible loopholes in the hypothesis leading to the Bell-like inequalities. All the experimental results are in full agreement with the predictions of standard Quantum Mechanics and more precisely of Quantum Electro Dynamics. See for instance [Lal12] for a recent and very complete review.

5.3.6 Contextual models

Let me discuss briefly the last possibility: relax completely the requirement of non-contextuality and consider fully contextual hidden variable models.

5.3.6.a - Bell's simple model

Full contextuality for a hidden variable model amounts to assume that the output a for the measurement of an observable A is determined by the full context of the measurement, i.e. all the other compatible measurements and operations that are performed independently on the system. Let me consider the simple case of a finite dimensional Hilbert space. One can reduce the situation to the case where all these operations are ideal projective measurements, and thus consider as "the context" a complete family [2] of compatible projectors $\mathcal{P} = \{P_i\}_{i=1,K}$ onto orthogonal subspaces $\mathcal{E}_i$

2. In some discussions the context may include additional parameters (not included into the hidden variables) like the details of the apparatus used, etc.

so that

$$P_i P_j = \delta_{ij} P_i \quad , \quad \mathbf{1} = \sum_i P_i \quad \text{i.e.} \quad \mathcal{H} = \bigoplus_i \mathcal{E}_i \quad , \quad \mathcal{E}_i \perp \mathcal{E}_j \tag{5.3.32}$$

The projectors are not necessarily of rank 1 (projector onto pure states) so that $K \leq n = \dim(\mathcal{H})$.

$$\text{context} \ := \ \{P_j\}_{j=1,K} = \boldsymbol{\mathcal{P}} \tag{5.3.33}$$

Then the output $a_i = 0,1$ of a measurement of some $P_i \in \boldsymbol{\mathcal{P}}$ is determined by the hidden variable $\mathfrak{v}$ through a function that depends on all the P_j's, not only on P_i.

$$a_i = f(P_i|\{P_j\};\mathfrak{v}) = f_i(\boldsymbol{\mathcal{P}};\mathfrak{v}) \tag{5.3.34}$$

The only constraint is that

$$f_i(\boldsymbol{\mathcal{P}};\mathfrak{v}) = 0,1 \quad , \qquad \sum_i f_i(\boldsymbol{\mathcal{P}};\mathfrak{v}) = 1 \tag{5.3.35}$$

and that, if a quantum pure state $|\psi\rangle$ corresponds to a hidden variable distribution p_ψ, one has

$$\langle\psi|P_i|\psi\rangle = \int_{\mathfrak{V}} p_\psi(\mathfrak{v})\ f_i(\boldsymbol{\mathcal{P}};\mathfrak{v}) \tag{5.3.36}$$

The l.h.s. corresponds to the Born rule and should be independent on the context, i.e. of the choice of the others P_j's, $j \neq i$.

As noticed by J. Bell in [Bel66], this is trivial to satisfy. Just take as hidden variable the quantum state vector itself, $|\psi\rangle$, plus a uniformly distributed real random variable $X \in [0,1]$.

$$\mathfrak{v} = (|\psi\rangle, X) \tag{5.3.37}$$

From the probabilities $q_j = q_j(\boldsymbol{\mathcal{P}};|\psi\rangle) = \langle\psi|P_j|\psi\rangle$ of getting $a_j = 1$ in the state $|\psi\rangle$ (given by Born's rule), one can decompose the interval $[0,1[$ into the successive disjoint intervals $\mathcal{I}_| = [X_{j-1}, X_j[$ with $X_j = \sum\limits_{k=1,j} q_k$. The output functions are defined as

$$a_i = f_i(\boldsymbol{\mathcal{P}};\mathfrak{v}) = \begin{cases} 1 & \text{if } X \in \mathcal{I}_i = [X_{i-1}, X_i[, \\ 0 & \text{otherwise.} \end{cases} \tag{5.3.38}$$

and the probability distribution for a quantum state $|\phi\rangle$ is the Dirac distribution on the space of quantum states (the projective hyperplane in $\mathcal{H}$) times the uniform distribution over $[0,1[$ for X

$$\int_{\mathfrak{V}} p_{|\phi\rangle}(\mathfrak{v}) = \int_{|\psi\rangle} \delta(|\psi\rangle, |\phi\rangle) \int_0^1 dX \tag{5.3.39}$$

Thus one reobtains easily the standard output quantum probabilities

$$p_i = \langle\phi|P_i|\phi\rangle \tag{5.3.40}$$

from the principle "Born rule in, Born rule out".

This model is very simple indeed, and was considered by J. Bell as too trivial and not physical. Indeed it amounts to standard quantum mechanics itself. The random variable X has no particular physical or "'ontological" status. It is just introduced to give an explicit representation of the quantum probabilities as probabilities on an abstract classical probability space. In a mathematical langage,X implements the Kolmogorov extension theorem, which states that any abstract probabilistic process – provided it satisfies some obvious self-consistency conditions – can be represented on some explicit probability space Ω. But on the other hand this model is so generic (well, one has to extend it to infinite dimensional Hilbert spaces, see below) that one may suspect that any contextual hidden variable model might be reduced to this one, if one wants to keep it fully compatible with quantum mechanics.

5.3.6.b - The de Broglie-Bohm pilot-wave model

This is a more elaborate model which contains the dynamics of the Schrôodinger equation. It was first proposed by L. de Broglie, and rediscovered and developed by D. Bohm. Let me recall the model for the simple and standard example of a non relativistic particle in a scalar potential U. In position space and in the Schrödinger picture a pure quantum state $|\psi\rangle$ is given by the wave function $\psi = \psi(\vec{q}, t)$. The basic idea is that the continuity equation for the probability density ρ and the probability current $\vec{j}$

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0 \tag{5.3.41}$$

with

$$\rho = |\psi|^2 \quad , \qquad \vec{j} = \frac{\mathrm{i}\hbar}{2m}(\psi\vec{\nabla}\bar{\psi} - \bar{\psi}\vec{\nabla}\psi) \tag{5.3.42}$$

can be rewritten as the flow equation for the density distribution $\tilde{\rho} = \tilde{\rho}(\vec{q}, t)$ of particles driven by a vector field $\vec{V} = \vec{V}(\vec{q}, t)$ that derives from the wave function itself

$$\frac{\partial \tilde{\rho}}{\partial t} + \vec{\nabla}(\vec{V}\tilde{\rho}) \qquad \text{with} \qquad \vec{V} = \frac{\vec{j}}{\rho} = \frac{\hbar}{m}\vec{\nabla}(\arg(\psi)) = \frac{\mathrm{i}\hbar}{2m}\left(\frac{\vec{\nabla}\bar{\psi}}{\bar{\psi}} - \frac{\vec{\nabla}\psi}{\psi}\right) \tag{5.3.43}$$

Thus a hidden variable model consists in taking as sub-quantum degrees of freedom (hidden variable) the wave function itself $\psi = \psi(\vec{q}, t)$ (the "pilot wave") and the position vector $\vec{x} = \vec{x}(t)$ of a "real" particle, that will be the observable degree of freedom.

$$\mathfrak{v} = (\psi, \vec{x}) \tag{5.3.44}$$

The particle behaves as a passive scalar carried along the time dependent flow $\vec{V}$. The vector flow $\vec{V}$ is given by the gradient of the phase of the wave function ψ, that obeys the Schrödinger equation ($U = U(\vec{q})$ is the external scalar potential)

$$\mathrm{i}\hbar\frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi + U\psi \tag{5.3.45}$$

The dynamics for $\vec{x}$ is

$$\frac{d\vec{x}}{dt} = \vec{V}(\vec{x}, t) \tag{5.3.46}$$

and the initial probability distribution $\tilde{\rho}_\psi(\vec{x}, t=0)$ for the particle in position space, that corresponds to the ensemble that describe a quantum state $|\psi\rangle$, has to be taken equal to the probability density function given by the Born rule

$$\tilde{\rho}_\psi(\vec{x},t) = |\psi(\vec{x},t)|^2 \qquad \text{at} \qquad t=0 \tag{5.3.47}$$

It is easy to see that this relation (relating the probability distribution for the particle $\vec{x}$ to the pilot wave function ψ) stays valid at all times t, if the dynamics of the hidden variable $\mathfrak{v}$ is given by the Schrödinger equation 5.3.45 for ψ and the pilot wave equation 5.3.46 for $\vec{x}$.

With these notations for a contextual hidden variable theory, the de Broglie-Bohm theory describes the observables of position of the quantum particle. Thus the "context" is the set of projections over position eigenstates

$$\text{Context} \;=\; \mathcal{Q} = \{P_{\vec{q}} = |\vec{q}\rangle\langle\vec{q}|;\ \vec{q}\in\mathbb{R}^d\} \tag{5.3.48}$$

The output rule is (the hidden variable being $\mathfrak{v} = (\psi, \vec{x})$)

$$f(P_{\vec{q}}|\mathcal{Q};\mathfrak{v}) = \delta(\vec{x}-\vec{q}) \tag{5.3.49}$$

and the probability distribution associated to a quantum state $|\phi\rangle$ is

$$p_\phi(\mathfrak{v}) = p_\phi(\psi,\vec{x}) = \delta(\phi,\psi)\,|\psi(\vec{x})|^2 \tag{5.3.50}$$

One has indeed for any function A of the position operator $\vec{Q}$

$$\langle\phi|A(\vec{Q})|\phi\rangle = \int_{\psi,\vec{x}} p_\phi(\psi,\vec{x})A(\vec{x}) \tag{5.3.51}$$

The interesting property of the model is that its dynamics reproduces the average quantum dynamics. This means that the expectation value of the momentum operator P on a state $|\phi\rangle$ is given by the probabilistic average of the classical momentum $\vec{p} = m\dot{\vec{x}}$ of the "classical" particle

$$\langle\phi|\vec{P}|\phi\rangle = m\int_{\psi,\vec{x}} p_\phi(\psi,\vec{x})\,\dot{\vec{x}} \tag{5.3.52}$$

The model can be trivially extended to more that one particle. Then the pilot wave function $\psi(\vec{x}_1,\vec{x}_2,\cdots)$ depends on the position vectors of all the particles, and takes into account the non-local correlations between the positions of the "classical" particles resulting from entanglement.

For a single particle in d=1 dimension, the de Broglie-Bohm model is in fact equivalent to the simple Bell's model discussed in the previous section. Indeed instead of the hidden variable x (the position of the particle), one can consider the variable X

$$X = \int_{-\infty}^{x} dy\,|\psi(y)|^2 \tag{5.3.53}$$

and switch to the hidden variables $\mathfrak{u}$

$$\mathfrak{v} = (\psi, x) \quad \to \quad \mathfrak{u} = (\psi, X) \tag{5.3.54}$$

For any quantum state the variable X is now uniformly distributed on the interval $[0,1]$ and it has no dynamics

$$p_\phi(\psi,X) = \delta(\phi,\psi)\,1 \quad , \qquad \dot{X} = 0 \tag{5.3.55}$$

5.3.6.c - Stochastic models and Nelson stochastic mechanics

A question often discussed for these models is whether there is a good reason to take as initial distribution $\tilde{\rho}_\phi$ for the position $\vec{x}$ the quantum probability distribution $|\phi|^2$. Indeed, the evolution process for ψ and $\vec{x}$ and the evolution equation 5.3.43 are valid for $\tilde{\rho} \neq |\phi|^2$ as well.

One solution is to modify the evolution equation for $\vec{x}$ by adding some randomness, so that the general evolution equation for $\tilde{\rho}$ contains a diffusion term and a drift term, in such a way that any initial distribution $\tilde{\rho}$ relax irreversibly towards the quantum probability distribution $\tilde{\rho}_\phi = |\phi|^2$ at large time. The asymptotic equilibrium dynamics at large time, although non-deterministic, is physically indistinguishable for the deterministic model, as far as position observables are concerned.

In such models, the evolution equation for x becomes a stochastic equation of the form

$$dx = V\,dt + D\,dt + \nu\,dB_t \tag{5.3.56}$$

where V is the driving term of the original model, D is a drift term, dB_t a Brownian process, and ν a diffusion constant (it may depend on x). The drift term D is adjusted to the diffusion term so that the evolution equation takes the form (for ν independent of x and t)

$$\frac{\partial\tilde{\rho}}{\partial t} + \vec{\nabla}(\vec{V}_\phi\,\tilde{\rho}) + \nu\,\rho_\phi\,\Delta(\tilde{\rho}/\rho_\phi) \tag{5.3.57}$$

It is clear that the dynamics is now irreversible, and that it should relax towards the distribution where the diffusion term $\Delta(\tilde{\rho}/\rho_\phi)$ vanishes, namely the quantum probability distribution $\tilde{\rho} = \rho_\phi = |\phi|^2$.

These kind of models have been proposed by D. Bohm and his collaborators, who looked for physical models where the dynamics of the particles and the diffusion process was the result of some non-trivial microscopic sub-quantum dynamics, resulting in short time relaxation towards the equilibrium quantum distribution.

A particular case is the so-called *stochastic quantum mechanics* proposed by E. Nelson [Nel66]. It amounts to a stochastic dynamics of the form 5.3.56, with a special choice for the diffusion coefficient

$$\nu = \frac{\hbar}{2m} \tag{5.3.58}$$

This particular choice has some advantages. For a free particle (external potential $U = 0$), the dynamics of the particle is a Brownian process with some characteristics and statistics similar to those of the trajectories in the Feynman path integral formulation. The momenta of the particle $p = m\dot{q}$ becomes a stochastic variable with short time correlations of a white noise, with the right variance in order to reproduce as a statistical law the Heisenerg uncertainty relations, etc.

Other mechanisms that might produce relaxation of a general distribution function towards the quantum one have been suggested, based on the idea of coarse graining over microscopic sub-quantum degrees of freedom and of the notion of effective wave functions. See [Val91a, Val91b] for a simple version, and [DuGZ92] for another more detailed formulation and for details.

5.3.6.d - An adiabatic argument

Finally, let me indicate another argument [Dav14] that justifies the choice of the quantum distribution measure $\rho_\phi = |\phi|^2$ as hidden variable distribution $\tilde{\rho}$. Let me start from a given quantum state ϕ_0, and to simplify take it to be an energy eigenstate of a time symmetric Hamiltonian H of the quantum system, so that the dynamics of the position hidden variable $\vec{x}$ is trivial ($\dot{\vec{x}} = 0$). Now let me perturbs adiabatically the dynamics along a cycle in the space of Hamiltonians, starting and ending at H

$$H \to H(t) = H + \epsilon\, \delta H(t\epsilon) \tag{5.3.59}$$

with $\delta H(0) = \delta H(1) = 0$, as done in the computation of the Berry phase. In the adiabatic limit $\epsilon \to$ after one cycle ($T = 1/\epsilon$) one ends in the same quantum state ϕ_0 (up to a phase, including the Berry phase, of course). But one can also study the adiabatic dynamics of the particle $\vec{x}(t)$ and argue that it is non trivial after one cycle, namely

$$\vec{x}(0) \to \vec{x}(T = 1/\epsilon) \neq \vec{x}(0) \tag{5.3.60}$$

The final position depends of the adiabatic cycle. Therefore, with the exception of the one dimensional case $d = 1$, adiabatic transformations mix the particle positions and the quantum distribution measure $\rho_\phi = |\phi|^2$ is the only probability distribution $\tilde{\rho}$ that is invariant under arbitrary adiabatic cyclic transformations, and hence natural under general unitary Hamiltonian evolutions.

5.3.6.e - The problems with contextual models

Of course the main feature of these models is that they are contextual. They are appropriate, and indeed equivalent to standard quantum mechanics, as long as one wants to describe the properties of a quantum system that depend (at a given time) of a given family of compatible observables. In the pilot-wave models these observables (the context) are the position observables, and these models single out the space of positions of the particles (the configuration space) as playing a specific role.

The original pilot wave model manages to treat also (some aspects of) the single momentum operator P, but none of these contextual models will succeed in treating simultaneously, together with the position operators Q, Q^2, Q^3, etc. all the momentum operators P, P^2, P^3, etc. and the combinations of P's and Q's, in particular the energy operator (the Hamiltonian) H. This impossibility is ensured by the no-go theorems that we discussed (Gleason's theorem, Kochen-Specker's theorem, etc.).

It is indeed possible to construct a hidden variable model that deals with the P, P^2, P^3, etc. but it will be uncorrelated (and ontologically incompatible) with the model for the Q, Q^2, Q^3, etc. In order to construct a consistent hidden variable model, one must first know which physical quantities one is going to measure, and the description of the possible outputs of your measurements in terms of preexisting actual values for the hidden variables will depend on this choice! In the pilot wave model, the position x's variables can be considered as the subsets of hidden variables that you can actually observe, the remaining wave function ψ, i.e. the pilot wave, being actually the really hidden variable that you cannot observe. This description is valid (and sometime quite sensible and useful for physics or chemistry), but only in the context of position observables. To give a firm ontological status to these x variable (namely, stating that they are "element of reality" that preexists the measurements and determine the results of the measurement) is not really possible.

It is sometimes advocated that the position observables have a special status and should be considered as really more fundamental. But I think that this point of view is difficult to defend, especially in high energy physics and in quantum field theory, where the position $\vec{x}$ as well as the time t are not – and cannot be – physical observables!

Finally let me mention that this issue of contextuality and of hidden variables has been also discussed for quantum measurement processes. See for instance [DGZ04]. One has to associate also hidden variables to the measurement apparatus, in order to construct a hidden variable model for quantum measurements. However, in fact one has to introduce hidden variables common to the measurement system and the measurement apparatus, and to assume that there are preexisting variables that correlates the choice of the observed quantity by the observer and the results of the measurements. In particular, this means that the basic independence assumption in deriving the Bell inequalities – namely that the two observer can choose independently whatever they measure – is violated. The former – there are elements of reality that "exist" but always conspire to reproduce the results of quantum mechanics – is often denoted "superdeterminism". I shall not discuss more this point here, but for many physicists superdeterminism leads to formidable philosophical problems and to no predictive powers.

5.4 Summary discussion on quantum correlations

Let me now try to summarize what can be learnt from the last two sections on quantum correlations and the notions of contextuality and non-locality.

The significance and the consequences of Bell's inequalities and their extensions, and of the Kochen-Specker-like theorem's for quantum theory have been enormously discussed, and some debates are still going on. It is not the purpose of these notes to review and summarize all these discussions that took place at the physics level and at the philosophical level. Let me just try to make some simple remarks.

5.4.0.f - Locality and realism

As discussed in section 5.3, the assumption of context-free value definiteness is clearly not tenable, from Gleason's theorem. This means that one must be very careful when discussing quantum physics about correlations between results of measurements. To quote a famous statement by Y. Peres: "Unperformed experiments have no results" [Per78].

Trying to assign some special ontological status to a (finite and in practice small) number of observables to avoid the consequence of the Kochen-Specker theorem may be envisioned, but raises other problems. For instance, if one wants to keep the main axioms of QM, and non-contextuality, by using a finite number of observables, one would expect the quantum logic formalism would lead to QM on a finite division ring (a Galois field), but it is known that this is not possible (see the discussion in 4.3.2). Note however that relaxing some basic physical assumptions like reversibility and unitarity has been considered for instance in [tH07].

It is also clear that non-local quantum correlations are present in non-separable quantum states, highlighted by the violations of Bell's and CHSH-like inequalities (and their numerous and interesting variants). They represent some of the most non-classical and counter-intuitive features of quantum physics. In connexion with the discussions of the "EPR-paradox", this non-local aspect of quantum physics has been often – and are still sometimes –presented as a contradiction between the principles of quantum mechanics and those of special relativity. This is of course not correct.

The concept of local realism, which is incompatible with quantum mechanics and known to be excluded by experiments, is indeed different from the concept of locality used in relativistic quantum field theory. As discussed in 5.3.5, local realism corresponds to the property of "local contextuality" for hidden variable models. This idea was advocated by Einstein and assumed to hold by Einstein, Poldovski and Rosen in their EPR 1935 paper. It means that two causally independent systems can be assigned separate and individual (but locally contextual) hidden variables ("elements of reality") and that classical correlations between these local hidden variables are sufficient to explain the quantum correlations between the two systems when they are entangled.

Locality in quantum field theory is something different, and corresponds to the concepts of local events (localized in space and time) and of the causal relations between these events as related to the geometry of space-time), in particular of there independence for space-separated points. This is the concept of locality and causality as formulated by A. Einstein in 1905 in the theory of special relativity, and then extended to general relativity. These requirements of causality and locality are necessary in quantum theory in order to formulate consistent relativistic quantum field theories, as briefly presented in 3.8. They are constraints on the observables (the operators) of the quantum theory, rather than constraints on the (tentative hidden variable description of the) quantum states. They imply that no information (causal effects) can propagate faster than light, and thus imply the "non-signaling" property of quantum information science. This is the reason why EPR-like experiments and the violations of Bell-like inequalities should not be considered as a signal of some "spooky-at-a-distance physical action" at work in quantum operations over entangled multipartite systems.

$$\text{locality in QFT} \neq \text{local realism}$$

In section 5.3 I discussed also the status of fully contextual models, that would correspond to some form of "non-local contextual realism", and explained why they are in fact quite problematic. I shall come back shortly to these issues in 5.6, when discussing the relations between the formalisms and the interpretations.

5.4.0.g - Chance and correlations

Now let me return to the quantum correlations discussed in 5.2. Contrary to classical physics, in quantum physics there is an irreducible quantum indeterminism and uncertainty in the description of a quantum system. Not all the physical observables can be characterized uniquely and independently at the same time. This essential feature reflects itself in the Heisenberg uncertainty principle, and can be formulated for instance as N. Bohr's "complementarity principle".

However, contrary to what could be expected, this does not mean that a quantum system is always more uncertain or "fuzzy" than a classical system. Indeed, the quantum correlations may very well be stronger than the corresponding classical correlations, when considering bipartites or multipartite systems. This is exemplified for instance by the quantum entropic inequalities 5.2.11 and 5.2.14 when compared to their classical analog, the entropic bound 5.2.12, and by the Tsirelson bound 5.2.26 compared to the B-CHSH inequality 5.3.30. Indeed, thanks to entanglement, quantum systems may be more correlated than what is expected classically when one assumes that correlations come only from classical, local and non-contextual "elements of reality" shared by the systems.

Nevertheless, the quantum correlations are still strongly controlled by the physical principles that we have discussed in the presentation of the formalisms, causality, reversibility and locality (in the causal sense), or by the information principle formulations shortly discussed in 5.1. The "super-strong" correlations (that can be build for instance using the Popescu-Rohrlich boxes discussed in 5.2.2) raise problems and it has not been possible to implement them in a physical theory. This can be represented by the little drawing of Fig. 5.6, where the set of quantum correlations (the red square) is shown to be larger than the set of classical correlations, but smaller that the set of all logically possible correlations.

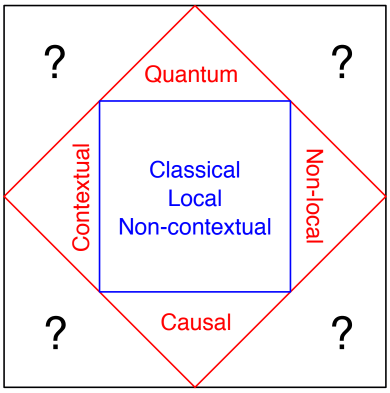

Figure 5.6: Schematic of the worlds of classical correlations, quantum correlations and "super-strong" unphysical correlations

This simple drawing illustrates why the theoretical works by J. Bell and its successors, besides their importance for our theoretical and philosophical understanding of what is and what is not quantum mechanics, turned out to have a significant and long term impact in science and technology. They played an important role in the rise of quantum information science, since they showed that, using quantum correlations and entanglement, it is possible to transmit and manipulate information, perform calculations and search processes, etc. in ways which are impossible by classical means, or which are much more efficient.

5.5 Measurements

5.5.1 What are the questions?

Up to now I have not discussed much the question of quantum measurements. I simply took the standard point of view that (at least in principle) ideal projective measurements are feasible and one should look at the properties of the outcomes. The question is of course highly more complex. In this section I just recall some basic points about quantum measurements.

The meaning of the measurement operations is at the core of quantum physics. It was considered as such from the very beginning. See for instance the proceedings of the famous Solvay 1927 Congress [BV12], and the 1983 review by Wheeler and Zurek [WZ83]. Many great minds have thought about the so called "measurement problem" and the domain has been revived in the last decades by the experimental progresses, which allows now to manipulate simple quantum system and implement effectively ideal measurements.

On one hand, quantum measurements represent one of the most puzzling features of quantum physics. They are non-deterministic processes (quantum mechanics predicts only probabilities for outcomes of measurements). They are irreversible processes (the phenomenon of the "wave-function collapse"). They reveal the irreducible uncertainty of quantum physics (the uncertainty relations). This makes quantum measurements very different from "ideal classical measurements".

On the other hand, quantum theory is the first physical theory that addresses seriously the problem of the interactions between the observed system (the object) and the measurement apparatus (the observer). Indeed in classical physics the observer is considered as a spectator, able to register the state of the real world (hence to have its own state modified by the observation), but without perturbing the observed system in any way. Quantum physics shows that this assumption is not tenable. Moreover, it seems to provide a logically satisfying answer [3] to the basic question: what are the minimal constraints put on the results of physical measurements by the basic physical principles [4].

It is often stated that the main problem about quantum measurement is the problem of the uniqueness of the outcome. For instance, why do we observe a spin 1/2 (i.e. a q-bit) in the state $|\uparrow\rangle$ or in the state $|\downarrow\rangle$ when we start in a superposition $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$? However by definition a measurement is a process which gives one single classical outcome (out of several possible). Thus in my opinion the real questions, related to the question of the "projection postulate", are: (1) Why do repeated ideal measurements should give always the same answer? (2) Why is it not possible to "measure" the full quantum state $|\psi\rangle$ of a q-bit by a *single* measurement operation, but only its projection onto some reference frame axis?

Again, the discussion that follows is very sketchy and superficial. A good recent reference, both on the history of the "quantum measurement problem", a detailed study of explicit dynamical models for quantum measurements, and a complete bibliography, is the research and review article [ABN12].

3. If not satisfying every minds, every times...
4. Well... as long as gravity is not taken into account!

5.5.2 The von Neumann paradigm

The general framework to discuss quantum measurements in the context of quantum theory is provided by J. von Neumann in his 1932 book [vN32, vN55]. Let me present it on the simple example of the q-bit.

But before, let me insist already on the fact that this discussion will not provide a derivation of the principle of quantum mechanics (existence of projective measurements, probabilistic features and Born rule), but rather a self-consistency argument of compatibility between the axioms of QM about measurements and what QM predicts about measurement devices.

An ideal measurement involves the interaction between the quantum system $\mathcal{S}$ (here a q-bit) and a measurement apparatus $\mathcal{M}$ which is a *macroscopic object*. The idea is that $\mathcal{M}$ must be treated as a quantum object, like $\mathcal{S}$. An ideal non destructive measurement on $\mathcal{S}$ that does not change the orthogonal states $|\uparrow\rangle$ and $|\downarrow\rangle$ of $\mathcal{S}$ (thus corresponding to a measurement of the spin along the z axis, S_z), correspond to introducing for a finite (short) time an interaction between $\mathcal{S}$ and $\mathcal{M}$, and to start from a well chosen initial state $|I\rangle$ for $\mathcal{M}$. The interaction and the dynamics of $\mathcal{M}$ must be such that, if one starts from an initial separable state where $\mathcal{S}$ is in a superposition state

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle \tag{5.5.1}$$

after the measurement (interaction) the whole system (object+apparatus) is in an entangled state

$$|\psi\rangle\otimes|I\rangle \qquad \to \qquad \alpha|\uparrow\rangle\otimes|F_+\rangle + \beta|\downarrow\rangle\otimes|F_-\rangle \tag{5.5.2}$$

The crucial point is that the final states $|F_+\rangle$ and $|F_-\rangle$ for $\mathcal{M}$ must be *orthogonal* [5]

$$\langle F_+|F_-\rangle = 0 \tag{5.5.3}$$

Of course this particular evolution 5.5.1 is unitary for any choice of $|\psi\rangle$, since it transforms a pure state into a pure state.

$$|\psi\rangle\otimes|I\rangle \qquad \to \qquad \alpha|\uparrow\rangle\otimes|F_+\rangle + \beta|\downarrow\rangle\otimes|F_-\rangle \tag{5.5.4}$$

One can argue that this is sufficient to show that the process has all the characteristic expected from an ideal measurement, within the quantum formalism itself. Indeed, using the Born rule, this is consistent with the fact that the state $\alpha|\uparrow\rangle$ is observed with probability $p_+ = |\alpha|^2$ and the state $\alpha|\downarrow\rangle$ with probability $p_- = |\beta|^2$. Indeed the reduced density matriices both for the system $\mathcal{S}$ and for the system $\mathcal{M}$ (projected onto the two pointer states) is that of a completely mixed state

$$\rho_{\mathcal{S}} = \begin{pmatrix} p_+ & 0 \\ 0 & p_- \end{pmatrix} \tag{5.5.5}$$

For instance, as discussed in [vN32][vN55], if one is in the situation where the observer $\mathcal{O}$, really observe the measurement apparatus $\mathcal{M}$, not the system $\mathcal{S}$ directly, the argument can be repeated as

$$|\psi\rangle\otimes|I\rangle\otimes|O\rangle \qquad \to \qquad \alpha|\uparrow\rangle\otimes|F_+\rangle\otimes|O_+\rangle + \beta|\downarrow\rangle\otimes|F_-\rangle\otimes|O_-\rangle \tag{5.5.6}$$

5. as already pointed out in [vN32]

and it does not matter if one puts the fiducial separation between object and observer between $\mathcal{S}$ and $\mathcal{M}+\mathcal{O}$ or between $\mathcal{S}+\mathcal{M}$ and $\mathcal{O}$. This argument being repeated ad infinitum.

A related argument is that once a measurement has been performed, if we repeat it using for instance another copy $\mathcal{M}'$ of the measurement apparatus, after the second measurement we obtain

$$|\psi\rangle\otimes|I\rangle\otimes|I'\rangle \qquad \to \qquad \alpha|\uparrow\rangle\otimes|F_+\rangle\otimes|F'_+\rangle+\beta|\downarrow\rangle\otimes|F_-\rangle\otimes|F'_-\rangle \tag{5.5.7}$$

so that we never observe both $|\uparrow\rangle$ and $|\downarrow\rangle$ in a successive series of measurements (hence the measurement is really a projective measurements). The arguments holds also if the outcome of the first measurement is stored on some classical memory device $\mathcal{D}$ and the measurement apparatus reinitialized to $|I\rangle$. This kind of argument can be found already in [Mot29].

The discussion here is clearly outrageously oversimplified and very sketchy. For a precise discussion, one must distinguish among the degrees of freedom of the measurement apparatus $\mathcal{M}$ the (often still macroscopic) variables which really register the state of the observed system, the so called *pointer states*, from the other (numerous) microscopic degrees of freedom of $\mathcal{M}$, which are present anyway since $\mathcal{M}$ is a macroscopic object, and which are required both for ensuring decoherence (see next section) and to induce dissipation, so that the pointer states become stable and store in a efficient way the information about the result of the measurement. One must also take into account the coupling of the system $\mathcal{S}$ and of the measurement apparatus $\mathcal{M}$ to the environment $\mathcal{E}$.

5.5.3 Decoherence, ergodicity and mixing

As already emphasized, the crucial point is that starting from the same initial state $|I\rangle$, the possible final pointer states for the measurement apparatus, $|F_+\rangle$ and $|F_-\rangle$, are orthogonal. This is now a well defined dynamical problem, which can be studied using the theory of quantum dynamics for closed and open systems. The fact that $\mathcal{M}$ is macroscopic, i.e. that its Hilbert space of states in very big, is essential, and the crucial concept is *decoherence* (in a general sense).

The precise concept and denomination of quantum decoherence was introduced in the 70's (especially by Zeh) and developed and popularized in the 80's (see the reviews [JZK$^+$03], [Zur03]). But the basic idea seems much older and for our purpose one can probably go back to the end of the 20′ and to von Neumann's quantum ergodic theorem [vN29] (see [vN10] for the english translation and [GLM$^+$10] for historical and physical perspective).

One starts from the simple geometrical remark [vN29] that if $|e_1\rangle$ and $|e_2\rangle$ are two *random unit vectors* in a N dimensional Hilbert space $\mathcal{H}$ (real or complex), their average "overlap" (squared scalar product) is of order

$$\overline{|\langle e_1|e_2\rangle|^2} \simeq \frac{1}{N} \quad , \qquad N=\dim(\mathcal{H}) \tag{5.5.8}$$

hence it is very small, and for all practical purpose equal to 0, if N is very large. Remember that for a quantum system made out of M similar subsystems, $N\propto(N_0)^M$, N_0 being the number of accessible quantum states for each subsystem.

A simple idealized model to obtain a dynamics of the form 5.5.4 for $\mathcal{S}+\mathcal{M}$ is to assume that both $\mathcal{S}$ and $\mathcal{M}$ have no intrinsic dynamics and that the evolution during the interaction/measurement time interval is given by a interaction Hamiltonian (acting on the Hilbert space $\mathcal{H}=\mathcal{H}_\mathcal{S}\otimes\mathcal{H}_\mathcal{M}$ of $\mathcal{S}+\mathcal{M}$) of the form

$$H_{\mathrm{int}} = |\uparrow\rangle\langle\uparrow|\otimes H_+ + |\downarrow\rangle\langle\downarrow|\otimes H_- \tag{5.5.9}$$

where H_+ and H_- are two *different* Hamiltonians (operators) acting on $\mathcal{H}_\mathcal{M}$. It is clear that if the interaction between $\mathcal{S}$ and $\mathcal{M}$ takes place during a finite time t, and is then switched off, the final state of the system is an entangled one of the form 5.5.4, with

$$|F_+\rangle = \mathrm{e}^{\frac{t}{\mathrm{i}\hbar}H_+}|I\rangle \quad , \qquad |F_-\rangle = \mathrm{e}^{\frac{t}{\mathrm{i}\hbar}H_-}|I\rangle \tag{5.5.10}$$

so that

$$\langle F_+|F_-\rangle = \langle I|\mathrm{e}^{-\frac{t}{\mathrm{i}\hbar}H_+}\cdot\mathrm{e}^{\frac{t}{\mathrm{i}\hbar}H_-}|I\rangle \tag{5.5.11}$$

It is quite easy to see that if H_+ and H_- are not (too much) correlated (in a sense that I do not make more precise), the final states $|F_+\rangle$ and $|F_-\rangle$ are quite uncorrelated with respect to each others and with the initial state $|I\rangle$ after a very short time, and may be considered as random states in $\mathcal{H}_\mathcal{M}$, so that

$$|\langle F_+|F_-\rangle|^2 \simeq \frac{1}{\dim(\mathcal{H}_\mathcal{M})} \lll 1 \tag{5.5.12}$$

so that for all practical purpose, we may assume that

$$\langle F_+|F_-\rangle = 0 \tag{5.5.13}$$

This is the basis of the general phenomenon of *decoherence*. The interaction between the observed system and the measurement apparatus has induced a decoherence between the states $|\uparrow\rangle$ and $|\downarrow\rangle$ of $\mathcal{S}$, but also a decoherence between the pointer states $|F_+\rangle$ and $|F_-\rangle$ of $\mathcal{M}$.

Moreover, the larger $\dim(\mathcal{H}_\mathcal{M})$, the smaller the "decoherence time" beyond which $\langle F_+|F_-\rangle \simeq 0$ is (and it is often in practice too small to be observable), and the larger (in practice infinitely larger) the "quantum Poincaré recurrence time" (where one might expect to get again $|\langle F_+|F_-\rangle| \simeq 1$) is.

Of course, as already mentionned, this is just the first step in the discussion of the dynamics of a quantum measurement. One has in particular to check and to explain how, and under which conditions, the pointer states are quantum microstates which correspond to macroscopic classical-like macrostates, which can be manipulated, observed, stored in an efficient way. At that stage, I just paraphrase J. von Neumann (in the famous chapter VI "Der Meßprozeß" of [vN32])

"Die weitere Frage (...) soll uns dagegen nicht beschäftigen."

Decoherence is a typical quantum phenomenon. It explains how, in most situations and systems, quantum correlations in small (or big) multipartite systems are "washed out" and disappear through the interaction of the system with other systems, with its environment or its microscopic internal degrees of freedom. Standard references on decoherence and the general problem of the quantum to classical transitions are [Zur90] and[Sch07].

However, the underlying mechanism for decoherence has a well know classical analog: it is the (quite generic) phenomenon of *ergodicity*, or more precisely the *mixing property* of classical dynamical systems. I refer to textbooks such as [AA68] and [LL92] for precise mathematical definitions, proofs and details. Again I give here an oversimplified presentation.

Let us consider a classical Hamiltonian system. One considers its dynamics on (a fixed energy slice $H = E$ of) the phase space Ω , assumed to have a finite volume $V = \mu(\Omega)$ normalized to $V = 1$, where μ is the Liouville measure. We denote T the volume preserving map $\Omega \to \Omega$ corresponding to the integration of the Hamiltonian flow during some reference time t_0. T^k is the iterated map (evolution during time $t = kt_0$). This discrete time dynamical mapping given by T is said to have the *weak mixing property* if for any two (measurable) subsets A and B of Ω one has

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \mu(B\cap T^k A) = \mu(B)\mu(A) \tag{5.5.14}$$

The (weak) mixing properties means (roughly speaking) that, if we take a random point a in phase space, its iterations $a_k = T^k a$ are at large time and "on the average" uniformly distributed on phase space, with a probability $\mu(B)/\mu(\Omega)$ to be contained inside any subset $B \in \Omega$. See fig. 5.7

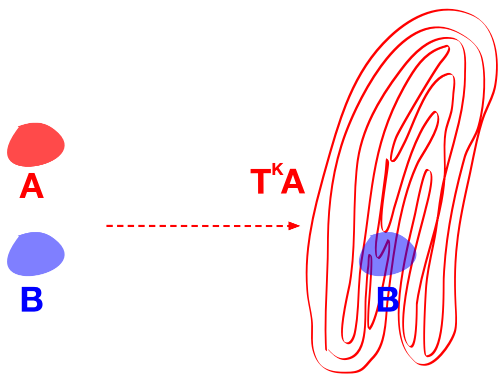

Figure 5.7: Graphical representation of the mixing property (very crude)

Weak mixing is one of the weakest form of "ergodicity" (in a loose sense, there is a precise mathematical concept of ergodicity).

Now in semiclassical quantization (for instance using Bohr-Sommerfeld quantization rules) if a classical system has M independent degrees of freedom (hence its classical phase space Ω has dimension $2M$), the "quantum element of phase space" $\delta\Omega$ has volume $\delta V = \mu(\delta\Omega) = h^M$ with $h = 2\pi\hbar$ the Planck's constant. If the phase space is compact with volume $\mu(\Omega) < \infty$ the number of "independent quantum states" accessible to the system is of order $N = \mu(\Omega)/\mu(\delta\Omega)$ and should correspond to the dimension of the Hilbert space $N = \dim(\mathcal{H})$. In this crude semiclassical picture, if we consider two pure quantum states $|a\rangle$ and $|b\rangle$ and associate to them two minimal semiclassical subsets A and B of the semiclassical phase space Ω, of quantum volume δV, the semiclassical volume $\mu(A \cap B)$ corresponds to the overlap between the two quantum pure states through

$$\mu(A \cap B) \simeq \frac{1}{N} |\langle a|b\rangle|^2 \tag{5.5.15}$$

More generally if we associate to any (non minimal) subset A of Ω a mixed state given by a quantum density matrix ρ_A we have the semiclassical correspondence

$$\frac{\mu(A \cap B)}{\mu(A)\mu(B)} \simeq N\, \mathrm{tr}(\rho_A \rho_B) \tag{5.5.16}$$

With this semiclassical picture in mind (Warning! It does not work for all states, only for states which have a semiclassical interpretation! But pointer states usually do.) the measurement/interaction process discussed above has a simple semiclassical interpretation, illustrated on fig. 5.8.

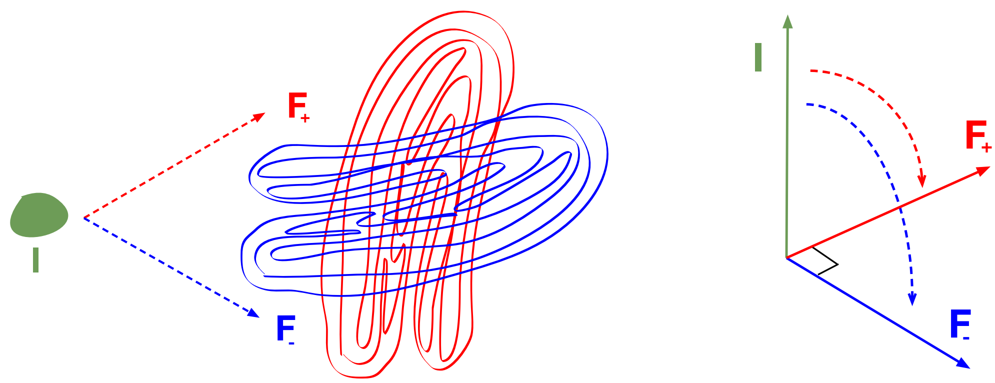

Figure 5.8: Crude semiclassical and quantum pictures of the decoherence process 5.5.10-5.5.12

The big system $\mathcal{M}$ starts from an initial state $|I\rangle$ described by a semiclassical element I. If the system $\mathcal{S}$ is in the state $|\uparrow\rangle$, $\mathcal{M}$ evolves to a state $|F_+\rangle$ corresponding to F_+. If it is in the state $|\uparrow\rangle$, $\mathcal{M}$ evolves to a state $|F_-\rangle$ corresponding to F_-. For well chosen, but quite generic Hamiltonians H_+ and H_-, the dynamics is mixing, so that, while $\mu(F_+) = \mu(F_-) = 1/N$, typically one has $\mu(F_+ \cap F_-) = \mu(F_+)\mu(F_-) = 1/N^2 \ll 1/N$. Thus it is enough for the quantum dynamics generated by H_+ and H_- to have a quantum analog the classical property of mixing, which is quite generic, to "explain" why the two final states $|F_+\rangle$ and $|F_-\rangle$ are generically (almost) orthogonal.

5.5.4 Discussion

As already stated, the points that I tried to discuss in this section represent only a small subset of the questions about measurements in quantum mechanics. Again, I refer for instance to [ABN12] and [Lal12] (among many other reviews) for a serious discussion and bibliography.

I have not discussed more realistic measurement processes, in particular the so called "indirect measurements procedures", where the observations on the system are performed through successive interactions with other quantum systems (the probes) which are so devised as to perturb as less as possible the observed system, followed by stronger direct (in general destructive) measurements of the state of the probes. Another class of processes is the "weak measurements", which are a series of measurements that perturb weakly the measured system, combined with some adequate postselection process in order to define the "weak value" of an observable. Such measurement processes, as well as many interesting questions and experiments in quantum physics, quantum information sciences, etc. are described by the general formalism of POVM's (Positive Operator Valued Measure). I do not discuss these questions here.

In any case, important aspects of quantum measurements belong to the general class of problems of the emergence and the meaning of irreversibility out of reversible microscopic laws in physics (quantum as well as classical). See for instance [HPMZ96].

The quantum formalism as it is presented in these lectures starts (amongst other things) from explicit assumptions on the properties of measurements. The best one can hope is to show that the quantum formalism is consistent: the characteristics and behavior of (highly idealized) physical measurement devices, constructed and operated according to the laws of quantum mechanics, should be consistent with the initials axioms.

One must be careful however, when trying to justify or rederive some of the axioms of quantum mechanics from the behavior of measurement apparatus and their interactions with the observed system and the rest of the world, not to make circular reasoning.

5.6 Formalisms, interpretations and alternatives to quantum mechanics

5.6.1 What about interpretations?

In these notes I have been careful up to now not to discuss the interpretation issues of quantum mechanics. There are at least two reasons.

Firstly, I do not feel qualified enough to review and discuss all the interpretations of quantum mechanics that have been proposed and all the philosophical questions raised by quantum physics since its birth. This does not mean that I consider these question to be unimportant (nor that I never think about these).

Secondly, these lecture notes are focused on the presentation of the mathematical formalism of "standard quantum mechanics", and on the main approaches to present and justify physically the formalism. The point of view chosen is that of most physicists, chemists, mathematicians, computer scientists, engineers, etc who study or exploit the ressources of the quantum world. This "operational point of view" consists in considering quantum mechanics as a theoretical framework that provides rules to compute the probabilities to obtain a given result/response when measuring/manipulating a quantum system as a function of its state. The concepts of "observables", "states" and "probabilities" being defined within the principles of the formalisms considered.

Of course this point of view may be already considered as a interpretational prejudice. So let me nevertheless come back to these issues of interpretations and try to make a few simple – and probably naïve – remarks.

As already explained in the introduction, I find the presentation and the discussion about the "interpretations of quantum mechanics" a bit confusing (even in some excellent reviews and textbooks) because they mix three different levels and topics: (i) the formalisations (mathematical formulations) of standard quantum theory, (ii) the various interpretations of the formalisms and of their principles for quantum theory, (iii) alternate theories, that try to obviate or solve in different ways some of the questions raised by quantum mechanics, but which are different physical theories of the quantum world, leading to different predictions than standard quantum mechanics in some regimes. These different concepts are usually presented at the same level, and denoted generically "interpretations of quantum mechanics", although in my opinion they may be quite different things.

5.6.2 Formalisms

Formulations such as the "canonical formalism", the path integral formalism, the algebraic and the quantum logic formalisms are in my opinion different mathematical representations of the same physical theory (or rather of the same physical framework in which one formulates the physical laws and physical models, in the same sense that classical physics is the framework for the physical laws of classical models). Their starting principles might be different, but they will lead to the same results or equivalent ones at the end. These different representations have their pros and cons, with different levels of mathematical rigor and of operability. However it is expected that they are not contradictory within their common domains of validity. Even if a real mathematician can claim – for instance – that no consistent quantum field theory has been rigorously constructed yet in 4 dimensional Minkowski space time, most physicists believe that this goal will be reached one day. Perhaps this will be done using the algebraic formalism, with some path integral and renormalization group. Perhaps some more powerful mathematical formalism will have to be used, quite different from the formalisms that we know at the moment.

5.6.3 Interpretations

Interpretations of quantum mechanics, such as the "Copenhagen interpretation" or the "many-worlds interpretation", are something slightly different. They do not challenge the present standard mathematical formulations of the theory, but rather insist on a particular point of view or a particular formulation of quantum mechanics as the best suited or the preferable one to consider and study quantum systems, and the quantum world, and they incorporate some more thoughts on the meaning of these principles. Thus in my opinion, rather than being formalizations of quantum mechanics, they should be considered as different and particular choices of points of view and of philosophical options to think about quantum mechanics and practice it.

This does not mean that I am a conscious adept of post-modern relativism...

Let me first discuss the example of the so-called "Copenhagen Interpretation". Remember however that there is no clear cut definition of what "The Copenhagen Interpretation" is. Although they exist since the birth of quantum mechanics, the term "Copenhagen interpretation" was introduced only in 1955 by Heisenberg. I refer to the paper by Howard [How04] for an historical and critical review of the history, uses and misuses of the concept. There are various points of view "from Copenhagen", some purely frequentists (quantum mechanics should apply to ensembles of systems only), some based on a more Bayesian concept and use of probabilities (quantum mechanics may be applied to single non-repeatable experiments on a single system). All this class of interpretations insist anyway on the fact that quantum mechanics deals only with the result of experiments and make predictions for the results of these experiments, not on some non-accessible underlying "reality". The presence of an "observer" (that does not need to be a conscious being) is thus important, either as a principle, or through theories of measurement processes or decoherence. Therefore I think the various Copenhagen interpretations may be considered as

Copenhagen = "quantum mechanics from a pragmatist point of view"

where "pragmatism" should be understood in the philosophical sense of pragmatism. Indeed these interpretations seems to be the most used (explicitly or implicitly) in experimental quantum physics and its applications.

At the other extreme one finds the "Many Worlds Interpretations". Again there are many variants, starting from the original proposal of Everett, its revised version by Wheeler and its presentation by B. Dewitt, who coined the term "many worlds" (see [DG73] for references), to the most recent views (see for instance [Deu97]). The basic idea (as far as I understand) is to take seriously the concept of "the wave function of the universe" as the underlying concept of quantum mechanics, without reference to the observer, and to reinterpret or rederive the probabilistic features of quantum mechanics (in particular the Born rule and the projection postulate) as "relative" to some parts of this universal wave function, corresponding to some specific state of the observer for instance. From what I have seen, it seems that most proponents of the MW interpretations (but not all) agree on the fact that it does not allow to prove ab initio the Born law of probabilities. However, the MW point of view is often used in the field of cosmology and of quantum gravity and quantum cosmology, where the concept of "wave function of the universe" has to be tackled (for instance to discuss the Wheeler-DeWitt equation). It is also quite popular in some quantum information circles. I am not going to review or discuss more these interpretations, and refer to the recent proceedings+discussion book [SBKW10], which contains a modern presentation of the subject, and stimulating (and often contradictory) discussions and points of view by proponents and opponents. The role given to a (in practice unobservable) universal wave function that represents some "underlying reality" (but not in the sense of Einstein or Bell), while keeping the mathematical apparatus of quantum mechanics, is the central point for these interpretations. Therefore I think that most Many Worlds Interpretations may be considered as

MW = "quantum mechanics from a realist point of view"

Again, here realism is to be understood in its philosophical sense.

There is a whole spectrum of proposed interpretations that lie between these two main classes. Let me mention the “coherent history formulations" that insist on the fact that on should discuss and formulate quantum mechanics only in terms of “coherent histories”, namely coherent semiclassical processes that admit some semiclassical description (the existence of such histories being justified by models of decoherence and coarse graining processes). I refer to [Gri02] for details and references. This approach is also used in quantum gravity and quantum cosmology.

Another class of interpretations are the “modal interpretations” that try to make ontologically consistent the hidden variable theories based on pilot wave models by assigning a “modal status of reality” to the classical variables (such as the position of the particles driven by the wave function) so as to make them consistent with standard quantum mechanics. See [LD14] for a review of theses approaches.

There are many other (often overlapping) classes of interpretations that do not challenge in fact the mathematical formalism of quantum mechanics. A probably naïve and amateurish way to consider this “many-interpretations world of quantum mechanics” is to rather consider them as various point of view, from different philosophical perspectives, of the quantum formalism. Quantum mechanics is an impressive, beautiful and quite self consistent theoretical framework to describe physical phenomena, and it deserves several vistas to be appreciated and understood in its full majesty!

5.6.4 Alternatives

The interpretations that rely on the mathematical formulations of standard quantum mechanics should be clearly distinguished from another class of proposals to explain quantum physics that rely on modifications of the principles of quantum mechanics, and are thus different physical theories. These modified or alternative quantum theories deviate from “standard” quantum mechanics and should be experimentally falsifiable (and sometimes are already falsified).

This is the case of the various hidden-variables proposals, such as the de Broglie-Bohm pilot wave theories. It is often stated that the later formulations are equivalent to standard quantum mechanics, and fall into the previous category of interpretations. As discussed in section 5.3.6, this is correct only in the restricted context of position observables (for the simple case of non-relativistic particles), and generaly when dealing with a specified family of compatible observables, since these hidden variable models are contextual. However giving an ontological status to the position variable x of the particle driven by the pilot wave ψ, by considering x as an element of reality that can be related to what is observed, changes in my opinion the status of this model. It is now a different theoretical model than quantum mechanics, since some observables of quantum mechanics, like the impulsion p, are not described by this model, while the velocity of the particle $\dot{x}$ is a different variable that cannot be directly described by quantum mechanics.

This is also the case for the class of models known as “collapse models”. See [GRW85, GRW86] for the first models. In these models the quantum dynamics is modified by non-linear terms so that the evolution of the wave functions is not unitary any more (while the probabilities are conserved of course), and the "collapse of the wave function" becomes now a dynamical phenomenon. These models are somehow phenomenological, since the origin of these non-linear dynamical effects is quite ad hoc, they may be for instance justified by some quantum gravity effects. They predict a breakdown of the law of quantum mechanics for the evolution of quantum coherences and for decoherence phenomenon at large times, large distances, or in particular for big quantum systems (for instance large molecules or atomic clusters). Hence they can in principle be tested experimentally. At the present day, despite the impressive experimental progresses in the control of quantum coherences, quantum measurements, study of decoherence phenomenon, manipulation of information in quantum systems, no such violations of the predictions of standard QM and of unitary dynamics have been observed.

5.7 What about gravity?

Another really important issue that I do not discuss in these lecture notes is quantum gravity. Again just a few simple remarks.

It is clear that the principles of quantum mechanics are challenged by the question of quantizing gravity. The challenges are not only technical. General relativity (GR) is indeed a non-renormalizable theory, and from that point of view a first and natural idea is to consider it as an effective low energy theory. After all, history tells us that in the development of nuclear and particle physics there has been several times (in the 30′, the 40′, the 60′...) theoretical false alarms and clashes between experiments and theory. Each time this led many great minds to question the principles of quantum mechanics themselves. However further development and understandings of the standard formalism (and experiments of course) allowed to solve these problems, so that quantum mechanics came out unscathed and even stronger. Since the 1970′s and the construction of the standard model of strong and electroweak interactions the principles of quantum mechanics are not challenged any more.

However with gravity the situation is different. For instance the discovery of the Bekenstein-Hawking entropy of black holes, of the Hawking radiation, and of the "information paradox" shows that fundamental questions remain to be understood about the relation between quantum mechanics and the GR concepts of space and time. Indeed even the most advanced quantum theories available, quantum field theories such as non-abelian gauge theories the standard model, its supersymmetric and/or grand unified extensions, still rely on the special relativity concept of space-time, or to some extend to the dynamical but still classical concept of curved space-time of GR. It is clear that a quantum theory of space time will deeply modify, and even abolish, the classical concept of space-time as we are used to. Let me stress two important points.

Firstly, the presently most advanced attempts to build a quantum theory incorporating gravity, namely string theory and its modern extensions, as well as the alternative approaches to build a quantum theory of space-time such as loop quantum gravity (LQG) and spin-foam models (SF), rely mostly on the quantum formalism as we know it, but change the fundamental degrees of freedom (drastically and quite widely for string theories, in a more conservative way for LQG/SF). The fact that string theories offers some serious hints of solutions of the information paradox, and some explicit solutions and ideas, like holography and AdS/CFT dualities, for viewing space-time as emergent, is a very encouraging fact.

Secondly, in the two formalisms presented here, the algebraic formalism and the quantum logic formulations, it should be noted that space and time (as continuous entities) play a secondary role with respect to the concept of causality and locality/separability. I hope this is clear in the way I choose to present the algebraic formalism in section 3 and quantum logic in section 4. Of course space-time is essential for constructing physical theories out of the formalism. But the fact that it is the causal relations and the causal independence between physical measurement operations that are essential for the formulation of the theory is in my opinion a very encouraging signal that quantum theory may go beyond the classical concept of space-time.

Nevertheless, it may very well happens that (for instance) the information paradox is not solved by a sensible quantum theory of gravity, or that the concepts of causality and separability have to be rejected. Indeed one might imagine that no repeatable measurements are possible in a quantum theory of gravity, or that space being an "emergent" concept two separate sub-ensembles-of-degrees-of-freedom may never be considered as really causally independent. Then one might expect that the basic principles of quantum mechanics will not survive (and, according to the common lore, should be replaced by something even more bizarre and inexplicable...).