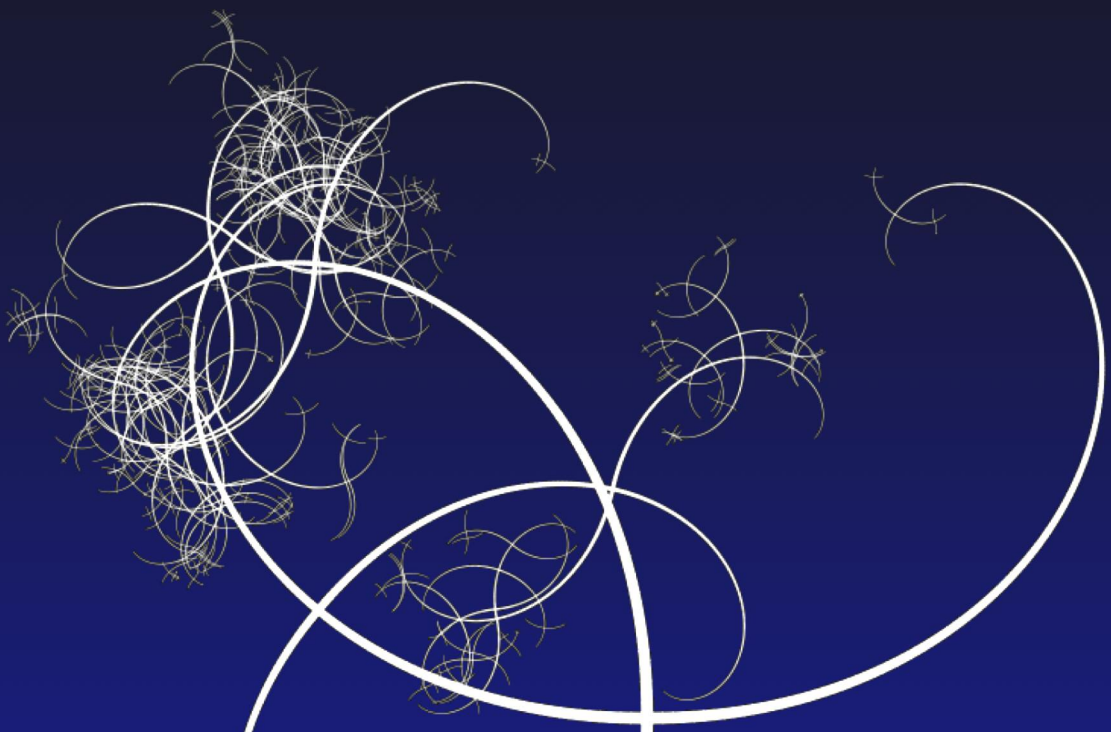
NISSUNA UMANA INVESTIGAZIONE SI PUO DIMANDARE VERA SCIENZIA
S'ESSA NON PASSA PER LE MATEMATICHE DIMOSTRAZIONI
LEONARDO DA VINCI



vol. 4 no. 3-4 2016

Mathematics and Mechanics *of* Complex Systems

Gianfausto Dell'Antonio

QUANTUM MECHANICS: LIGHT AND SHADOWS (ONTOLOGICAL PROBLEMS AND EPISTEMIC SOLUTIONS)

msp

MATHEMATICS AND MECHANICS OF COMPLEX SYSTEMS
Vol. 4, No. 3-4, 2016
dx.doi.org/10.2140/memocs.2016.4.423

QUANTUM MECHANICS: LIGHT AND SHADOWS (ONTOLOGICAL PROBLEMS AND EPISTEMIC SOLUTIONS)

GIANFAUSTO DELL'ANTONIO

We discuss several problems that arise in the Copenhagen interpretation of quantum mechanics, in an attempt to come to grips with what E. T. Jaynes has called *the quantum omelette*.[1]

1. Introduction

In this contribution in honor of Lucio Russo, friend and admired colleague, I will present some remarks on the status of quantum mechanics, a theory through which we try to understand the world of atoms and molecules.

The research in this field has led to extraordinary successes; in fact, our present technology is to a large extent based on our description of this world.

We are convinced that we have found the key that opens the door to its full content, and that in the future our task is only (!) to unravel ways to solve complicated equations, maybe with the aid of a computer.

But quantum mechanics (QM) also poses conceptual problems and has been the arena of debates since the times of the founding fathers.

These problems are related to the meaning of QM as a physical theory.

The problematic relations between ontology and epistemology have been debated in western culture since the time of Plato, but the debate has seen a new life with QM since the basic foundational elements of classical physics are not valid in QM.

I will consider here only the traditional presentation of QM based on the formulation given by the Copenhagen school and developed among others by Born, Jordan, Heisenberg, and Schrödinger.

Generally speaking, this is the only presentation taught in universities and known to the majority of physicists.

Another presentation, which I will call *pilot wave theory*, originated by de Broglie and brought to a high mathematical standard by S. Goldstein and D. Dürr, gives

Communicated by Raffaele Esposito.
PACS2010: 03.65.Ta.
Keywords: quantum mechanics, observables, tracks, Born's rule.

[1] See [Jaynes 1990].

a totally different representation of the building blocks of theory and has its own conceptual difficulties. I shall come back briefly to this theory.

It is a basic assumption of western science that existence and reality can be represented through a *metaphysical system* which, while creating its own fundaments, serves the purpose of *representing reality*.

One should recognize that the origin of physics is the idea that *reality exists* and is at least partially accessible to our inspection.

At the same time, humans shape their experiences not only through their senses but also through their metaphysical and categorial *presuppositions*.

Causality, identity, and noncontradiction are not regarded as platonic concepts that humans *discovered* in the world as ideas but rather as prior conditions for human understanding.

The categorial representation of reality, in particular Newtonian space-time, limits and configures in a definite manner Newtonian physics.

On the other hand, there is David Hume's analysis of the inductive nature of science and the impossibility of grounding the notion of causation in experience. Causation as such is never found in the observable word; it is rather a metaphysical presupposition which allows the subject to *make sense* of observations.

In the same way, *identity and noncontradiction* are the conditions that constrain our observations.

Positivism relies on the distinction between empirical terms (*empirically given* in physical theories and experiments) and *theoretical terms* which are their translations into simple statements.

In this way, *true knowledge* (episteme) is replaced by *objective knowledge* with humans' *shaping* (experience).

One may say that the machian critic of Newton mechanics paved the way for quantum mechanics.

In fact, the multiple representations provided by different modern theories suggest that the successes of human understanding must be regarded as *creations* rather that *discoveries*.

When discussing the breakdown of the foundations of theories in the twentieth century, Wolfgang Pauli [1994] remarked that *the modern physicist regards with skepticism* philosophical systems which, while imagining that they have definitively recognized the a priori conditions of human understanding itself, have in fact succeeded only in setting up *the a priori conditions* of the system of mathematics and *the exact sciences of a particular epoch*.

Still we believe that physical theories are not only mirrors that reflect our own beliefs.

Within physical discourse, a cornerstone is counterfactual reasoning. If a theory makes predictions which agree reasonably well with experimental or observational

results, scientists are inclined to believe that its logical and mathematical structure reflects the structure *of the real world* in some way, even if the philosophers remain permanently skeptical [Griffiths 2002].

In this regard, the classical representation of physics was produced when Newton related the mathematical theory of calculus with physical notions such as space, time, force, particles, and mass.

This representation was extended by Maxwell relating the theory of partial differential equations with electromagnetism and introducing in physics the notions of *charge* and *fields*.

This approach, contrary to the pythagorean–platonic view, regards mathematics *as a nonrepresentative discipline*. It is only physics which, *making use of the mathematical formalism*, attempts to discuss physical reality.

In this respect, it can be said that quantum mechanics represents a dissolution of this classical representation of the world.

2. Origins of QM

Conventionally one places the beginning of quantum mechanics at Planck's formulation of the quantum postulate. The theory soon went beyond its original formulation, which aimed to justify the spectrum of black-body radiation.

It was observed that in the *quantum world* some physical quantities, in particular the energy of the states of the atoms, seemed to be forced to have a discrete spectrum. This is of course totally different from the classical world as we perceive with our senses.

On the basis of the relation between energy and momentum of *particles of light* established by Einstein, this provided a *quantization* also of the frequency of the radiation emitted or absorbed.

The energy of the states of the atoms and some quantities related to a pair of states (emission or absorption of light) were the only information accessible to experimenters.

Following the positivistic rule that only observable quantities should be considered within a theory, W. Heisenberg advanced the first closed formulation of the theory, matrix mechanics.

The theory was not designed to talk about *trajectories of particles*; following the Einstein dictum "it is only the theory that tells you what you can observe", Heisenberg derived the *uncertainty principle from matrix mechanics and the quantum postulate*.

Notice that in this postulate enter quantities (position and momentum) that are well defined in classical physics, but are attributed to matrices in this formulation of QM.

Through a more accurate analysis, for which the contributions of M. Born and P. Jordan and the reference to the dispersion relations of Kramers were essential, it was soon discovered that the theory could have a wider scope, still within atomic physics.

The theory of matrices was given a stronger mathematical flavor with the theory of operators in a Hilbert space. Notice that at the very same time and in the same place (Göttingen) Hilbert was developing his functional calculus (and Jordan was his assistant for some time).

Several other matrices entered in the relations that were derived with matrix mechanics. The ones that more frequently appeared in the analysis were powers of position matrices, powers of momentum matrices, and their real linear span.

If $a_{i,j}$ is a matrix that *represents* the observable A and ϕ_i is the vector associated to the state S in order to account for experimental data, $\sum_{i,j} \bar{\phi}_i a_{i,j} \phi_j$ has to be taken to be the expected value for a measurement of A in the state S.

The measurements resulted in a real number, and therefore, all matrices which were used were hermitian. It was natural (and mathematically more convenient) to agree that *all hermitian matrices* represent observables.

Notice that *observable* is here understood as *a definition*, without reference to the instrumental apparatus that can be used to measure its value. This is a clear violation of the positivistic rule, since for very few observables can a prescription be given to construct an instrument which can be used to measure them.

Therefore, the term observable, often used in the mathematical formulation of QM, refers to an *ontological description*.

In this formalism, it turns out that in order to represent something one is forced to consider matrices which are complex-valued. Therefore, observables can be represented by matrices with complex entries. Observations always lead to a real number, and therefore, not all matrices represent observables but only those that are *hermitian*.

Notice that the algebra of matrices is not commutative and the product of two hermitian matrices need not be hermitian. In the mathematical formulation of the theory, it was convenient to consider the entire algebra of matrices and not only the "observable" ones.

The theory of matrices developed by Born, Heisenberg, and Jordan gave a rational basis to very many problems in atomic physics.

At the same time, it made clear that classical mechanics and dynamics were not a valid instrument for the description of the structure of the atomic world, in spite of the fact that the description of the atomic reactions was given in classical terminology. Indeed it was heavily stressed by N. Bohr that all relevant information can only be transmitted and received within the formalism of classical mechanics.

Roughly at the same time, E. Schrödinger developed *wave mechanics* adapting ideas of L. de Broglie.

L. de Broglie had remarked that a *quantal fraction of reality* seemed to have a particle-like behavior in some experimental instances and a wave-like behavior (diffraction) in others.

Following an analogy with the variational principles of classical mechanics both in the lagrangian and hamiltonian formulations, he proposed that a quantal fraction of reality be described by a complex wave. It was complex because the analogy with classical (hamiltonian) mechanics required the introduction of a "momentum" and by the formalism of the Fourier transform a fraction of reality with definite momentum was represented by a monochromatic wave and therefore complex-valued.

Schrödinger adapted the de Broglie formalism to describe the structure of the atoms, in particular the hydrogen atom. He modified the wave equation proposed by de Broglie to take into account that the dynamics is nonrelativistic, and he proposed an equation which replicates the hamiltonian structure of classical mechanics and takes into account the intuition of de Broglie, i.e., that plane waves should be representative of sharp values of the momentum.

The Fourier transform gives the relation between a representation in which the position has *a relevant role in the description* and a representation in which *the relevant role* is given to momentum.

Later Born postulated that the square of the modulus of the wave gives the density of the probability that at a given time the system is in a spacial configuration and that the square of the modulus of its Fourier transform gives the density of the probability that the system is in a specific momentum configuration.

A great success of the analysis of Schrödinger was the proof that his time-independent equation gave exactly the energy levels found for the hydrogen atom. It was proven later by W. Pauli that this result can be obtained by purely algebraic analysis.

For more complex atomic structure, the analysis is not so simple and requires several approximations and estimates. Still the analysis of the atomic spectra through a solution of the Schrödinger equation led to excellent results (and also to interesting developments in the theory of partial differential equations).

With de Broglie and Schrödinger begins the formulation of quantum mechanics as *wave mechanics*; immediately after it was proved by several people (including Schrödinger himself) that the two formulations are equivalent (as a theory of operators in a separable Hilbert space).

Schrödinger's formulation employs the representation of the Hilbert space as a space of square-integrable functions over a measure space; Heisenberg's formulation chooses a specific orthonormal complete basis and looks at the operators as "matrices" in this basis.

Of course separable Hilbert spaces are isomorphic, and therefore, one can represent QM in any other realization of the Hilbert space, for example in the realization as square-integrable functions on the unit interval of the real line. But in this representation, the operators corresponding to simple physical quantities (e.g., position, momentum, and energy) have a very complicated presentation.

A major role in the theory is played by the rule that describes the value of an observable when the system is in a given state.

The structure of the formalism suggests that this operation be linear in the observables but sesquilinear in the wave function that represents the state, in order to ensure that the result of any measurement is a real number. This is formalized in Born's rule, which we will discuss shortly.

This leads, as Schrödinger immediately remarked, to a characteristic feature of the theory, the *superposition principle* which is better explained in the Schrödinger representation.

If ϕ and ψ are unit vectors in the Hilbert space, $\Phi = (\phi + \psi)/|\phi + \psi|$ is also a unit vector but in general $(\phi, A\phi) + (\psi, A\psi) \neq (\Phi, A\Phi)$.

There is therefore *interference* between the waves. This phenomenon is very well known for water waves and for electromagnetic waves, but here it is counterintuitive since the waves are *probability waves*.

The superposition principle has a more involuted description in the formalism of Heisenberg, but the two formulations are equivalent, and therefore, the superposition principle also affects this representation. Independently of the formalism chosen, this implies a relationship between states that cannot be explained in terms of familiar classical physical concepts.

Some researchers tried to escape the problem posed by the superposition principle by searching for *hidden variables* (the wave functions does not fully describe the state of the system).

The naive hidden variables theory turned out to not be practicable because they lead to inequalities (Bell's inequalities) that are disproved by experiments.

A totally different way out is the pilot wave theory (also named Bohm theory), which was initiated by de Broglie himself. It is a non-Newtonian theory in which particles move under the action of a vector field (pilot wave) which itself satisfies a Schrödinger equation defined on the configuration space of all the particles.

It is a theory of particles and not a hidden variable theory because the (point) particles are fully described by their position and momentum. It is not a Newtonian theory because no reference is made to forces; the motion of the particles is ruled by a vector field *which is not associated to the particles*.

Its presence is perceived only through its action on the particles. The evolution of the field is *independent of the particles* and is determined by a Schrödinger equation on the configuration space of all the points.

We shall not further discuss this interesting theory (which is also mathematically difficult as the vector field is singular) which, as remarked before, has its own interpretation problems.

3. Ontology and epistemology: the quantum omelette

Bohr's reaction to the difficulty of describing the ontological content of quantum mechanics *was to abandon the physical representation of quantum mechanics*, i.e., regard the formalism of quantum mechanics as *solely epistemological*.

There is a difference between discussing *what reality is* and *how humans acquire knowledge from experience*.

An ontological question is a question about the nature of existence and reality. It presupposes the existence of reality *and the possibility to represent it*.

An epistemological answer is related to the way in which humans connect to external reality. It is not regarded as a ground or goal of *understanding*. Its task is to describe how humans relate to experience. Physical theories are *economics of human experience* (Mach).

From this point of view, problems are not *out there* but are part of a definite viewpoint with definite *metaphysical assumptions and presuppositions* without which they cannot even be stated.

According to Jaynes [1990], our present the quantum-mechanical formalism is not entirely ontic but at the same time not entirely epistemological. It is a peculiar mixture describing in part nature and in part incomplete human information about nature, *all scrambled up* by Heisenberg and Bohr into *an omelette* that nobody knows how to unscramble.

Unscrambling the quantum omelette is a prerequisite for any advance in basic interpretation of the theory.

Quantum mechanics, in its Copenhagen interpretation, makes a process end with a choice, codified by Born's probability rule.

But if we take an ontological point of view, there can be no choice which determines what reality is: a subject cannot define by a choice, *within a physical representation of a theory*, what is physically real. Physical reality can be represented in an objective manner only if the subject (the experimenter) *plays no role within that representation*.

Einstein showed the inconsistencies with respect to physical reality (as understood in classical physics, the setting in which experiments and their outcomes are described) in which QM had been drawn through Bohr's complementarity approach.

What is considered *physically real* according to quantum mechanics?

Bohr's epistemological approach escapes ontological debates. Bohr's explained *how things had to be done*. Following this set of rules, one could recover from QM a rational account of classical phenomena [Bohr 1935; 1963].

When asked whether the quantum theory can be considered as somehow mirroring an underlying quantum reality, Bohr declared, "There is no quantum world. There is only an abstract quantum physical description" [Petersen 1963]. It is wrong to think that the task of physics is to find out *what nature is*. Physics concerns what we *can say about nature*.

Later he wrote, "Physics is to be regarded not so much as the study of something a priori given but as *development of methods of ordering and surveying human experience* which can be unambiguously communicated *in ordinary human language*".

Bohr always stood on the epistemic side and never discussed questions related to the ontology of the quantum realm.

It was Heisenberg, and later Born and Pauli, who, when stressing the successes of the new theory, incoherently mixed the epistemological complementarity scheme of Bohr with an ontological (Platonistic) approach which assumed a direct relation between the mathematical formalism and reality itself.

This led to the cooking of the quantum omelette.

This quantum-unrealistic position was consolidated at the Solvay conference (1927) and is now part of what every physicist learns and practices. It is the conceptual background of all the brilliant successes of QM in atomic, nuclear, and solid-state physics over the past ninety years.

Physicist have learned to think about the theory in a highly unrealistic way, to be at ease with wave functions and operators. This (unrealistic) way has brought about the most marvelous predictive successes in the history of science.

The triumph of this approach is exemplified by the fact that the *Copenhagen interpretation* is taught in all universities around the globe, while the mathematically equivalent pilot wave theory of de Broglie is seldom taught and is considered a curiosity.

Our students manipulate and draw wave functions as if they had an objective reality.

4. Copenhagen quantum mechanics

The quantum theory of Planck and Bohr was the basis on which one had to construct *a new theory to describe the dynamics of atomic systems*.

Let us recall that a mathematical model, according to J. von Neumann, is a *mathematical construction* that, supplemented by a *verbal language of correspondence*, provides a coherent basis for the description of a class of physical phenomena.

A model originates from a combination of experimental evidence, theoretical analysis, and mathematical analogies. A model obtains the status of a theory on the basis of the amount of physical phenomena it helps to organize in a coherent way.

When the class of phenomena described by a model in a somewhat unified way covers an entire field of physics, one speaks of *a theory*. A theory in general provides a different perception of what is relevant, both conceptually and from the point of view of the experiments, a different *paradigm*.

The passage from a *model* to a *theory* is also conditioned by cultural background, versatility to adapt to applications, and also prejudgements.

A theory indicates what experiments are worth performing and what the questions that can be meaningfully asked are.

Form this point of view, QM deserves to be called a theory; it has changed our perception of the world at the atomic scale, providing a unified physical and mathematical picture, is at the basis of new technology, and has stimulated the development of a relevant part of modern mathematics.

When a theory reaches acceptance by a majority, it tends to dismiss as "false" or "irrelevant" any other attempt to construct an alternative model. Researchers working in QM tend to dismiss alternative theories as *irrelevant* and *mental constructions* and have the tendency to dismiss as *futile* the research on the foundation of the theory.

The consensus that comes from extraordinary successes is taken as a sign of truth.

Still there are conceptual problems that come partly, as N. Bohr emphasized, from the fact that *the language which is used is borrowed from classical physics* (N. Bohr went to the extreme of stating that classical physics *is necessary* to describe quantum mechanics) and partly from the difficulty of reconciling the intrinsic probabilistic aspects of quantum mechanics with the deterministic features which we are used to associating to physical phenomena.

For these reasons, quantum mechanics is a theory which is mathematically self-consistent and very effective in its application, but *not conceptually complete*.

We have stated that the mathematical construction of a model (and of a theory) requires

- stating axioms (or postulates), in general derived from phenomenology and from some historical and cultural background (the structure of previous successful theories),
- deduction of some nontrivial consequences, typically under the form of theorems and equations, and
- determining a *verbal language* which associates the mathematical structures to measurable quantities; this empirical description is in everyday language and links the theory to experimental data.

In classical physics, the mathematical constructions are, e.g., the variational principles, the equations of hamiltonian or lagrangian dynamics of material points, and the equations of the dynamics of the continuum and the equations of electromagnetism as formulated by Maxwell.

The (scientific) verbal language, i.e., the correspondence between mathematical entities and quantities that can be measured, is *given for granted* in classical physics; this common agreement is a result of centuries of "experience" and is also due to the fact that we have *a daily experience of the classical world.*

No one doubts the objective meaning of terms such as *measure of a velocity* or *measure of a magnetic field*, and we regard the result of a measurement as independent of the experimenter and of the apparatus used.

For the phenomena at the atomic scale as described by QM, this *objectivity* fails and *the very concept of measurement* can become problematic.

One can try to overcome this problem by stating that macroscopic objects, such as a measurement apparatus, must be regarded as classical objects, obeying the laws of classical physics.

But this would divide the physical world into two separate incompatible parts, and it would be difficult to make precise each time to which world one refers.

Many efforts have been made to solve the *measurement problem* (i.e., the detailed description in QM of the process of measurement), and various mechanisms have been proposed to explain why in (most) macroscopic bodies one does not perceive the typical structures of QM.

In particular, they explain why it is difficult to perceive (outside specialized laboratories) the *superposition principle and the entanglement*. Some of these attempts have led to a better understanding of the conceptual structure of the formalism and of its interpretation, but a satisfactory answer has yet to be found.

Let us stress again that from the empirical point of view QM has had outstanding success in organizing, describing, and also in some cases *predicting* results of experiments in its range of validity, namely the (nonrelativistic) physics of atoms and molecules and their aggregates.

At the same time, the refinement of the formalism of QM has contributed greatly to the development of modern mathematics.

Still it must be remembered that this theory has its own *range of validity*, in particular that it is nonrelativistic and it is not applicable to phenomena which occur at very high energies.

5. States and observables in QM

We review the basic structure with a mathematical description of the concepts of *states and observables*.

In general terms, a state of a system is the result of a *preparation procedure*. In order to construct a model, one must think of an idealized procedure that results in a well defined state.

One may think that the definition of a state is such that when a system is in state "A" any experiment gives the same result if performed by different observers.

This is true in classical physics; a measurement can be done, at least in principle, without altering the state of the system. And one can describe the state of a composite system by separately describing its parts. To give an example, the "state" of the solar system is described by giving the position and velocity of the single planets.

This has led those in classical mechanics to consider as elementary states *the points in phase space* $\mathcal{M}$.

An observable is characterized by the value it takes on each state, i.e., by a real-valued function on the phase space: if $m \in \mathcal{M}$ and if f is continuous, the number $f(m)$ represents the result of the measurement of the observable described by f when performed on a system in the state described by m.

From a mathematical point of view, therefore, the states in classical mechanics are elements of the dual of the space of continuous functions, the duality given by $\{m, f\} \to f(m)$.

The meaning of the word *measurement* and the role of the measurement apparatus are not discussed further; their definitions are considered clearly established and universally accepted. And it does not depend on the observer.

In classical statistical mechanics, one also introduces more general states represented by positive measures μ that are absolutely continuous with respect to Lebesgue measure.

These states are linear positive functionals on essentially bounded functions; as a consequence one can include in the theory a larger class of observables, i.e., functions in $L^1(\mathcal{M})$.

As can be seen from this brief reminder, the definition of *pure state* in classical mechanics is linked to the possibility of considering continuous functions as *observables*.

Dynamics is given by means of differential equations for functions in phase space which are required to be *differentiable*.

When one tries to develop QM and its dynamics keeping some analogy with Hamiltonian mechanics, the first problem one faces is that in QM an equivalent of phase space does not exist, and therefore, it is difficult to decide a priori how to describe a pure state and characterize an observable.

This problem is solved differently in the two basic formulations of quantum mechanics, which we shall denote *Schrödinger QM* and *Heisenberg QM*.

Schrödinger quantum mechanics. In the formulation of quantum mechanics due to Schrödinger, the primitive elements are the (pure) states which are represented by (normalized) vectors in a separable complex Hilbert space (for one particle, $\mathcal{H} \equiv L^2(\mathbb{R}^d)$) where d is the number of degrees of freedom of the corresponding classical system.

This interpretation makes explicit use of the analogy between $|\phi(x)|^2$, $x \in \mathbb{R}^d$, and the classical Liouville distribution $\rho(x)$.

In Schrödinger's formulation, the observables are *a dual structure*; they are represented by operators on $\mathcal{H}$.

Since $\mathcal{H}$ is concretely represented as $L^2(\mathbb{R}^d)$, the observables are represented by operations on functions, typically by multiplication by another function and by differential operators.

In view of the analysis done by de Broglie, the operator $-i\frac{\partial}{\partial x_m}$ can be identified with the momentum of a particle. Also real functions of momentum space represent observables. For these observables, one can expect to define a possible measurement procedure.

Dynamics of the states is given by the Schrödinger equation; dynamics on the observables is defined by duality.

Since the set of operators which are sums of a function of position and a function of momentum is not invariant under time translations, in order to be able to describe the dynamics, *one is forced to increase the number of observables*.

If we require the average value of any observable in any state to be a real number, we should restrict ourselves to hermitian operators (more precisely to self-adjoint operators in order to have a functional calculus).

One is therefore led to state that the observables are in *one-to-one correspondence* with self-adjoint operators, in spite of the fact that for a generic self-adjoint operator one is not able to exhibit the experimental apparatus which may be used to measure the observable it represents.

According to Born's rule, $\int_{\Omega\subset\mathbb{R}^3}|\phi(x)|^2\,dx$ represents the probability that, performing a position measurement of a particle in the state described by $\phi(x)$, the outcome is that the particle is localized in the region Ω.

This implies that, if the observable A is represented by the function $A(x)$ in configuration space, then

$$(\phi,\, A\phi) \equiv \int |\phi(x)|^2 A(x)\,dx$$

is the average of the results one obtains if one measures the outcomes of a measurement of A.

In the same way, if the observable B is represented by the function $B(p)$ in momentum space, then

$$(\hat{\phi}, B\hat{\phi}) \equiv \int |\hat{\phi}|^2(p) B(p)\, dp$$

(where $\hat{\phi}(p)$ is the Fourier transform of $\phi(x)$) is the average of the results of the measurements of the observable B.

By polarization, one obtains the value of $(\phi, A\psi)$ for every observable A and for any pair $\phi, \psi \in \mathcal{H}$.

Heisenberg's quantum mechanics. In the formulation given by Born, Jordan, and Heisenberg (matrix mechanics), the primitive elements are the matrices that give the probability of transition from an atomic state u_n to another state u_m under the influence of an external field or under spontaneous decay

This leads one to consider as basic elements in the theory the observables represented by infinite matrices, i.e., linear operators on a separable Hilbert space.

The structure of the states plays a lesser role in this formulation of quantum mechanics. They are considered the result of an initialization and are distinguished by means of the value they give for the expected value of the observables.

This correspondence is linear for the matrices that represent observables, and so the states are linear functions of the observables continuous in a suitable topology, i.e., elements in a dual space.

Interference effects are not easy to describe in the Heisenberg formalism. A *concrete* analysis of entanglement and interference without reference to the Schrödinger representation is difficult.

On equivalence. We have already seen that the two representations are equivalent in the mathematical sense and correspond mathematically to dual structures.

The mathematical instruments used are different: mostly algebraic in Heisenberg's presentation and mostly function-theoretical in Schrödinger's.

A bridge between the physics of the two formulations of QM is given by *Born's rule*, which we will describe soon as an axiom.

In this rule, the states and the observables play equally important and symmetric roles.

But Schrödinger's formulation has special properties that come from the fact that one was naturally led to use the representation of the Hilbert space as (square-integrable) functions on the *configuration space*.

This special presentation introduces spurious elements which make QM *more visualizable in space* (one draws on the blackboard the *shape of the wave function*) at the price of introducing misunderstandings.

Indeed the wave function is not a measurable quantity; it is rather an abstract instrument which can be used to determine probabilities of real events.

Still some remnants of the visual picture survive. The use of periodic cells introduces homology, and smooth functions can belong to different homology classes.

If, as usually assumed in regular crystals, the wave functions are coherent over many cells, these homological properties are inherited by the state of the crystal and may influence the expectation values of specific observables.

Therefore, the homology class of a wave function (a priori an abstract object) may be measurable; indeed the topology of the wave function is at the root of the use of geometrical and topological methods in solid-state theory.

As a consequence, one must be prepared to recognize that there are "geometric" properties of the wave functions which correspond to measurable quantities.

6. The axioms

After these preliminaries, we can now state the axioms of quantum mechanics. In choosing the order of the axioms, we shall follow the point of view of Schrödinger.

Axiom 1. Pure states are represented by unit vectors in a separable Hilbert space $\mathcal{H}$. Vectors that differ by a phase represent the same pure state.

It follows that the (pure) states are represented by projection operators $P_\phi = |\phi\rangle\langle\phi|$ (in Dirac's notation).

Since a Hilbert space is a vector space, the *superposition principle* holds: if $\phi, \psi \in \mathcal{H}$, then also $a\phi + b\psi \in \mathcal{H}$ for $a, b \in \mathbb{C}$. We assume here that all Hilbert space vectors represent states, i.e., there are no "superselection rules".

Also in quantum mechanics, one can introduce nonpure states, called statistical mixtures. They are represented by sums of projection operators

$$\sigma = \sum_n c_n P_{\phi_n}, \quad c_n > 0, \ \sum_n c_n = 1.$$

Positive-trace class operators with trace 1 are called *density matrices*. Their relation with the pure states *is the same as in classical mechanics*.

Contrary to what happens in classical mechanics, *no pure state is dispersion-free for all observables*; this is due to the fact that their dual, the algebra $\mathcal{B}(\mathcal{H})$, is not commutative.

Recall that the dispersion of a state σ relative to a (symmetric) operator A is

$$\Delta_\sigma(A) \equiv \sigma(A^2) - (\sigma(A))^2.$$

A state σ is *dispersion-free relative to* A if and only if $\Delta_\sigma(A) = 0$.

For comparison, notice that in classical mechanics, where the role of $\mathcal{B}(\mathcal{H})$ is taken by continuous functions, *all pure states are dispersion-free with respect to each observable*.

The dual of the pure states, under the duality given by P_ϕ, $A \to \operatorname{tr} P_\phi A$, is $\mathcal{B}(\mathcal{H})$, the set of bounded closed operators $\mathcal{H}$.

Axiom 2.

- The observables in quantum mechanics are represented by the self-adjoint operators on a separable (complex) Hilbert space $\mathcal{H}$.
- The *mean value* of the measurement of the observable represented by the self-adjoint operator A in the state represented by P_ϕ (the projection operator on the one-dimensional subspace spanned by the vector ϕ) is given by

$$\langle A \rangle_\phi \equiv (\phi, A\phi) \equiv \operatorname{tr}(A P_\phi)$$

where the symbol tr stands for *trace*, a function defined as usual for finite-rank matrices and extended by sum convergence in the case of infinite matrices.

Notice that in this in the formulation of Born's rule we assumed that the *measurement* is an abstract procedure that requires no further analysis. In this respect also Axiom 2 is ontic.

Notice that if $A \in \mathcal{B}(\mathcal{H})$ the correspondence

$$A \to \operatorname{tr}(\sigma A)$$

defines a linear continuous functional on $\mathcal{B}(\mathcal{H})$.

Therefore, we could start, as in the Heisenberg point of view, with the definition of the observables as the real part of the algebra of all bounded operators on a complex Hilbert space $\mathcal{H}$ and consider the states as derived quantities by duality.

Axioms 1 and 2 describe the *mathematical content of QM* (and also its metaphysical content).

We introduce now two axioms that represent the *verbal part of the model*, i.e., the rules which must be used to associate measurable quantities to the mathematical entities in Axioms 1 and 2.

Axioms 3 and 4 connect the mathematical formalism to the outcome of laboratory experiments and are therefore *of epistemic character*.

A step in this direction has already been made in Axiom 2 by Born's rule, but nothing has been said so far about the description of a *single measurement*.

In particular until now we did not speak of the effect that has on a state described by σ, the measurement of an observable a described by an operator A.

Axiom 3. Let the operator A describe the observable a, and assume that A has purely discrete simple spectrum, i.e., the eigenvalues are different from each other and the eigenfunctions ψ_i^A form a complete orthonormal basis.

If one performs a measurement of the *observable a* in a state represented by a vector $\phi \in \mathcal{H}$, $|\phi|_2 = 1$, the outcome can be only one of the eigenvalues a_k of A. The *probability* of the outcome a_k is $p_k^A = |(\psi_k^A, \phi)|^2$, where ψ_k^A is the eigenfunction of the operator A associated to the k-th eigenvalue.

We remark that this statement is *compatible* with Born's rule. Indeed from Axiom 3 it follows that the average of the results of the measurements of a when the state is described by the vector ϕ is $(\phi, A\phi) \equiv \mathrm{tr}(P_\phi A)$ where we have denoted by P_ϕ the orthogonal projection on the vector ϕ.

For observables which are represented by operators with partly continuous spectrum, the formulation of Axiom 3 is slightly more complicated; we don't detail here the obvious modifications. Axiom 3 is probably too ambitious as formulated. Given a generic symmetric bounded operator A, it is difficult even in principle to give a prescription for the construction of a measuring instrument which measures the observable associated to A.

For example it is difficult to indicate the instrument that measures the observable associated to $\xi_\Omega \hat{\xi}_\Sigma \xi_\Omega$, where ξ_Ω is the operator of multiplication by the indicator function of the domain Ω in configuration space and $\hat{\xi}_\Sigma$ is multiplication by the indicator function of the domain Σ in Fourier space.

Axiom 3 refers to the *possible results* of a measurement and to the *probability* with which they are obtained.

There is no indication of the effect of the state of the system after measurement.

In classical mechanics, it is assumed that, at least in principle, it is possible to perform measurements on a system *without altering its state*. In quantum mechanics, *this is not possible*. The interaction with the measuring apparatus *alters in general the state of the system in a way that cannot be predicted*. But one assumes:

Axiom 4 (projection postulate). If a_i is a nondegenerate eigenvalue, with eigenfunction ψ_i^A of the operator A associated to the observable a, and if the measurement of a has given a_i as a result, *immediately after the measurement*, the state of the system is described by the vector ψ_i^A.

The formulation *immediately after*, although imprecise, takes into account the fact that the operator A may not commute with the hamiltonian and therefore the eigenstates are not invariant in time.

Since the evolution under the Schrödinger equation is a continuous process in time, this effect is negligible if the time elapsed between two measurements is negligible.

We notice that Axiom 4 is needed to give *objective meaning* to the measurement process, i.e., the measurement codifies an objective property of the system *after the measurement*.

But Axiom 4 has far-reaching consequences since the algebra of the observables is nonabelian in quantum mechanics.

Suppose that at time $t_1 > t_0$ we perform a second measurement now of the observable b associated to the operator B which does not commute with A and has eigenvectors $\{\psi_k\}$.

According to the rules of quantum mechanics, we obtain the result b_k with probability $|(\phi_1, \psi_k)|^2$; if the result is b_k, we conclude, by Axiom 4, that the system immediately after the new measurement is in state ψ_k.

Now we perform again a measurement of A at time $t_2 > t_1$. The result will be a_h, $h \neq 1$, with probability $|(\psi_h, \phi_1)|^2 < 1$: the system *has a finite probability to be in a state different from* ϕ_1.

This implies that it is impossible to determine (even approximately) the state of the system *if one does not have complete control of the environment*.

For comparison, notice that in classical mechanics all observables (i.e., all functions on phase space) *take at any given time a definite value on all pure states*.

7. The semiclassical limit

The problem of finding a path relating the classical description and the quantum one is an *ontological problem* which aims to provide a *physical explanation* of what the relation between the classical and quantum realms is, *both of which are presupposed to be physically real*.

This problem cannot be understood in epistemological terms alone because there is no reference of the theory to something happening *within physical reality*. It is occasionally referred to as the problem of the quantum-to-classical limit.

Apart from the mathematical formulation which we shall outline shortly, the limit should be described in terms of a physical representation since it tries to explain what is going on beyond abstract mathematical formalism.

Therefore, the path must be represented in physical terms. The question analyzed from the point of view of Bohr's interpretation is ill posed: how can we possibly argue that there is a limit that can be explained with physical reality?

The quantum-to-classical limit in quantum mechanics is not so much a mathematical problem which seeks to relate incompatible formalisms but rather *a physical problem* which should provide a physical explanation for the connection between these seemingly incompatible descriptions.

The search of the physical explanation for the path from the classical to the quantum (or conversely) is *a strictly representative enterprise* [Jaynes 1990].

In some ways, this problem is similar to the problem of the description of a constrained system in mechanics. On one side is the classical microscopic point of view (world), where all systems satisfy the classical equations of Hamilton and Maxwell.

On the other side is the world in which constraints are considered as objective elements, worthy of a classification and an explicit (physical) description.

In this case, within classical dynamics, the connection between the two *physical* worlds is made through approximations, stating the physical approximation scheme as carefully as one is able.

The same view is taken in classical mechanics about statistical mechanics or thermodynamics, which have their own strict internal rules. It is suggested that there is a bridge to the world of classical mechanics; the bridge is made of approximations and changes of scales.

One may wonder whether there is a comparable *practical bridge* between the classical world and the quantum one.

Mathematically one can construct such a bridge provided one is willing to play with Planck's constant $\hbar$, which appears in the quantum-mechanical formalism.

Since Planck's constant has a well defined value (in suitable units) in our physical world, any such limit must be seen *as a mathematical exercise*.

One can prove mathematically that in a very precise sense *two classes of solutions* of the Schrödinger equation have as its limit when $\hbar \to 0$ *two classes of solutions* of classical dynamics, of Liouville and Hamilton type, respectively.

Generally speaking, the initial data should be *well localized* both in configuration space and in momentum space, compatibly with the rules of the Fourier transform (semiclassical wave packets) or otherwise much more localized in momentum space, and depend smoothly on the spacial coordinates (WKB states)

Remark that due to the structure of the Schrödinger equation one can trade the smallness of Planck's constant for a large value of the mass.

As a consequence, the barycenter of a very massive quantum-mechanical body (e.g., the earth) moves in a gravitational field roughly in the same way as a classical point particle with the same mass. This *saves* classical celestial mechanics.

But in general the description of dynamics is entirely different in classical and quantum mechanics.

8. Principle of decoherence

Taking for granted that the description of the motion of a stone is totally different from that of very small bodies such as atoms and electrons, one may expect that this difference is due to the (relative) complexity of the stone and that the peculiar features of quantum mechanics, detectable at the level of atoms, is *averaged out* and therefore no longer relevant for the motion of the stone.

This would prove *a physical way* to connect the classical and quantum worlds and would imply that *for all practical purposes* quantum-mechanical bodies of *macroscopic size* can be correctly described by classical physics.

The first problem of this approach is that apparently it is not only the size of the body which determines whether it has a *quantum behavior*.

Quantum-electronic devices, e.g., superconductors, can have the size of a meter and still must be described by quantum mechanics. It is still debated whether macromolecules can be described with classical mechanics or whether for them the quantum-mechanical properties must be taken into account.

As for superconductors, their *coherence* (tendency to show a quantum behavior) can be seen mathematically as a result of their almost pure periodic structure. Mathematically this is translated into the fact that the wave function is correlated over many elementary cells and therefore the analytic and topological structure is stable over a long distance.

Amorphous materials are more subject to interference effects so that the quantum structure is averaged out over a relatively short distance and is not effective over a long distance.

This averaging out is at the basis of the *principle of decoherence* which is often considered as a solution to the problem of the quantum-classical divide.

We shall briefly describe this principle through a typical example. We shall see that, contrary to a statement which is often made, some traces of the quantum behavior remain and in particular this principle does not solve the measurement problem (intrinsic indeterminacy in the measurement process).

Decoherence should be regarded as a consequence of a continuous process of correlations between the quantum system under study and the environment.

From a mathematical point of view, decoherence is linked to *partial trace* or *conditioning*.

Conditioning in quantum mechanics has properties similar to those of the operation with the same name in classical probability theory, but one should notice that, contrary to what happens in classical probability theory, in quantum mechanics *complete information about the state of the system does not imply knowledge of the state of each component*.

Suppression of information relative to the environment should lead to writing *effective equations* for a relevant subset of the measurable quantities of the subsystem.

If the subset can be described in classical terms, we expect that these equations are the equations of classical physics. The structure of the interaction should determine the subset of observables for which this *reduction* is possible.

The dynamics that one obtains should describe the evolution of these observables independently of the evolution of the environment *for almost all its configurations* and for a sufficiently long time.

It must be said at the onset that this program, on the mathematical side, is still in its infancy, in spite of its extreme conceptual interest.

Roughly speaking, the mechanism of decoherence is as follows. Assume that the initial state of the total system *observed object + environment* is

$$\Psi \in \mathcal{H}_{\text{tot}} = \mathcal{H}_{\text{obs}} \otimes \mathcal{H}_{\text{env}}, \quad \Psi = \psi \otimes \phi,\ \psi \in \mathcal{H}_{\text{obs}},\ \phi \in \mathcal{H}_{\text{env}}.$$

If one measures an observable $A \in \mathcal{B}(\mathcal{H}_{\text{obs}})$, the mean of the values will be $(\Phi, (A \otimes I)\Phi)$ (we have introduced the natural immersion of $\mathcal{B}(\mathcal{H}_{\text{obs}})$ in $\mathcal{B}(\mathcal{H})$).

If the hamiltonian of the total system is H, at time $t > 0$, the measurement of A in the state Φ will give

$$(\Phi, e^{itH}(A \otimes I)e^{-itH}\Phi).$$

Due to the interaction between the two systems, there *does not exist* in general an operator $K \in \mathcal{B}(\mathcal{H}_{\text{obs}})$ such that for all $A \in \mathcal{B}(\mathcal{H}_{\text{obs}})$ and for a generic state Φ

$$(\Phi, e^{itH}(A \otimes I)e^{-itH}\Phi) = (\Phi, (e^{itK}Ae^{-itK} \otimes I)\Phi).$$

One can hope (maybe even expect) that, if the environment has a large number of degrees of freedom and the interaction is very weak, the interaction has *mainly* the effect of modifying the state of the system, making it appear as a classical Liouville state.

In this case, the *coherence* which is at the root of the superposition principle *is hidden* by the lack of control of the environment. This would provide the bridge between the quantum world and the classical one.

This description scheme has not yet been developed; only special cases have been treated rigorously, and only *strong qualitative arguments* have been given in sufficiently general cases.

Strong qualitative arguments have been given, e.g., to show that *decoherence* for a quantum system can be produced by the interaction with a large number of light particles, and in this case position variables emerge as a "pointer basis" (variables which have a classical behavior).

One finds an overview of these considerations, e.g., in [Robert 1998; Wheeler and Zurek 1983].

Arguments have also been given to describe the decoherence for a quantum system *in a thermal bath*, i.e., interacting with a large number of particles in thermodynamic equilibrium at fixed temperature. In this direction, an approach to the mathematical description of decoherence has been pursued in [Hornberger 2009] in the framework of the algebraic formulation of quantum mechanics.

In spite of these developments, there are still many unsettled points in a mathematical theory of decoherence.

9. Mechanisms of decoherence

As a possible mechanism of decoherence, consider the system composed of a case of a great number of quantum particles of very small mass ϵ scattering one after the other and independently off a quantum particle of mass 1.

Let $\phi(x)$, $x \in \mathbb{R}^3$, be the wave function of the quantum particle before collision, and denote by y the coordinate of the first light particle.

The dynamics of the first collision is given by the Schrödinger equation (in units $\hbar = 1$)

$$i\frac{\partial \phi}{\partial t} = H\phi, \quad H = -\frac{1}{2}\Delta_x - \frac{1}{2\epsilon}\Delta_y + V(y-x),$$

where Δ is the laplacian and the potential V is regular and compactly supported.

The initial state of the system is $\Phi \equiv \phi(x)\psi(y)$, and we are interested in

$$(e^{itH}\Phi, (A \otimes I)e^{itH}\Phi)$$

where A is an observable of the system which is represented by an operator with kernel $A(x, x')$.

Setting $\xi = x - y$ and $\eta = x + y/\epsilon$ and noticing that all laplacians commute, it is not difficult to see that e^{itH} has the form

$$e^{itH} = e^{i(t/2)\Delta_x} e^{i(t/2\epsilon)\Delta_\eta} e^{-i(t/2\epsilon)\Delta_y} e^{it(1/2M)\Delta_y + V_x(y)}, \quad V_x(y) = V(y-x).$$

Consider now the case in which the particle with coordinate y is very light (ϵ is very small), and set $t = \epsilon s$.

Keeping into account that $A \otimes I$ commutes with H_y^0,

$$(e^{itH}\Phi, (A \otimes I)e^{itH}\Phi) = (W_s\Phi, (e^{-i\epsilon s H_x^0} A e^{i\epsilon s H_x^0} \otimes I)W_{s,x}\Phi), \quad H_x^0 = -\tfrac{1}{2}\Delta_x,$$

where

$$W_{s,x} = e^{-i(s/\epsilon)(1/2)\Delta_y} e^{i(s/\epsilon)((1/2)\Delta_y + V_x(y))}.$$

When $\epsilon \to 0$, the operator W_s converges to the wave operator W_x for the scattering of the light particle off the heavy one with a potential $V_x(y)$.

One therefore has

$$(e^{itH}\Phi, (A \otimes I)e^{itH}\Phi) = \int a_t(x, x')\zeta_t(x, x')\bar{\phi}(x)\phi_t(x)\, dx\, dx'$$

where $a_t(x, x')$ is the integral kernel of $e^{-itH_1^0} A e^{itH_1^0}$ and $\zeta_t(x, x')$ is a positive-type function (i.e., $\int \zeta(x, x')\bar{f}(x')f(x)\, dx\, dx' \geq 0$ for all f).

The function ζ is given explicitly as

$$\zeta_t(x, x') = (W_{x'}\phi, W_x(t)\phi).$$

As a function of $u = x - y$ and $v = (x + y)/2$, for each value it reaches its maximum when $u = 0$.

Notice that $\zeta_0(x, x')\bar{\phi}(x')\phi(x)$ is the kernel of the density matrix before the interaction.

The function ζ_t has the following properties:

(a) $\zeta_t(x, x) = 1$.

(b) If $t > 0$, $|\zeta_t(x, x')| < 1$ if $0 < |x - x'| < K_t$ where the constant K_t depends on the initial datum and on the potential.

(c) $\bar{\zeta}_t(x, x') = \zeta_t(x'x)$.

Property (b) indicates that due to the interaction the integral kernel is slightly (ϵ is very small) more concentrated on the diagonal and slightly more spread out due to the dispersive properties of the Schrödinger equation.

After the interaction with the first particles, the kernel of the density matrix will be

$$\rho'(x', x) = \zeta_t(x', x)\rho(x', x).$$

Assume now that the successive interaction with the N particles of the environment are independent and take place at times $\{T_k\}$ with intervals of order ϵ^{-1} in such a way that it is entirely a sequence of independent events.

In this case, the modifications to the kernel of the density matrix can be considered as independent and after N collisions the kernel of the density matrix will be

$$\rho_N(x,' x)\bar{\phi}(x')\phi(x)\zeta_1(x', x)\zeta_2(x', x)\cdots\zeta_N(x', x).$$

Since all functions ζ_k, $k = 1, \ldots, N$, have the properties (a), (b), and (c), we conclude that if N is very large the kernel of the density matrix after the very many collisions is concentrated for all macroscopic times on the diagonal and therefore is represented by a classical Liouville measure.

It is not difficult to show that under these conditions the propagation of the heavy particle is described within a good approximation by the Lagrange equation in configuration space.

In the terminology of the theory of decoherence, this may be described as the choice of a *preferred basis* (the configuration space).

But notice that in general the support on the diagonal is not of the order $\sqrt{\hbar}$ and therefore the (approximate, classical) dynamics is that of a Liouville distribution.

The procedure we have described is largely heuristic, and to have a rigorous result one should proceed much more carefully, establishing exact formulae and giving accurate estimates of the terms which one neglects.

This is a very difficult task that nobody has completed so far.

10. Experiments on decoherence

From the experimental point of view, interesting and very refined experiments have been performed in particular by the group of S. Haroche at the École Normale Supérieure. A typical experiment is described in [Raimond 2014].

In this experiment, rubidium atoms (mice) initially in a circular Rydberg state f are injected one at a time in a photon box (cat), an open cavity with reflecting walls that can keep for 1 millisecond a specified number of photons of wavelength 6 millimeters.

The electromagnetic field in the cavity is prepared in a state of 9 or 10 photons; this state approximates reasonably well a coherent state for which a (semi)classical description is possible.

In particular one can define the *frequency* of the radiation field. In this state the cavity-cat is sleeping.

The field is prepared in a (quasi)monochromatic state with frequency resonating with that of the transition between the f and the g states of the rubidium.

The rubidium atom (mouse) which is injected in the cavity is in a state "f"; this state has small dipole momentum and therefore does not disturb the cat (change the number of photons).

While crossing the cavity, the atom-mouse is subjected to a laser beam to induce a transition to another state "g"; the superposition of the two states has a large electric dipole.

This produces emission and absorption of photons and modifies the distribution of the number of photons in the cavity-cat which is now *suspended* between two normal states f and g (and this makes its presence visible to the cat).

The environment is in this case represented by the walls of the cavity that "interact" with the field present in the cavity because of imperfections.

The number of photons in the cavity (the status of the cat) can be monitored by sending into the cavity a second rubidium atom. If the cat has remained in a suspended state, it will interact with the atom-mouse.

The apparatus therefore permits one to tell whether the cavity is in a coherent state and thereby permits one to measure the *decoherence induced by the environment* (by impurities in the walls of the photon box). The experiment reveals that the amount of decoherence depends on the delay with which the second atom is injected into the cavity.

Call *decoherence time* the time after which the description of the photon box as a coherent state produces a relative error of $1/10^4$. The results of the experiment which is roughly described above indicate that the decoherence time is on the order of 0.1 milliseconds.

Decoherence *is therefore effective* in a small time even in this carefully organized experiment.

One can expect that in the case of a real cat the survival time of coherence is several orders of magnitude smaller and therefore *coherence cannot be seen* under normal everyday life conditions (i.e., for true cats).

This and related experiments show that a result (the cat suspended between two *physical states*) that seems to be counterintuitive when one has little control of the environment can be observed in a laboratory in which maximum control is possible.

11. The measurement problem and tracks in a cloud chamber

The difficulty in unscrambling the quantum omelette is clearly shown in the measurement problem, which cannot be solved by the theory of decoherence.

As Bohr emphasized, measurement is done with classical instruments and the result is expressed in classical language.

One often says that the instrument *interacts* with the object to be measured, but there is an ambiguity in this statement.

In mathematical terms, the term *interaction* refers to the description of the dynamics by means of the structure of the equations. It belongs therefore to the world of mathematical quantum mechanics.

On the other hand, the word *measurement* (distinct from interaction) must refer to a process that takes place in the real world and leads to unambiguous results.

According to Bohr [Robert 1998], *the unambiguous interpretation of any measurement* must be essentially framed in terms of classical physics theories, and we may say that the language of Newton and Maxwell remains the language of physicists for all time.

Therefore, the analysis of the quantum measurement process is the key to recovering a rational account of physical phenomena.

We have stated that the main problem in QM is the distinction between *interaction* (a mathematical structure) and *measurement* (a physical process).

We exemplify these difficulties by considering a simple phenomenon, the occurrence of tracks in a cloud chamber. It is simple enough to admit an almost complete mathematical description, and at the same time it contains all the interpretation problems in QM.

A Wilson cloud chamber is a vessel that contains vapor which is in a supersaturated state. Under a small *local* perturbation, it can locally make a transition to a liquid state (droplet).

It is an experimental fact that an α-decay produces in a cloud chamber *at most one track* (sequence of liquid droplets placed on a line that is straight or slightly curved if a magnetic field is present (a trajectory of a classical particle)).

Different decay events produce tracks that point in random directions. This seems to contradict the description of decay in QM: according to Gamow if the decay takes place at rest, a spherical wave is produced and moves radially according to Schrödinger's equation.

The presence of the track makes the result appear *as if* the interaction with the supersaturated vapor turned the wave into a particle. We want to place this effect in the context of Schrödinger's QM.

Notice that in this experiment the *experimenter* is the cloud chamber (or rather the supersaturated gas).

The problem of *justifying the presence of a track of droplets* within Schrödinger's quantum mechanics goes back to the early days of QM together with the question of whether the presence of the track can be considered proof that a *real* α-particle is produced in the decay.

Mott [1929] was among the first to attempt a systematic description using properties of the solutions of the *time-independent* Schrödinger's equation. A rather detailed account of the history of the problem and of various attempts to find a solution can be found in a recent book by Figari and Teta [2014].

The analysis given by Mott is based on stationary and nonstationary phase techniques in the *time-independent* formulation of Schrödinger's equation; it goes in the right direction but is incomplete in several ways.

To improve the analysis, we rely on semiclassical theory [Dell'Antonio 2015].

We shall see that *the properties of the initial wave function* allow for the introduction of a semiclassical formalism in which the interaction of the wave *with a single atom* can be regarded as semiclassical inelastic scattering. We stress that this description does not have a universal character and *depends essentially on the mathematical properties of the initial state*.

In this mathematical formulation, the α-wave *before the production of the first droplet of the track* can be regarded (mathematically) as fragmented into (coherent) semiclassical wavelets, each of high momentum, moving radially away from the point at which the decay has taken place.

The linear size of each wavelet is comparable to that of the atoms. Each wavelet moves according to the laws of QM: its barycenter moves on a classical path (curvilinear if a constant magnetic field is present), and its dispersion is of order $\sqrt{\hbar}$ (in natural units) and increases slowly in time.

No physical significance should be attached to this mathematical exercise. One can compare it to the description of light as composed of *light rays* with a major difference: the α-wave *is a probability wave*.

The α-wavelets move coherently; the entire wave keeps its spherical structure in accordance to Huygens' principle.

When one of the wavelets interacts with an atom, the coherence with the other wavelets is lost: the combined system wavelet + ion (+ emitted electron) belongs now to a different subspace of the Hilbert space in which the entire system (emitted wave and atoms of the vapor) is described.

If one regards the resulting subsystem as isolated, the result is an entangled state of the wavelet, of the wave functions of an ion, and perhaps of the emitted electron.

The interaction wavelet-atom can be considered as independent of the environment, and mathematically it can be regarded as inelastic scattering.

Therefore, QM describes the system after the interactions *but before the production of the first droplet*, as a collection of very many coherent triples each composed of the wave function of an ionized atom, a semiclassical wavelet, and the wave function of an electron.

Each triplet belongs to a different sector of the Hilbert space, and there is no interference between them.

Note that *mathematically* the ionization of an atom is the result of *an interaction* described in QM by a unitary propagation within a huge Hilbert space.

On the other hand, the production of the first droplet is a *random macroscopic event*, the result of a chain of processes of magnification which can probably best be described within statistical mechanics. The local phase transition is due to the modification produced by the ion in the electronic structure of the nearby atoms. QM can at most be used to determine *the probability* that a droplet be formed.

Its relation with the Hilbert space description in QM is the *measurement problem*; how does the quantum-mechanical system choose the (probability) wavelet that produces the droplet? Is it *a random choice*?

After the production of the droplet, the remaining wavelets no longer enter the description of the system. It is improper to say that *they have disappeared* because as probability waves they had *no physical existence* even before the formation of the droplet.

The selection process is *probability preserving* (since the outcome occurs with probability 1) and *nonlocal* (since the initial wave function is extended and the final result is localized in a small cone). It is *not described* by the Schrödinger equation.

Since the interaction provides strong entanglement among the component of each of the triplets of probability waves indicated above, *one may assume* that the process of measurement selects not only an ion but also the corresponding α-wavelet, *although no actual measurement of the wavelet is done*.

After the interaction, the chosen wavelet is still a (semiclassical) probability wave. It can *interact* with the atoms in its path. Since the wave function of the

wavelet has support on the order of magnitude of the square root of Planck's constant, the interaction can be regarded by a macroscopic observer as having taken place at one point.

The momentum of the wavelet is *essentially* concentrated along a vector that is directed from the point of decay to the point at which the first droplet is formed.

The production of further droplets is again a macroscopic phenomenon not described by the Schrödinger equation. Since the interaction is local, the incoming (probability) wavelet is well localized and the exchange of momentum is negligible; its barycenter has essentially a classical motion.

The wavelet therefore moves as a classical object (an α-particle) with roughly the same energy and momentum of the wave emitted in the decay.

The question of whether *after the production of the first droplet* the remaining droplets in the track are produced by a probability wavelet or by a particle is devoid of objective meaning.

In any case the interaction of the wavelet-particle with the other atoms leads to the formation of a straight line of droplets (or a curved line if a magnetic field is present). Due to the semiclassical nature of the wavelet, the direction of the track is determined by the position of the source *and of the first droplet*.

Notice that after each collision *the shape* of the wavelet may change and from the point of view of mathematics at every interaction the wavelet changes sector in the abstract Hilbert space.

The quantum aspects of this description are limited to the fact that, *although the initial state is completely known*, one can give only *the probability* that a track is produced in a given direction.

We stress that one sees droplets *only if ionization takes place and the ion triggers the magnification mechanism*. Without this mechanism (which can be described at most by quantum statistical mechanics), the event is *not recognized* by the macroscopic observer as *a measurement* and it must be considered only as an interaction.

One therefore has to invoke the presence of a step in which *probability* is turned into *occurrence*. This step is beyond QM and has not been understood so far. This ambiguity is at the heart of the measurement problem in QM.

Notice that also from a bayesian point of view (updating of information) this is a difficult problem since the updating is (presumably) done by the supersaturated gas. On the other hand, critics of the description of a measurement process have always remarked *that a Ph.D. is not required to a make a measurement*.

12. Some mathematics

For completeness we give some details of the mathematics involved. More details can be found in [Dell'Antonio 2015].

According to quantum mechanics, the α-wave ϕ_0 produced in the decay can be presented as a complex-valued spherically symmetric function with support in a small neighborhood of the origin and with a phase proportional to $M|v_0||x|/\hbar$ where $M \in \mathbb{R}^+$ and $v_0 \in \mathbb{R}^3$ are the parameters (mass and radial velocity) which characterize the wave produced in the decay.

We take natural units in which Planck's constant $\hbar$ is very small, and we assume that $|v_0|$ is very large.

It is convenient to have a different (equivalent) *presentation* of the initial data as a function on the product of a small interval $I \subset \mathbb{R}^+$ and a fibered two-dimensional sphere S^2, with fibers perpendicular at each point to the sphere.

This presentation is particularly adapted to the introduction of a semiclassical structure since both the free evolution in time and the wave packets are obtained by the convolution of the wave function with a gaussian kernel.

For concreteness we shall write

$$\phi_0(x) = Ce^{-|x|^2/(2\hbar)} \int_{S^2} d\omega\, e^{iMv_0(\zeta(\omega),x)/\hbar}, \quad x \in \mathbb{R}^3,$$

where C is a normalization constant and $\zeta(\omega)$, $\omega \in S^2$, is the unit vector orthogonal to S^2 at the point ω and directed opposite to the center of the sphere. The wave is produced with high momentum, and therefore, we take $1 \ll v_0$.

If there are no interactions, the wave evolves according to the free Schrödinger equation; the evolution is described by the convolution with a suitable gaussian kernel.

Taking into account that $|v_0|$ is very large, it is easy to see that at a later time T the wave is approximately localized in a corona of mean radius v_0T and width on the order of $\sqrt{\hbar}$.

Therefore, up to a small error, the wave function $\phi_T(x)$ at time T can also be presented as a function on the product of an interval on the positive real axis and the fibered unit sphere.

We shall assume that the (mathematical) interaction of the wave with the atoms is of very short range and is nontrapping. This will allow us to consider the result of the interaction as an inelastic scattering event. If the atoms are sufficiently separated from each other, we can consider the interactions as *independent scattering events*.

The fact that several further ionizations are seen (forming a track of droplets) suggests that the interaction between the atom and the semiclassical wave is rather strong, and therefore, it is advisable to avoid using perturbation theory (a contact interaction may be a better choice).

The waves in the Schrödinger picture are probability waves and carry no objective reality; they are tools to give the probability distribution of the outcomes *if*

a measurement of a given observable is performed. Understanding the process by which this probability is turned into a specific outcome constitutes the *measurement problem* that is still unsolved in spite of its conceptual relevance.

A detailed description in quantum mechanics of the interaction of the emitted wave with the atoms in the cloud chamber is beyond reach. We therefore make some simplifying assumptions and approximations. We use natural units in which Planck's constant $\hbar$ is very small.

Before the interaction, the wave satisfies the free Schrödinger equation and therefore the solution at time T is given by the convolution of the initial data with a gaussian kernel. Under the assumption that Mv_0 is very large, the presentation we have used provides the following description of the wave at any time $T > 0$:

$$\phi_T(x) = F_T^{\hbar}(|x|) \int_{S^2} d\omega\, e^{iMv_0(\zeta(\omega),x)/\hbar} + R_T(x), \quad x \in \mathbb{R}^3.$$

The function $F_T^{\hbar}(\rho)$ is supported in a spherical corona of depth $\sqrt{\hbar}$ and radius $|v_0|T$. The residual term $R_T(x)$ has L^2 norm of order $\sqrt{\hbar}$ and decreases fast in time. We will neglect this term in the following analysis and will take the L^2 norm of ϕ_T to be 1.

We make use of natural units in which Planck's constant h is very small. The essential support of the wave function of an atom has linear size of order $\sqrt{\hbar}$. All quantities will be evaluated up to a relative error of order $\sqrt{\hbar}$.

One can consider separately the evolution of small fragments, *wavelets*, of the α-wave, of linear size $\sqrt{\hbar}$.

The density of the atoms in the cloud chamber is such that each fragment interacts with *at most one* of the atoms.

In the presentation of the α-wave given in (2.2), the fragments are obtained using elements of a smooth partition of the unit sphere. Each element ξ has support of linear size $O(\sqrt{\hbar})$, and its initial condition at time T is

$$\phi_\xi(x, T) = \Phi_\xi(x, T) + R_T, \quad \Phi_\xi(x, T) = F_T(|x|) \int_{S^2} \xi(\omega) e^{iMv_0(\zeta(\omega),x)/\hbar}\, d\omega.$$

$F_T(\rho)$ has support in a neighborhood of v_0T of linear dimension $O(\sqrt{\hbar})$, and we shall neglect R_T, which is smaller in norm by a factor $O(\sqrt{\hbar})$ with respect to Φ_ξ.

We prove now that the solution with initial condition $\Phi_\xi(x, T)$ is localized together with its (quantum) Fourier transform in a domain of linear size $\sqrt{\hbar}$ and therefore represents a semiclassical wavelet. We take ξ to be localized around the point $(0, 0, 1)$. By construction the function $\Phi_\xi(x, T)$ is then supported in a neighborhood of linear size $O(\sqrt{\hbar})$ of $(0, 0, v_0T)$. Recall that the quantum Fourier transform is the Fourier transform written in units of $\hbar^{-1}$. Notice that we make an

error of order $\hbar$ in substituting the support of ξ on the sphere of radius $v_0 T$ with its projection on the tangent plane. Up to an error of order $\hbar$, we can therefore write

$$\tilde{\Phi}_{\xi}(x, T) = \int \tilde{\xi}(y) F_T(|x|) e^{iM(v_0 x_3 + y_1 x_1 + y_2 x_2)/\hbar}\, dy_1\, dy_2$$

where $\tilde{\xi}(y)$ has support in $|y| < C_2\sqrt{\hbar}$.

The Fourier transform is easy to compute; it has support in the ball of radius $O(\sqrt{\hbar})$ and center $(0, 0, Mv_0)$. Therefore, under free evolution, $\Phi_{\xi}(x, T)$ behaves as a semiclassical wave packet. Since v_0 is very large, if the interaction is not trapping, the wave remains in the cloud chamber for a very short time Δ.

We can use this semiclassical picture during the time in which the interaction with the atoms takes place. Standard phase-space analysis shows that, if ϕ_T^{ξ} and ϕ_T^{η} have supports separated by a finite distance d, the same is true (up to an error of order $\hbar$) for a time $T \le t \le T + \Delta$.

If the phenomenon we describe were scattering of a semiclassical wavelet by a potential $V(x)$, regularity and no trapping properties of the potential would give a description of the event as semiclassical scattering. Regularity conditions on the potential must be imposed in order for dynamics to preserve the semiclassical structure. In our case the scattering is *inelastic* because the final state also contains an electron.

Consider first the interaction with a single atom with wave function ψ_Y with essential support in a neighborhood of linear size of the order $\sqrt{\hbar}$ of a point $Y \in \mathbb{R}^3$. As a result of the interaction, ionization occurs; we assume that the wave function of the resulting ion remains localized in a neighborhood of Y of linear size $O(\sqrt{\hbar})$.

We have assumed that the interaction is not trapping and its range is of order $\sqrt{\hbar}$. Since the speed with which the wave moves is very large, the interval of time Δ in which the interaction takes place is very short. Due to our assumption on the density of the atoms, we can assume that the fragmentation of the wave is such that during the interval of time Δ *only one of the fragments* interacts with the atom in Y. Under this assumption after the interaction, this fragment is localized again in a region of linear size $\sqrt{\hbar}$ near the atom in Y.

Since the momentum of the incoming fragment was localized around $Mv_0\hat{Y}$ in units of $1/\hbar$ and the loss of momentum in the interaction is very small on this scale (the ionization energy is comparatively small), energy-momentum conservation implies that also the momentum of the outgoing fragment is sharply localized around $Mv_0\hat{Y}$.

The outgoing fragment is therefore represented by a semiclassical wavelet, with approximately the same mean momentum as the incoming wavelet and approximately the same variance.

Consider an atom localized near the point $P \in \mathbb{R}^3 \equiv (0, 0, r_0)$. Under free motion

the wave function $F_T^\hbar$ overlaps the wave function of the atom for a very small time interval $\Delta_0 T$. Since the interaction is of very short range and nontrapping, the wave function with initial data $F_T^\hbar$ overlaps the wave function for a very short time. Therefore, only a small part of the incoming wave contributes to the interaction with the atom in P.

This suggests a (mathematical) decomposition of the incoming spherical wave into fragments (wavelets), each of which can interact with only one of the atoms. We will prove that, *due to the properties of the initial α-wave*, the wave can indeed be seen as decomposed into small (coherent) fragments (wavelets) each propagating as a semiclassical wave packet (its barycenter follows a classical path). The dispersion is of order $\sqrt{\hbar}$ both in space and momentum (the latter in units of $\sqrt{\hbar}$).

Notice that we are manipulating *mathematical objects* (probability amplitudes) that enter into the *mathematical framework* by which quantum mechanics describes *outcomes* of experimental observations.

Before the formation of the first droplet (and after very many interactions), the partition in wavelets is a mathematical exercise. The *macroscopic production* of the droplet selects one the ions (the seed for the production of the droplet). This selection process is nonlocal and *is not described by the Schrödinger equation*.

The measurement process also selects the wavelet associated to the ion (although one measurement is performed on it). The selected wavelet is still a probability wave. Momentum conservation together with the semiclassical approximation imply that also this outgoing α-wavelet can be treated semiclassically (but its shape may have been changed by the interaction).

The outgoing wavelet interacts with the atoms on its path giving rise to further ionizations. Each ionized atom is a seed for production of a droplet; this originates the track. All other probability wavelets *now have probability* 0; therefore, there is only one track.

As already mentioned, the *mathematical device* of partitioning does not lead per se to anything physical. It reveals a detectable phenomenon *due to the macroscopic mechanism of production of a droplet*.

Notice that in this description the semiclassical wave packet entangled with the selected atom may lead to measurable effects *although no measurement is performed on it*.

As a result of the measurement (a probability-preserving nonlocal map), one of the ions is selected and the corresponding wavelet acquires probability 1. *This distinguishes measurement from interactions*.

We now generalize the analysis to take into account that there are many atoms in the cloud chamber, uniformly distributed and sufficiently separated so that each of them interacts with at most one of the wavelets and the interactions can be regarded as independent.

Each wavelet interacts with at most one atom. The outcome of the interaction is an entangled state made of the wave functions of an ion and of the outgoing wavelet (and of an electron).

We conclude that the interaction of the α-wave with the atoms in the cloud chamber can be mathematically described *before the measurement* as a sequence of *disjoint and independent* interactions of semiclassical probability wavelets with the atoms of the cloud chamber.

After the interaction, the wavelets move incoherently and the wave functions of the atoms are turned into the wave function of an ion. The interaction time is so short that we are justified in substituting the interaction with the scattering map.

The act of measurement (the cloud chamber measures *the position of the first droplet produced*) selects, according to Born's rule, one and only one of the ions to be the seed of the process of formation of the first droplet of liquid. The exact mechanism behind this selection has not been understood so far; it can be best described within quantum statistical mechanics.

The ion selected modifies the wave function of the nearby atoms. Since the vapor is supersaturated, this gives rise locally to a phase transition with production of a liquid droplet.

We *assume* that the measurement process also has the effect of keeping, as part of the description of the system *after the measurement*, also the wavelet entangled with the selected ion (although no direct measurement is performed on the wavelet).

The wavelet which is selected may originate on its path further ionizations, and this gives the visible track. Notice that the (position) measurement *of the first droplet in the path* is the only one which is represented by a (unitary) *nonlocal transformation*. The process of production of the remaining part of the track is essentially local.

After the measurement only one of the wavelets enters in the description of the system; it has essentially the energy-momentum of the entire incoming wave. Since it is well localized in position, it can be described as a particle (the α-particle). The remaining (probability) wavelets *no longer enter the description of the system*.

In conclusion, the analysis we have performed of the production of tracks in a cloud chamber shows that *interaction* should not be confused with *measurement*.

We briefly note the relation of this analysis with the problem of *decoherence*.

Before the interaction with the atoms, the semiclassical wavelets were coherent. After the interaction with an atom, the wavelet is entangled with the wave function of an ion. The coherence with the rest of the wave is no longer detectable (it would require a detailed knowledge of the wave functions of the atoms and of the emitted electrons).

If the environment contains N atoms which are placed sufficiently far apart so that the interaction of the wave with each atom can be treated as independent, the

interaction produces *N mutually incoherent triples* each representing an entangled state of an ion, a wavelet, and the emitted electron.

This decoherence between the triples *is entirely different* from the decoherence of a slow-moving quantum wave as a result of *very many interactions* with the ambient space. Decoherence in the cloud chamber experiment is related to *a single interaction with an atom*.

13. Quantum mechanics: Born's rule as conditional probability and information-theory analysis

Combined with the projection postulate, Born's rule says that, when one knows that a measurement corresponding to an observable a associated to a symmetric operator A with discrete spectrum has taken place but one does not know the result, the following information is gained.

If the initial state is described by density matrix ρ_{in}, then the density matrix ρ_{fin} of the final state is given by

$$\rho_{\text{fin}} = \sum_i \text{tr}(P_i \rho_{\text{in}} P_i)$$

where $A = \sum_i \lambda_i P_i$ is the spectral decomposition of the operator A. One has by definition $\sum_i P_i = I$ and $A = \sum_i \lambda_i P_i$.

This formulation no longer requires that the initial state of the system be pure.

It is interesting to notice that the formula *can be interpreted in information theory* as saying that ρ_{fin} represents *the most probable state* that one may have after a measurement of the observable a in the state described by ρ_{in} [Vedral 2002].

We clarify what this statement means.

According to von Neumann, information is measured by relative entropy, and the most probable state is the state which corresponds to minimal entropy *relative to the initial state*.

Following Wiener, we consider the amount of information to be the negative of the quantity defined as entropy.

We therefore take the negative of the relative entropy function $D(\rho, \sigma)$ as a measure of the relative information about the quantum state σ that can be derived from knowledge of the quantum state ρ:

$$D(\rho, \sigma) \equiv \text{tr}(\rho \ln \rho - \rho \ln \sigma).$$

The function D is nonnegative and can be considered a *nonsymmetric* distance.

The most probable outcome state is by definition the state which minimizes the distance $D(\rho, \sigma)$ for all allowed σ.

In the case of measurement of the observable a which is represented by the operator $A = \sum \lambda_i P_i$, the reference states are the density matrices which belong

to the set

$$\Sigma_A \equiv \{\sigma : [P_i, \sigma] = 0\} \quad \text{for all } P_i.$$

The last equation is equivalent to the condition $[A, \sigma] = 0$, i.e., the requirement that the state is obtained as a consequence of the measurement of the observable a.

We must minimize $D(\rho, \sigma)$ over Σ. This amounts to selecting the quantum state that is least distinguishable from the original state *among all the states that satisfy the constraint of being produced by the measurement of* a.

We consider only the case in which the Hilbert space is finite-dimensional. The same results are obtained if A is compact.

The set Σ_A is defined by a linear relation, so it is a simplex.

$D(\cdot, \cdot)$ is jointly convex in both arguments so that $(\cdot, \rho)$ is convex for all ρ.

Since the problem is finite-dimensional, the following holds: if the function f is (Gateau) differentiable and strictly convex on a simplex, and the directional derivatives (we are in a finite-dimensional setting) at a point b are all 0, then b is the *global* minimum of f.

We can parametrize Σ_A noticing that every element is of the form

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

where Λ is a trace-1 matrix with positive entries and U is a unitary operator $U = \pi_i U_i$ where U_i is the identity on the range of $(I - P_i)$.

This means that $[U, P_i] = 0 = [\Lambda, P_i]$.

Therefore, writing σ_i for σ restricted to the range of P_i, we have for every function f on Σ_A

$$f(\sigma) = \bigoplus_i f(\sigma_i)$$

(i.e., functions act blockwise on Σ_A).

Consider first the variation along the directions parametrized by U.

We look for the variation in the direction parametrized by one-parameter subgroups. Call L the generator.

We then compute

$$\frac{d}{dt}\phi_t^* \operatorname{tr} AU|_{t=0} = \frac{d}{dt} \operatorname{tr}(Ae^{tL})U|_{t=0} = \sum_{i,j} A_{i,j} L U_{i,j} = \operatorname{tr}(ALU)$$

where we have denoted by ϕ_t^* the adjoint action. In the same way, one computes

$$\frac{d}{dt}\phi^* \operatorname{tr}(AUBU^*)|_{t=0} = \operatorname{tr}(ALUBU^*) - \operatorname{tr}(AUBU^*L).$$

It is easy to verify that

$$[L_i, P_j] = 0, \qquad L_i P_j = \delta_{i,j} L_i.$$

The derivatives take the form

$$\partial_{L_i} \sum_j \operatorname{tr}(P_j \rho P_j U_j \ln \lambda_j U_j \ln \lambda_j U_j^*) = \operatorname{tr} L_i [\ln \sigma_i, \rho_i].$$

If σ and $\sum P_i \rho P_i$ can be diagonalized simultaneously, the derivatives vanish.

This is also a necessary condition since the commutator $[\ln \sigma_i, \rho_i]$ is traceless and L_k and iL_k span the all-traceless matrices in the i-th block.

We must consider next the variation with respect to Λ restricting to the case when σ and $\sigma_i P_i \rho P_i$ can be simultaneously diagonalized.

Let μ_k^σ and μ_k^ρ be the eigenvalues of σ and of $\sum_i P_i \sigma P_i$.

If $\mu_i^\rho \neq 0$ and $\mu_i^\sigma = 0$, one has $D(\rho, \sigma) = \infty$ so that this cannot be a minimum.

One has

$$\partial_{\lambda_k^\sigma} - \partial_{\lambda_k^\mu} \sum_m \lambda_m^\rho \ln \lambda_m^\sigma = 0,$$

which implies that $\lambda_k^\rho / \lambda_k^\mu$ is independent of k.

So the ratio of the eigenvalues of σ and of $\sum_i P \sigma_i P$ is fixed. Since they are both of trace 1, they coincide.

It follows that the state $\sigma = \sum P_i \rho P_i$ is the unique minimum of the relative entropy, i.e., it is the *unique* state that is least distinguishable from the original state among all states which are compatible with the observation of the observable a.

The results was later generalized by Kostecki [2014], who proves that minimization of the (Araki) quantum entropy is equivalent to the Lüders rule (a rule for updating information)

$$\rho \to \frac{\sum_{j \in J} P_j \rho P_j}{\sum_j \operatorname{tr}(P_j \rho P_j)}$$

where J is a subset of a countable set corresponding to an orthogonal decomposition and $\sum_i P_i = I \in \mathcal{B}(\mathcal{H})$ (the spectral sequence of an operator A on $\mathcal{H}$) and the domain of this equation is restricted by the condition that $\sum_i \operatorname{tr}(P_i \rho) \neq 0$ (the measurement of A succeeded).

This result has a strong bayesian flavor, and this leads us to the next topic.

14. Quantum bayesianism (QB)

I close this paper with a brief discussion of *quantum bayesianism*.

This vision of quantum mechanics, put forward by C. Fuchs and A. Peres [2000] and then by C. Fuchs, N. Mermin, and A. Schack [Fuchs et al. 2015], explores further the path laid by Bohr.

The approach of these authors was to follow the Bayes–de Finetti interpretation of probability as updating of information in order to account for QM. They call this approach quantum bayesianism (in short QB-ism).

They agree with Bohr that *the primitive concept of experience* is fundamental for the understanding of QM, but contrary to Bohr, QB-ism explicitly takes the *subjective* view of probability stressed by Bayes and de Finetti; i.e., probabilities are assigned to an event by an agent and are *particular to that agent*.

These authors state that QM does not describe *physical reality*. It provides *every single agent* with an algorithm for computing probabilities for macroscopic events (such as detector clicks) that are consequences of the agent's interactions with his world.

The agent has in general no control of the reaction, and the result of the experiences leads to an *upgrading* of the picture and of the expectations.

Still, an important component of the agent's experience is the impact of the efforts of other agents to communicate in speech or writing their own experiences. Science is a collaborative effort to find, through individual actions and verbal communications, a model for what is common to all our constructed external worlds.

To reify the (common) external world is a sound strategy *for all practical purposes*, but when subtle scientific concept are at stake, such as *quantum state*, it pays to trace back our description to our experience of the external world.

In a letter to Sommerfeld, Schrödinger [2011, p. 490] already stated, "Quantum mechanics forbids statements about what really exists — statements about the object. Its statements deal only with the object-subject relation".

And Niels Bohr [1934, p. 18] added, "in our description of nature the purpose is not to disclose the real essence of the phenomena but only to track down [...] relations between the manifold aspects of our experience".

Failing to recognize the foundational role of personal experience creates puzzles and paradoxes.

This strict definition of the scope of quantum theory is the only one ever needed whether by experimenters or by theorists.

In a letter to Peierls, J. Bell commented, "One can learn quantum mechanics the way one learns how to ride a bicycle, without really knowing what one is doing. But it is impossible to make sense of either without taking into account of what people actually do with them".

QB-ism shares with the Copenhagen interpretation the statement that the *quantum state* of a system is not an objective property of that system but only a mathematical tool to think about it.

A fundamental difference is that QB-ism (like Bayes) explicitly introduces *each user* of quantum mechanics into the story, together with the world external to the user.

A measurement is any action that an agent takes *to elicit a response*. Given a measurement outcome, the quantum formalism guides the agent *in updating* the probabilities for subsequent measurements.

From this point of view a measurement does not, as the term unfortunately suggests, reveal a preexistent state of the system.

Quantum mechanics is a powerful tool that any agent can use to organize his own experience. That this tool is used with spectacular success is an important *objective* fact about the world we live in.

But quantum mechanics itself does not deal directly with the objective world: it belongs to our experience of that objective world.

This is entirely different from the standard versions of quantum mechanics; for example Landau and Lifshitz [1965, pp. 2–3] state, "By measurement [...] we understand any process of interaction between classical and quantum objects, occurring [...] independently of any observer".

Bohr renounces this extreme attitude, but still individuals enter the story *only as proprietors of a large classical apparatus*, and the apparatus objectifies the diverse family of users. Replacing the single user with the apparatus introduces the ill defined *shifty split* much criticized by J. Bell.

This is a split between *classical and quantum*, *macroscopic and microscopic*; the split is shifty because its location can be freely shifted.

Because the outcomes of the Copenhagen experiments are *classical*, they are considered automatically real. In this interpretation, words like *macroscopic* are used to indicate the *objective*, nonquantum character of the outcome of a measurement.

In QB-ism measurement has a broader meaning; every action constitutes a measurement, and every outcome is a private experience that can be communicated in classical terms. The famous story about Wigner's friend is transformed from a paradox to a basic dictum.

The Copenhagen school holds that a quantum state encapsulates our knowledge. QB-ism replaces *knowledge* with *belief*, the belief of the person who made a state assignment, the belief of the implications of further experiences.

An important difference is the meaning of "*with probability* 1". As in the theory of Bayes, it reflects only the willingness to accept bets. It does not imply the existence of *a deterministic mechanism*.

This point was made long ago by D. Hume in his critique of induction: in physics we believe in induction only because it has worked over and over again.

That probability-1 assignments are as any other assignment is essential for the coherence of QB-ism.

References

[Bohr 1934] N. Bohr, *Atomic theory and the description of nature*, Cambridge University, 1934.

[Bohr 1935] N. Bohr, "Can quantum-mechanical description of physical reality be considered complete?", *Phys. Rev.* **48**:8 (1935), 696–702.

[Bohr 1963] N. Bohr, "The unity of human knowledge", pp. 8–16 in *Philosophical writings of Niels Bohr*, vol. 3: Essays 1958–1962 on atomic physics and human knowledge, Ox Bow, Woodbridge, CT, 1963.

[Dell'Antonio 2015] G. F. Dell'Antonio, "On tracks in a cloud chamber", *Found. Phys.* **45**:1 (2015), 11–21.

[Figari and Teta 2014] R. Figari and A. Teta, *Quantum dynamics of a particle in a tracking chamber*, Springer, Berlin, 2014.

[Fuchs and Peres 2000] C. A. Fuchs and A. Peres, "Quantum theory needs no 'interpretation'", *Phys. Today* **53**:3 (2000), 70–71.

[Fuchs et al. 2015] C. A. Fuchs, N. D. Mermin, and R. Schack, "Reply to Nauenberg", *Amer. J. Phys.* **83**:3 (2015), 198.

[Griffiths 2002] R. B. Griffiths, *Consistent quantum theory*, Cambridge University, 2002.

[Hornberger 2009] K. Hornberger, "Introduction to decoherence theory", Chapter 5, pp. 221–276 in *Entanglement and decoherence*, edited by A. Buchleitner et al., Lecture Notes in Physics **768**, Springer, 2009.

[Jaynes 1990] E. T. Jaynes, "Probability in quantum theory", pp. 381–403 in *Complexity, entropy and the physics of information* (Santa Fe, NM, 1989), edited by W. H. Zurek, Santa Fe Institute Studies in the Sciences of Complexity **8**, Addison-Wesley, Redwood City, CA, 1990.

[Kostecki 2014] R. P. Kostecki, "Lüders' and quantum Jeffrey's rules as entropic projections", preprint, 2014. arXiv 1408.3502v1

[Landau and Lifshitz 1965] L. D. Landau and E. M. Lifshitz, *Quantum mechanics: non-relativistic theory*, 2nd ed., Course of Theoretical Physics **3**, Pergamon, Oxford, 1965.

[Mott 1929] N. F. Mott, "The wave mechanics of α-ray tracks", *P. Roy. Soc. A* **126** (1929), 79–84.

[Pauli 1994] W. Pauli, *Writings on physics and philosophy*, edited by C. P. Enz and K. von Meyenn, Springer, Berlin, 1994.

[Petersen 1963] A. Petersen, "The philosophy of Niels Bohr", *B. Atom. Sci.* **19**:7 (1963), 8–14.

[Raimond 2014] J.-M. Raimond, "Exploring the quantum world with photons trapped in cavities and Rydberg atoms", Chapter 11, pp. 395–426 in *Quantum machines: measurement and control of engineered quantum systems* (Les Houches, France, 2011), edited by M. Devoret et al., Lecture Notes of the Les Houches Summer School **96**, Oxford University, 2014.

[Robert 1998] D. Robert, "Semi-classical approximation in quantum mechanics: a survey of old and recent mathematical results", *Helv. Phys. Acta* **71** (1998), 44–116.

[Schrödinger 2011] E. Schrödinger, *Eine Entdeckung von ganz außerordentlicher Tragweite: Schrödingers Briefwechsel zur Wellenmechanik und zum Katzenparadoxon*, edited by K. von Meyenn, Springer, Berlin, 2011.

[Vedral 2002] V. Vedral, "The role of relative entropy in quantum information theory", *Rev. Modern Phys.* **74**:1 (2002), 197–234.

[Wheeler and Zurek 1983] J. A. Wheeler and W. H. Zurek (editors), *Quantum theory and measurement*, Princeton University, 1983.

Received 1 Sep 2016. Revised 6 Dec 2016. Accepted 12 Jan 2017.

GIANFAUSTO DELL'ANTONIO: gianfa@sissa.it
Mathematics Department, University Sapienza, Rome, Italy

and

Mathematics Area, SISSA, Trieste, Italy

msp

MATHEMATICS AND MECHANICS OF COMPLEX SYSTEMS

msp.org/memocs

MEMOCS (ISSN 2325-3444 electronic, 2326-7186 printed) is a journal of the International Research Center for the Mathematics and Mechanics of Complex Systems at the Università dell'Aquila, Italy.

Cover image: "Tangle" by © John Horigan; produced using the *Context Free* program (contextfreeart.org).

PUBLISHED BY

mathematical sciences publishers

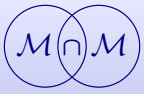
nonprofit scientific publishing

http://msp.org/

Mathematics and Mechanics of Complex Systems
vol. 4 no. 3-4 2016

Special issue in honor of
Lucio Russo

MEMOCS is a journal of the International Research Center for the Mathematics and Mechanics of Complex Systems at the Università dell'Aquila, Italy.



2326-7186(2016)4:3;1-A