REFLECTIVE METAPHYSICS:
UNDERSTANDING QUANTUM
MECHANICS FROM A KANTIAN
STANDPOINT

Michel Bitbol

ABSTRACT

Instead of either formulating new metaphysical images (as realists would do) or rejecting any metaphysical attempt (as empiricists would do), the case of quantum mechanics might well require from us a complete redefinition of the nature and task of metaphysics. The sought redefinition can be performed in the spirit of Kant, according to whom metaphysics is the discipline of the boundaries of human knowledge. This can be called a “reflective” conception of metaphysics. In this paper, each one of the most popular “interpretations” of quantum mechanics is shown to be naturally associated with a variety of Kant-like reflective metaphysics. Then, the two major “paradoxes” of quantum mechanics (the measurement problem and the EPR correlations) are reformulated by way of this reflective attitude, and they are thereby “dissolved”. Along with this perspective, quantum mechanics becomes one of the most elegant and understandable theories of the history of physics in addition of being one of the most efficient. The only point that must be clarified is why it looks culturally so difficult to accept a reflective and non-ontological standpoint on physical theories.
1 Introduction

A preliminary condition for any inquiry into quantum metaphysics is to be clear about what is meant by “metaphysics”. The key etymological component of this word is meta-. This prefix is usually translated as “after” by due reference to the order of Aristotle’s writings, but it is more conveniently understood as “beyond”. According to E. Levinas (2003, p. 22), “The metaphysical desire does not strive towards return, for it is desire of a land where we were not born; a land foreign to any nature, which was not our homeland and towards which we will never move”. The primary aim of metaphysics is then to go beyond, with an almost dreamlike flavor, yet with some faith in rational rigour. More precisely, this aim is to expand the scope of knowledge beyond its empirical area of validity (so as to reach the met-empirical land where “we were not born”). It is to get some insightful information about what exists beyond appearances, beyond law-like relations between phenomena, beyond the formal order of the theoretical predictions that bear on the outcome of well-defined classes of experimental and technical interventions. And at the same time, at least in the recent program of “experimental metaphysics” (Cohen et al. 1997), the aim is to look for some good empirical arguments to prefer this or that representation of the met-empirical realm. In the domain of quantum physics, this urge to “go beyond” is felt all the more strongly since:

(1) The only consensual part of the theory is a formal skeleton enabling one to calculate the probability of various experimental outcomes at any time, given the initial preparation (Peres, 1995; Schwinger, 2001).

(2) This formal skeleton is often complemented with bits and pieces of former pictures of the world borrowed from classical physics, but connected to one another in an unfamiliar and unruly way.

A recurring complaint is that, as long as we are left without any truly coherent representation of the world and of its “ontological furniture” compatible with the quantum formalism, we cannot claim that we truly
“understand” quantum mechanics. The well-known paradoxes (or enigmas) of quantum mechanics are taken as evidence of this felt lack of understanding. At the same time, none of the many representations of the quantum world that have been formulated until now has gained unanimous acceptance. Ontology is strongly underdetermined (essentially between particles, fields, or deeper dispositions); there is recurring debate about the meaning or even existence of non-locality; there is ongoing disagreement about whether and how one should take the measurement process as an object of description, etc. To sum up, what characterizes the current state of philosophical research about quantum mechanics is a combination of (i) urgent need for pictures of what is allegedly “beyond” the empirical domain, and (ii) persistent failure to gain general agreement about any such picture.

The situation is somehow reminiscent of the one Kant discovered in the aftermath of Newton’s theoretical achievement. Kant was struck by the fact that, in his time, logic, mathematics, and (Newtonian) physics, had followed “the secure path of a science”, as witnessed by the unique historical development they had undergone and by the widespread agreement of scientists about them. By contrast, “Countless times, in metaphysics, we have to retrace our steps”; and moreover metaphysics remains “a combat arena … in which not one fighter has ever been able to gain even the smallest territory and to base upon his victory a lasting possession” (Kant, 1996, p. 20). After having tried, during the so-called pre-critical period of his philosophy, to formulate one more metaphysical picture of the world by imposing it a strong condition of compatibility with Newton’s mechanics, Kant realized (under the influence of Hume’s Treatise) the weakness of any such attempt, and he then dramatically changed his strategy. Instead of trying to look through the physical theory towards the world supposedly described by it, he asked about the cognitive “conditions of possibility” of the remarkable success of this theory. Accordingly, instead of either formulating a new metaphysics (as a dogmatist would do) or rejecting any metaphysical attempt (as an
empiricist would do), he completely redefined the nature and task of
metaphysics in agreement with his choice to focus attention on cognition.
Kant recognized, in other terms, that the land sought by metaphysics can
be foreign to us not due to the excessive distance we have with it, but
rather due to its excessive proximity to us. In the latter case, however, the
prospects of obtaining reliable results is much better.

In this article, my aim is to characterize the nature and task of
quantum metaphysics in a Kantian spirit. Of course, this does not imply
imitating Kant’s solutions, which were adapted to the physics of his time,
but only to adopt his basic attitude, which is reflective. The roadmap is
then as follows. In section 1, I analyze Kant’s redefinition of metaphysics
in the spirit of his critical philosophy, so as to make a similar move for
quantum metaphysics. In section 2, I show how each one of the most
popular “interpretations” of quantum mechanics is associated with a
brand of Kant-like reflective metaphysics. Then, in section 3, a renewed
understanding of quantum mechanics (and a dissolution of two so-called
“paradoxes”) is gained by way of this reflective attitude.

2 Critique of metaphysics and renewal of
metaphysics.

Kant’s attitude to metaphysics was more balanced than usually accepted.
It combined:

- a sharp rejection of standard metaphysical pretensions of providing
  access to some suprasensible realm,
- a nuanced recognition of the function of metaphysical quest in human
  knowledge, and
- a new conception of what is and what should be metaphysics in a
  scientific age.
Kant’s rejection of metaphysical excess was partly directed against his own pre-critical beliefs. The last (and perhaps clearest) expression of these beliefs can be found in his *Inaugural dissertation* of 1770: “It is plain that what is sensuously thought is the representation of things as they appear, while the intellectual presentations are the representations of things as they are”. In other terms, sensibility gives access to *phenomena*, whereas reason opens the way towards things-in-themselves. But Kant was by no means naive about the powers of reason (or *powers of syllogistic inference*, that he distinguished from the powers of conceptualization that belong to our *understanding*). He knew that the proper use of inference is only to connect propositions bearing on contents of *experience*, or on generalizations of judgments of *experience*, and certainly not to push the inquiry entirely *in abstracto*. Kant then soon became utterly diffident about any possibility of breaking up the bounds of experience towards the properties of things “as they are in themselves” by way of reason. After all, he noticed, the mere use of the rules of general logic, the purely formal workings of concepts and principles, is unable to provide us with the slightest indication about objects, since they only point towards the operations of our thought (Kant, 2004a, Introduction). The proof that this cognitive use of reason is indeed unwarranted is provided by the so-called “antinomies”, or conflicts of reason with itself documented in the *Critique of Pure Reason*: two perfectly convincing reasonings can be provided in favor or against two archetypal metaphysical theses, without any way of deciding between them. Our preference for one thesis against other options is to be ascribed to our pre-rational “interests” (or values), rather than to rational evidence. Reason, if left alone at an excessive distance of the empirical material it is meant to order, looks unable to reach any univocal conclusion.

Kant however nuanced his condemnation of the cognitive value of metaphysics, by recognizing two things: (i) that metaphysics is an unavoidable by-product of the work of reason, a sort of extrapolated shadow cast by the otherwise operationally valuable faculty of inference;
and (ii) that metaphysics may serve as a guide and impulsion for sound scientific research. The extrapolation of the connecting power of conceptualization and inference is natural, in so far as, when they use this power, subjects seek both completion of the system of knowledge and ultimate explanations of observed phenomena. Reason has no spontaneous clue as to where it should stop its task of connecting propositions to one another, or embedding particular propositions into general principles, and it then pursues this task until it has formulated an all-encompassing representation, even when many elements of this representation look arbitrary. In particular, reason sees no limit to its urge for causal explanations, and it then pursues the search of causes until it has posited some sort of ultimate or primordial cause, even though no proof of the existence of this ultimate cause can be given. It is the assignment of critical philosophy to set these limits, and to circumscribe the domain of validity of rational inquiry. It is its specificity to understand that “nothing is actually given to us but perception and the empirical advance from it to other possible perceptions” (Kant, 1996, B521-522, op. cit., p. 508). Trying to bypass empirical advance and directly refer to a transcendent entity construed as the end of this advance is a typical move of reason, but, according to critical philosophy, this move is only tantamount to giving a totalizing name to the ever-developing “progression of experience” itself.

Kant nevertheless accepts that this inclination of reason towards extrapolating its power can be useful in many circumstances of scientific research. It is useful as a “regulation” of research, as a guide and incentive for scientists, even if it has no constitutive value for knowledge. In the introduction of his Prolegomena to any Future Metaphysics (a title that clearly indicates that metaphysics is by no means rejected as such, but only in its present state and definition) Kant thus criticized Hume’s exclusively negative judgment about metaphysical thought. According to him, Hume “overlooked the positive injury which results, if reason be deprived of its most important prospects, which can alone supply to the
will the highest aim for all its endeavor”. Speculative metaphysics is in fact useful to a certain extent because it provides research with images of possible prospects; and aiming at such prospects in turn strengthen the will in its quest for knowledge. This position is not very far from the “motivational realism” that has sometimes been ascribed to Einstein (Fine, 1986, p. 109) or Duhem (Darling, 2003).

Kant’s motto is that, despite its stemming from the “extravagant claims of speculative reason” (Kant, 1997, Introduction), metaphysics should not be rejected but disciplined. It should be given an epistemological rather than ontological status, so much so that ontology itself is seen as an epistemological tool. At the very end of Kant’s work of reconstruction, metaphysical statements are then no longer seen as representations of something “out there”, but as rules in a grammatical pre-ordering of experience. They are to be taken as formal tools of our intellect, and their value is to be assessed by due reference to the abilities of this intellect. Hence, metaphysics becomes nothing else than a reflective analysis of the powers and credence of reason. At first sight, this move may seem to be at variance with all the former philosophical tradition, but it has many forerunners in medieval philosophy. For instance, Duns Scot considered that one is not allowed to speak of transcendent entities unless the words and concepts that are used to this purpose are accepted as valid; he thus thought that metaphysics could not be developed independently of a critical analysis of the ability of knowledge to exceed sensible experience (Honnefelder, 2002, p. 25). If metaphysics is to become a science of primordial Being, it cannot avoid to wonder about the significance of primordial Cognition. The pretension of metaphysics to reach a transcendent realm is conditional to its capacity of gaining access to the transcendental domain of the necessary presuppositions of knowledge (Honnefelder, 2002, p. 3).

Kant went as far as possible in this direction, to conclude, even before the first Critique, that “metaphysics is the science of the boundaries of human reason” (Kant, 1899, p. 113). If metaphysics is to become as well-
established as a mathematical or physical science, he explained at length, this disciplin has to deal neither with things in themselves, nor with selected objects of experience, but with the *preconditions* for knowledge of any object of possible experience. Conversely, if one inquires into the causes of the failure of old-fashioned dogmatic metaphysics, it becomes clear that this is due to insufficient analysis of the scope and validity of the instruments of reason that it uses without control: “metaphysics has an hereditary failing, not to be explained, much less set aside, until we ascend to its birth-place, pure reason itself” (Kant, 1997, Appendix).

True, the dream of metaphysics is to strive, by way of reason, from the sensible to the supra-sensible realm (Kant, 2002, 351-424), but at the end of the day, this only means that “metaphysics does not bear on objects, but on knowledge” (Kant, 1902-1983, vol XVIII, Reflection 4853). The recent (post-leibnizian) specialization of metaphysics in systematic construction and evaluation of possible worlds, far from being a refutation of Kant’s redefinition, is an excellent illustration of the way it works: a world can be said “possible” if its components are “compossible” relative to certain rational rules. This kind of hypothetical byproduct of metaphysics thus displays nothing else than the very workings of reason.

We can therefore conclude that there is a legitimate domain of metaphysics. But this domain is neither speculative nor empirical; it is reflective. If lucidly analyzed, metaphysics reflects on the use and limits of speculation, and on the preconditions of empirical knowledge. By doing this job overtly, Kant-like metaphysics clarifies to a considerable extent the status of physical theories. For example, it shows that crucial parts of the structure of classical mechanics can directly be derived from the preconditions of objective knowledge of spatio-temporal objects (called material bodies) (Kant, 2004b). Far from having a purely empirical origin, a large fraction of this structure is provided in advance by the very project of objectifying something of what is given through our sensibility.
This reading of classical mechanics had the power to dispel many (now almost forgotten) qualms of thinkers of the end of the seventeenth century about: (i) the purely mathematical nature of Newtonian mechanics, (ii) its partial lack of intelligibility, and (iii) its unfamiliar features. Thus, according to Kant, the mathematical nature of Newton’s theory was justified neither by a narrow-minded interest for precision, nor by the neo-platonic belief that the book of nature was written in mathematical characters by its Creator. It was in fact a pre-requisite of any attempt to prescribe a law-like order bearing on possible sequences of events beyond any particular actual event (Bitbol et al. 2009); and this prescription, in turn, was a condition of objective knowledge. Similarly, the lack of intelligible elucidation of the ultimate causes of gravitation was not due to any incompleteness of that particular theory, but to the all-pervading fact that our intelligence is not fit for anything else than connecting phenomena to one another; no other explanation than systematic mathematical connections of contents of experience can then be provided by it. As for the unfamiliar (and ontologically surprising) features that pervaded classical mechanics, they could also easily be traced back to epistemological constraints. One such prominent feature is Galileo’s principle of relativity which asserts, against common sense and against Aristotle, that “(uniform) motion is like nothing”. But the principle of relativity is made indispensible by the fact that space being no object of experience by itself, absolute motion in space is not an object of experience either (Kant, 2004b, Chapter 1, Axiom). Indeed, since physics is not concerned by anything else than objects of possible experience, it can confer no ontological status to absolute motion.

My wish here is to apply reflective metaphysics in order to clarify and to dispel some alleged “paradoxes” of quantum mechanics. But in order to do so, the resources of transcendental philosophy must be exploited to their maximal extent, even over and above what Kant could figure out. There are many ways to do that. One way is to make use of strategies that (unlike the Critique of Pure Reason and the Metaphysical Foundations of
Kant developed for purposes which were not immediately connected to physics. This is especially the case of the *Critique of Judgment* that was initially meant to elucidate the status of aesthetic judgment and biological teleology. Another way is to broaden and relativize the definition of the *a priori* forms of the *Critique of Pure Reason* in the wake of neo-Kantianism. One can thus adopt a pragmatic definition of the *a priori* instead of a purely intellectual one (Pihlström, 2003). According to this definition, an *a priori* form is no longer a universally necessary intellectual condition for objective knowledge, but a pragmatic condition locally and provisionally necessary for the determination of some intersubjectively shared domain of experimental or technological intervention. Such pragmatic *a priori* forms closely correspond to Reichenbach’s general definition of a “constitutive principle” whose validity is relative to a certain domain of practice (Reichenbach, 1965).

3 A reflective interpretation of interpretations

Combining a reflective attitude and a relativized conception of the *a priori* is bound to renew our conception of the so-called “interpretations” of quantum mechanics. Along with the reflective conception of metaphysics, an interpretation of quantum mechanics is not to be construed as a tentative representation of the world, but rather as an expression of some favorite aspect of the procedure of constitution of objectivity. Here, constituting objectivity only means picking out a significant and sufficiently stable feature of phenomena, so as to perform efficient actions that yield predictable events. The fact that there are many such interpretations can be traced back to the many “symbolic forms” (Cassirer, 1953) that are available to circumscribe a domain of stable phenomena significant for a certain range of interventions. And the fact that below the level of these interpretations there is a unique formalism (say the Dirac-Von Neumann structure) demonstrates that it is possible to
extract a sort of invariant of invariants, an abstract symbolic form setting
the frame for a highly general type of efficient actions. I will now
evaluate some interpretations of quantum mechanics along this line:
showing which kind of “symbolic form” it imposes in advance, and
which significant aspect of phenomena it helps us to pick out. I will
successively document the information-theoretic interpretation(s), the
various brands of realism of the wave function, Bohr’s complementarity,
and Bohm’s hidden variables.

a) The information-theoretic interpretation of quantum mechanics
might well be the deepest one (Fuchs, 2009), because (in the same way as
Heisenberg’s original formulation in terms of matrices of observable
values) it immediately points towards the conditions of possibility of any
theoretical elaboration. In fact, it comes closer to an integral
reconstruction of the basics of quantum formalism from first principles
than to any (unavoidably partial) “interpretation” (Grinbaum, 2007).

Firstly, the necessity of a dichotomy of the domain of discourse into a
transcendental background and a set of objectified entities is immediately
apparent when information is at stake. Indeed, information and material
bodies are mutually dependent: material bodies are reconstructed
abstractions out of sensorial and experimental information, whereas
information requires some material body as data support. One must then
decide at some point which of the two concepts is taken as primitive,
namely as pre-conditional. This very act of choosing to put part of the
domain of discourse into bracket and take it as pre-conditional is typical
of transcendental epistemology.

Secondly, it is indeed possible to reconstruct a large part of the
structure of quantum mechanics (essentially the structure of orthomodular
lattice) from elementary axioms bearing on experimental information
(Grinbaum, 2005). The axioms that are used to that purpose are especially
instructive, since they state the limits of available experimental
information when conjugate variables are measured (Grinbaum, 2003).
But this is exactly the kind of result which is to be expected from
reflective metaphysics in the spirit of Kant: bringing out what, in the form of our knowledge, is due to nothing else than the limits of the (mental and technological) instruments of this knowledge.

b) One of the most widespread metaphysical position associated with quantum mechanics is realism of the wave function, or realism of the state vector. Even though it is rarely expressed as such because every physicist knows (or should know) that, in practice, a state vector is little more than a mathematical tool for computing probabilities of experimental events, this interpretation is latent in several research programs of quantum mechanics. Two research programs that (explicitly or implicitly) rely on realism of the state vector are Everett’s or many-worlds interpretations of quantum mechanics, and decoherence theories. Everett’s most extensive version or his conception was entitled *The Theory of the Universal Wave Function* (DeWitt and Graham, 1973). In this text, realism of the wave function (or state vector) was overtly endorsed and traced back to Schrödinger. This realist reading was taken as a good reason to reject the postulate of reduction, and to take the terms of linear superpositions of eigenstates of observables at face value (namely *qua* really existing in a corresponding number of branches or “worlds”).

As for decoherence theories, they are often interpreted as a method to display how classical appearances emerge from quantum reality. Indeed, decoherence shows (by way of “environment-induced superselection”), that the structure of quantum micro-states entails the macro-state structure observed in the laboratory. If one assumes that linear superpositions of vectors in a Hilbert space faithfully represent physical reality, then everything else, including the classical behavior of macroscopic objects, is nothing but a superficial appearance to be derived from it. Other readings of decoherence are available, including an empiricist and a transcendental reading (Bitbol, 2007), but the dominance of this standard reading bears witness of the latent realist construal of state vectors.
Now, this transcendent metaphysics of state vectors can be reinterpreted straightaway in terms of reflective metaphysics. An early remark in this direction was made by P. Mittelstaedt. According to Mittelstaedt’s thoroughly Kantian analysis, in quantum mechanics state vectors play the role of *substances* (namely of permanent bearers of a restricted set of objective properties) that was played by material bodies in classical physics (Mittelstaedt, 1976). His remark can be developed in three steps. (1) In quantum mechanics, a probabilistic predictor (the state vector) has much better characteristics of stability and law-like dependence than any cluster of events in space-time. (2) Stability and law-like dependence are exactly the two crucial features that define objectivity in Kant’s approach. (3) Therefore, in quantum mechanics, what is objectified is no longer a spatio-temporal entity, but a predictive symbol called the state vector.

Such observation was later worked out, and turned into a program of “transcendental deduction” of quantum mechanics (Bitbol, 1997; 1998). Two constraints, imposed on theoretical prediction, were taken as departure points: (i) in microphysics predictions bears on contextual phenomena; and (ii) in spite of this, we seek a unique predictor valid across experimental contexts. The second constraint is crucial, since it imposes invariance of the predictor; it is precisely tantamount to a clause of constitution of objectivity. At this point, one can show that the basic Hilbert-space structure of quantum mechanics is the simplest formalism that respects the two former constraints.

A further achievement is to derive the Schrödinger equation from epistemological constraints. The preliminary step consists in relying on a well-known theorem according to which the general form of Schrödinger’s equation arises if one assumes that evolution operators are *unitary* and that they have the structure of a one-parameter group of *linear operators*. Now, these conditions are easily spelled out in terms of epistemological constraints: (i) unitarity is required because it ensures that probabilities add up to 1 at all time; (ii) the one-parameter group
structures ensures that successive state vectors are connected by a univocal rule (thus fulfilling the role of the category of causality). Taken together, the conditions impose the stability and lawlikeness of the probabilistic predictor; they are, once again, conditions for its objectivity. This being granted, the formalism of quantum mechanics appears as an expression of a set of conditions that establish the inter-contextual and inter-subjective invariance of probabilistic prediction. If supplemented (through the correspondence principle) with the older clauses of constitution of objectivity which underly classical mechanics, it becomes a full-blown physical theory. We conclude that reflective metaphysics is able to elucidate the meaning of the quantum formalism and to make good epistemological sense of certain speculative attempts of standard metaphysics (such as realism of wave functions).

c) Another metaphysical thesis about quantum mechanics is commonly ascribed to Bohr (although it is almost certain that Bohr never held it). It is the so-called “macro-realist” construal of properties and objects at our scale. According to it, there is a fundamental, nay ontological, difference between macrophysics and microphysics. Macrophysics describes intrinsically existing objects and properties, mutually exclusive to each other, whereas microphysics can only consist in an abstract “symbolism” used to predict experimental phenomena whose instrumental conditions are stated by way of macrophysical concepts.

Actually, Bohr did not ascribe any ontological significance to the use of macrophysical concepts; he only thought experimenters must use them if they want to communicate unambiguously about their preparations and their results. In other terms, he took macrophysical concepts as a condition of possibility of intersubjectively acceptable knowledge, thus ascribing them a transcendental significance, in Kant’s sense. But the domain of validity of this latter condition of possibility is intentionally restricted. It does not extend to the whole field of microphysics, but only to the direct environment of mankind, at the scale of the instruments used
by human scientists. It is an anthropocentric condition of possibility for the experimental conditions of possibility of microphysical research. In other terms, it is a second-order condition of possibility of microphysical knowledge. This idea was clearly formulated by Heisenberg: "What Kant had not foreseen was that these a priori concepts can be the conditions for science and at the same time have a limited range of applicability" (Heisenberg, 1990, p. 78). The macroscopic concepts that directly precondition knowledge of our direct environment and of the domain of classical physics have limited range of applicability, but at the same time they indirectly precondition knowledge of any research beyond this domain. Here again, we see how reflective metaphysics can clarify a situation (such as the alleged ontological difference between macro- and micro-objects) that is blurred by ordinary speculative metaphysics.

Yet, the clarification is incomplete if we have no reflective understanding of the domain of research that extends beyond the circle where macroconcepts (and especially the categories of the Critique of Pure Reason) apply. But nothing prevents us from making this further step. We must realize that Kant’s philosophy has also enough resources to formulate constructive propositions on what does not directly fall under the joint rule of the forms of intuition and categories of pure understanding (Palmquist, 1990). One major resource consists in construing the structure of quantum mechanics, not as an expression of the constitutive function of categories, typical of the Critique of Pure Reason, but as a formal projection of a program of unity of the system of nature, typical of the Critique of Judgment (Pringe, 2007; 2009).

To see how this can be done, let us first notice that, after the widespread puzzlement of physicists about the so-called “wave-corpuscle duality”, a new unexpected kind of unity was restored by Bohr’s concept of “complementarity”. According to it, these two exclusive representations (i) are relative to different types of experimental devices or different types of associated classical concepts, and (ii) jointly characterize “one and the same object”. However, the hypothetical object
towards which the two complementary representations are supposed to converge cannot be said to simultaneously possess the two corresponding properties. No constituted object is “behind” the contextual phenomena. Bohr’s “object” is only a regulative device used as a unifying symbol. In the same way as in Kant’s Critique of Judgment, what is used in this case is a purely «symbolic analogy», instead of the standard constitutive «analogy of experience» which would only be available for proper objects of intuition (Kant, 1987, p. 227).

My only objection to this analysis based on Kant’s third Critique, is that (unlike what Pringe claims in his book) it is by no means the only acceptable transcendental reading of quantum mechanics. It does not exclude other readings of this kind, as documented in points a) and b), but only complements them by its close adaptation to one particular interpretation of quantum mechanics, namely Bohr’s. This analysis which makes use of an extrapolation of the program of unity of the system of nature beyond the domain of validity of the categories of pure understanding, is just one more partial reflective metaphysical view of quantum mechanics. It only clarifies the pre-conditions of a certain procedure of constitution of objectivity used when ordinary methods are no longer available: the procedure that consists in focusing two or more experimentally incompatible phenomena on a single symbolic object construed as their common ground. But, as we have seen in point b), this option can be replaced with another procedure that consists in objectifying the probabilistic predictor of phenomena instead of their symbolically construed ground.

d) Bohm’s original theory of 1952 is likely to be the most metaphysical (in the strongest, speculative, sense) of all readings of quantum mechanics. It posits free particle trajectories in space-time, that are unobservable in virtue of the theory itself. Indeed, the theory is thoroughly contextualist. It involves holistic processes in which the mere presence of a macroscopic apparatus partakes of the definition of the alleged trajectories. Therefore, a truly “independent” trajectory is in
principle out of reach of experiment; only isolated spatial and
kine\ntical coordinates, codetermined by the “quantum potential” of the
macroscopic apparatus, can be observed. As Bohm himself acknowledges,
“This is reminiscent of Bohr’s notion of wholeness, but it differs in that
the entire process is open to our ‘conceptual gaze’ and can therefore be
analyzed in thought, even if it cannot be divided in actuality without
changing its nature” (Bohm and Hiley, 1993). The conceptual gaze on
possible or imagined processes is thus intrinsically cut from any actual
test.

If there can be no justification to such representation, there must at
least be an exceptionally strong motivation in its favor. It is at this point
that reflective metaphysics comes in once again to clarify the origin of
transcendent metaphysical speculation. The motivation of Bohm-like
speculation about unobservable trajectories is a reaction against Bohr’s
and Heisenberg’s renunciation to “pictures” of microphysical processes
in space-time. It is an attempt to reactiv\nte the immemorial dialectic of
the theory of knowledge as documented by Kant: (spatio-temporal)
intuition and concepts; concepts applied to intuitive (spatio-temporal)
contents; concepts geometrically constructed in “pure” intuition (namely
in an abstract spatio-temporal framework awaiting to be fulfilled by the
senses). This urge towards reactivation of the function of pure intuition is
so strong that it can accommodate very unusual features. For instance, it
has been shown that interpreting neutron interferometry in terms of
Bohm’s theory requires to accept that particles are “bare particulars”,
endowed with a position but with no other property localized in the
vicinity of this position (Brown, Dewdney and Horton, 1995). Recovery
of the ability of our understanding to operate on pure (spatio-temporal)
intuition is a project which, for many physicists, seems to justify almost
any conceptual boldness. If this provides physicists with a heuristic tool
in certain experimental situations such as the calculation of complex
interferometric patterns, so much the better; but if limitation of thought
to an old pattern becomes a hindrance, one should be ready to dispense
with it. This readiness is favored by reflective metaphysics, and considerably hampered by speculative metaphysics (for who would accept to relinquish an image that is believed to tend to a faithful representation of the world as it is?)

4 Do “paradoxes” dissolve under the reflective gaze?

A major strength of scientific endeavor is to ensure a certain consensus about the nature and formulation of the problems to be solved. On the basis of this consensus, every effort can be devoted to specialized solutions, with no interest for residual doubts concerning the general meaning of research. Such a consensus, such silencing of radical questioning, and such concentration of work on accepted issues, is what defines a paradigm in Kuhn’s sense. The very functioning of scientific institutions, such as journals, doctoral studies, and recruitment in research organizations, depends on instituting a paradigm. Now, of course, before scientific communities can reach this state, a large debate must be accepted in order to define the boundaries and the reliability of the consensus. But the aim of this debate is to reach a (hopefully quick) conclusion. Philosophy is invited in the debate, but only with a short-term mission.

A recent phenomenon, promoted by analytic philosophy and technical philosophy of science, is the attempt at establishing paradigms even in philosophy. The advantage, here again, is a common view of problems, the possibility of specialized research, and the belief that cumulative progress is accomplished by piling up specialized contributions making intense reference to one another. This, however, has a drawback: the necessity of relying on another disciplin whenever one feels the need to question the basis of that consensus. Let us call this other disciplin “boundary-philosophy”, because its task is to question the boundaries of thought, rather than to develop thoughts within accepted boundaries.
Besides, even in physics, closing too early the initial phase of foundational discussions may be damaging. For, if the foundational issues have not been completely clarified from the outset, and if a rigid set of problems has been accepted without enough discussion about their relevance, one’s lasting inability to solve some of these problems can be perceived as a symptom of permanent “crisis”. It can even trigger diffidence against a theory despite its remarkable practical success. Yet this so-called “crisis” of the theory may well boil down to a crisis of the too quickly accepted formulation of its problems, or to a crisis of the prejudice associated to it in the name of immediate efficiency. In this case, philosophers, who were too hastily banished from the scene of science, are invited again. But, according to the type of philosophers, the results to be expected are not the same. Ordinary philosophers will accept the formulation of problems, even when they trace them back to more and more elementary assumptions. By contrast, boundary-philosophers will question even this formulation and may well come to the conclusion that the consensual problems are ill-posed or do not exist. Specialists of analytic philosophy and analytic metaphysics belong to the first type; but continental philosophers trained to Kant’s “copernican revolution”, and interested in reflective metaphysics, usually belong to the second type.

Two traditional “paradoxes” of quantum mechanics (the measurement problem and the EPR paradox) will now be examined in the spirit of boundary-philosophy and reflective metaphysics: wondering whether the problems arise, and whether the paradoxes exist at all.

a) Should we try to “solve” the measurement problem? Bohr was almost incredulous when he read or heard the formulations of the problem which were offered by his colleagues. He could see no such problem in quantum mechanics. According to him, one has no reason to wonder how linear superpositions of eigenstates can fit with the univocity of experimental outcomes, since (i) these superpositions partake of a mere “symbolism” that has no ambition to “represent” real events in space time but only to evaluate their probabilities, and (ii) the existence
of well-defined univocal outcomes is a necessary presupposition of any experimental research, that has no need of being “derived” from the formalism. This blunt denial has rarely been accepted, even by Bohr’s closest collaborators as Rosenfeld (1965). But is such lack of acceptance of Bohr’s straightforward “dissolution” of the measurement problem really justified? I rather suspect it is due to an irresistible Platonician tendency to reify formal elements (such as the state vector), and also to misunderstandings about the remarkable structure of quantum probabilistic valuations.

The tendency to reify state vectors manifests itself in the use of the very word “state”. The “grammar” (in Wittgenstein’s sense) of the word “state” requires that this is the state of something; that it belongs to something; that it characterizes this something independently of anything else. Such grammar, and the conception associated to it, is sufficient to generate one of the major aspects of the measurement problem. To see this, let’s compare two sentences bearing on Schrödinger’s thought experiment:

(1) “After Schrödinger’s preparation, the measurement chain, including the cat, is in the state: \( |\Psi\rangle = 2^{-1/2} (|1\rangle |\text{dead}\rangle + |0\rangle |\text{alive}\rangle)\)”
(2) “It is found experimentally that the state of the cat is ‘alive’”

What generates a contradiction between these two sentences is the common word “state”, and also, in the background, the verb “to be” that is suggestive of ontology. The state of the measurement chain (and of the cat which partakes of it) cannot be superposed and well-defined at the same time. Hence the usual question about what triggers a transition of the state of the measurement chain (especially the cat) from (1) to (2). Hence also the usual puzzlement when it is seen that, irrespective of the number of systems that are allowed to interact with the initial microscopic system (here a radioactive atom), no such transition occurs unless it is imposed by hand. The feeling that there is a contradiction is attenuated when (as suggested by Van Fraassen), one makes a distinction between the dynamic state which is referred to in (1), and the value state
which is referred to in (2). This feeling even completely vanishes if, in agreement with Bohr, one considers: (i) that $|\Psi>$ is no “state” at all (especially not a state of the overall system) but only a mathematical “symbol” enabling one to calculate the probability of an outcome arising from a well-defined experimental environment, and (ii) that “alive” is no statement of what the cat is, but rather of what is found about it in observation. In the latter case, no sudden transition of the state vector after the experimental process is needed, but only a redefinition of it for a practical purpose of easy probabilistic valuation bearing on further experiments that may be performed on the system. As Schrödinger cogently noticed in 1935, when he commented the Copenhagen views, $|\Psi>$ does not change during a measurement; instead, “it is born anew, is reconstituted, is separated out from the entangled knowledge that one has” (Schrödinger, 1983). But how credible is this old-fashioned Bohrian view? How acceptable is the claim that the so-called state vector does not represent the state of a system, and that it rather expresses a probabilist assessment globally applicable to a complete experimental set up? Doesn’t it trivialize, or ignore, the remarkable work of the generations of physicists who formulated elaborate formalisms (from consistent histories to decoherence) in order to solve the measurement problem?

To begin with, every physicist should know that the state vector does not represent any inherent feature of the system to which it is ascribed. A thought experiment, designed by C.F. von Weizsäcker (von Weizsäcker 1974 : Discussion in : M. Jammer 1974) in 1931, and later reinvented by J.A. Wheeler (1983) under the name “delayed choice experiment”, clearly demonstrates this. Let’s assume that a photon interacts with an electron. After the interaction, an entangled state vector is tentatively associated with the system (electron + photon). If we wish to ascribe selectively a state vector to this electron, a good strategy is to evaluate the position and/or the momentum of the photon by using a microscope. Indeed, this measurement yields disentanglement of the former state
vector. But, according to whether one prepares the microscope in a configuration aimed at measuring the photon’s momentum with maximum precision, or in a distinct configuration aimed at measuring the photon’s position with maximum precision, one must ascribe two completely different state vectors “to the electron”. The state vector “of the electron” at a certain time depends on the choice of a measurement to be performed on the photon arbitrarily later. Unless one is ready (as Wheeler was) to assert that preparing the microscope has retrospectively, suddenly, and inscrutably altered the “state of the electron”, this is enough to conclude that the state vector should by no means be ascribed to it taken in isolation. The state vector rather characterizes the overall experimental preparation, including the microscope’s configuration. It expresses the appropriate probabilistic valuations of experiments performed on any electron having undergone this preparation. In other terms, it reflects future propensions determined by the global preparative situation which includes the electron, the photon, and the microscope, but certainly not the “state” of the electron taken apart. No wonder that each time the preparation is redefined, the “state” vector relevant for predicting further measurements on the electron has also to be redefined.

The second point concerns probabilities. Authors of textbooks often discard the claim that state vectors merely represent instruments of probability valuations, by noticing that these “probabilities” do not add up but rather interfere in a wave-like manner; and that no “ignorance interpretation” of the probability valuation is therefore allowed. The only sound remark is the second one. Quantum probabilities are not based on an underlying ontology of monadic properties (which one just happens to ignore), but on a background of relational observables whose act of measurement has to be implemented to make them exist. Predicting the value of a relational observable must take into account the fact that such value is necessarily conditional on the effective intervention of an experimental contraption, and has no existence “an sich” as long as no such contraption is operated. This relational status of observables and
quantum probabilities has some consequences that account for many *prima facie* surprising features of quantum predictions. The major surprising feature is that “amplitudes add up, but not probabilities”. It can be traced back to the fact that quantum probability functions are not defined on a single boolean algebra of propositions ascribing properties to objects, but on an orthoalgebra that contains several boolean algebras as substructures. Each one of these boolean algebras connects experimental propositions (rather than ontological propositions) whose validity is relative to one particular instrumental context. The resulting probability function is called a “generalized probability function” (Hughes, p. 222-223). It does not afford a probability valuation of each event in a single space of spontaneous occurrences, but as many scales of probability valuations as instrumental contexts relative to which measurements can be defined. Moreover, the many scales overlap as long as no one of them has been selected by the effective operation of an experimental device; the additivity of amplitudes precisely expresses this overlapping.

This being granted, a solution (or rather *dissolution*) of the measurement problem boils down to finding a way to articulate the indefinite chain of relational statements of the quantum theory to the absolute statements that are used in experimental work. An articulation of this kind can easily be found, provided one realizes that the latter absolute statements are in fact indexical; provided one realizes that these statements are only “absolute” relative to us, to our scale, to the open community of experimenters to which we belong (Rovelli, 1996; Bitbol, 2008). At this point, one is bound to realize the ineliminability of *situatedness* from the apparently neutral descriptions of quantum mechanics, and to accomplish thereby the reflective move typical of Kant’s renewed definition of metaphysics.

Finally, we must raise the issue of the vast amount of theoretical work that has been done in the past eighty years in view of solving the measurement problem. Is the remarkable value of this work annihilated
by the realization that the measurement problem does not really have to be solved? By no means.

(i) It is clear that if this work had not been done, and if, in spite of many accomplishments, it had not brought out how untractable the measurement problem is, the old Bohrian “dissolution” would not have seemed so unavoidable. The interest of coming back to this approach would not have been so strong if so many options had not been explored with mitigated success: “What is the solution to the measurement problem? I say it is this: on measurement of X with eigenstates $|\psi_i \rangle$, outcome $x_i$ is observed with probability $|\langle \psi | x_i \rangle|^2$, where $|\psi\rangle$ is the initial state. This is what we return to, so it will do for a beginning as well” (Saunders, 1994).

(ii) As Kant had predicted, in this case as in many others, a metaphysical representation is used as a guide and incentive of research; but the clarification of the results of this research can only arise from a study in reflective metaphysics. A good example is given by decoherence theories. As we have seen previously, they were formulated with the speculative hope that they would solve the measurement problem and show how classical appearances emerge from a quantum reality. They have proved to be extremely useful in many areas of applied physics (especially quantum computation), and they have been corroborated experimentally. Yet, many physicists now accept that decoherence provides us with no complete solution of the measurement problem, is so far as the ultimate transition from many potentialities to one actuality is out of its scope (Lyre, 1999). The only thing which is still wanting is then a proper interpretation of the true significance of decoherence theories, apart from their dubious realist motivation. This interpretation can be afforded by a reflective analysis: what decoherence provides is a proof of self-consistence of the system made of the
quantum probabilistic formalism and the classical-like interpretation of the instruments’ working and readings. Indeed, decoherence theories demonstrate that a classical structure of probability valuations which is liable to the ‘ignorance’ interpretation can emerge, approximately and under certain assumptions, from a quantum system of probability valuations which is definitely averse to any such ‘ignorance’ interpretation (Bitbol, 2009; Osnaghi, 2009). Here, in good agreement with the spirit of Kant-like reflective metaphysics, decoherence is taken not as a description of some piece of “reality out there”, but as an exposure of the interplay of our theoretical and experimental practices.

b) According to a popular belief, “quantum mechanics is non-local”. In fact, this belief derives not from quantum mechanics itself, but from Bell’s theorem about hidden variable theories able to predict the same phenomena as quantum mechanics. This celebrated theorem proves that no local hidden variable theory is compatible with quantum mechanics. Therefore, one must renounce either locality or property realism (which underpins the standard hidden variable program). A later theorem, also formulated by Bell, is even more stringent, since it imposes to renounce either locality or event realism (Bell, 1981). Now, recent work provides fresh arguments against the idea that non-locality is a consequence of quantum mechanics (Mermin, 1998; Fuchs and Peres, 2000; Deutsch and Hayden, 2000), and puts more and more pressure on the assumption of (property or event) realism even in its non-local variety (Gröblacher et al., 2007; Suarez, 2008). So, it becomes likely that one can make good sense of quantum distant correlations in a local framework of thought, provided property or event realism are completely relinquished. As we shall now see, this only requires to be clear about the cognitive context of the assertion that “events have occurred”.

As I insisted many years ago in my first paper about the foundations of quantum mechanics (Bitbol, 1983), two space-like separated events
can only be detected in a space-time framework located at the intersection of the future light-cones of these event. Comparing them is thus only possible in this latter space-time framework, and ascribing them a correlation in the space-like separated region where they occurred can only be done retrospectively, from the standpoint of this space-time framework. In classical thought, this would have had no importance, since nothing here prevents one from accepting that events occur “by themselves”. But in quantum physics, no event should be ascribed autonomy. In this case, every event is tantamount to an observable value- ascription, and an observable is only defined relative to an effective instrumental possibility of assessing it. In quantum physics, the instrumental context is not only a way of getting access to an event; it is a way of generating it. True, the two space-like separated events of a Bell-like experiment can be said to occur relative to a local apparatus, but the event of their comparison occurs relative to a third apparatus which necessarily operates somewhere at the intersection of the future light-cones of these event. Asking for a non-local explanation of the EPR correlation of these events is thus tantamount to forgetting that such correlation in fact only makes local sense, in the future space-time framework where the event of comparison between them can be generated. The same kind of idea has recently been developed on the background of Rovelli’s relational quantum mechanics (Smerlak and Rovelli, 2007). In relational quantum mechanics, saying that a quantum state vector is superposed or reduced is entirely relative to an observer’s framework. Then, one cannot say that the state vectors of the subsystems on which space-like separated measurements were performed, are reduced in the absolute as soon as these measurements are completed. Each one of their reductions only makes sense relative to a local observer, and reduction of their joint state vector (which contains the information about their correlation) only makes sense relative to an observer who can compare them, namely an observer situated somewhere in the intersection of their future light-cones. Quantum mechanics here appears as local and
complete, provided one does not try to figure out a domain of absolute properties, events, and states, that is not relevant for it. Once again, problems arise from speculative metaphysics, and they are easily dissipated in terms of reflective metaphysics.

5 Conclusion

As Kant wrote to his friend Markus Herz in May 11, 1781: “(The Critique of Pure Reason) contains the metaphysics of metaphysics” (Kant, 1902-1983, vol XI, p. 269). Critical philosophy goes beyond metaphysics by inquiring beneath its preconditions. By doing so, it annihilates standard metaphysical claims of knowledge, and at the same time it elucidates the origins of many paradoxes of rational inquiry that remained unfathomable as long as they were hidden in the smoke of conceptual reifications. This work of clarification is useful in every sector of scientific research, where it has recently taken the form of a “philosophy of experiment” inspired by pragmatism (Pickering, 1995; Hacking, 1983). But it is truly indispensable in quantum mechanics, where the preconditions of knowledge are also preconditions for the emergence of the empirical material of this knowledge. As soon as this elementary and massive fact is taken into account, the apparent paradoxes disappear, and quantum mechanics becomes one of the most elegant and understandable theories of the history of physics in addition of being one of the most efficient. The only point that makes such realization difficult to obtain is that the type of understanding gained in the process is by no means ontological, but reflective. A completely new conception of the aims and nature of physics would be needed in order to fully accept this kind clarification which implies, as Bohr foresaw, a transformation of the very sense of “understanding”. However, a redefinition of physics deep enough to reach this aim amounts to a change of culture, not only a change of methods. Many signs, in epistemology and in our civilization...
as a whole, indicate that we might be on the verge of a change of this size.

CREA, CNRS/ Ecole Polytechnique
ENSTA/CREA 32, Bd. Victor, 75015 Paris
Email: michel.bitbol@polytechnique.edu

REFERENCES


Kant, I., 1987, Critique of Judgment, Hackett, Indianapolis.


Classical mechanics in a nutshell. The classical state. 2. Historical Origins of Quantum Mechanics: Black-body radiation, the photoelectric effect, the Compton effect. Why an electron is not a wave. 4. The Quantum State: How does the electron get from A to B? Representation of the De Broglie wave as a point moving on the surface of a unit sphere. Reflective Metaphysics: Understanding Quantum Mechanics from a Kantian Standpoint. Michel Bitbol. Philosophica 83:53-83 (2010). A Philosopher's Understanding of Quantum Mechanics: Possibilities and Impossibilities of a Modal Interpretation. Pieter E. Vermaas - 1999 - Cambridge University Press. Towards a Micro Realistic Version of Quantum Mechanics, Part II. Nicholas Maxwell - 1976 - Foundations of Physics 6 (6):661-676. Bohmian Trajectories as the Foundation of Quantum Mechanics. Attempts to understand quantum mechanics as descriptive and explanatory, to modify it such that it can be so understood, or to argue that no such understanding is necessary, can all be taken as versions of the project of interpreting quantum mechanics. The problematic nature of quantum mechanics stems from the fact that the theory often represents the state of a system using a sum of several terms, where each term apparently represents a distinct physical state of the system. What’s more, these terms interact with each other, and this interaction is crucial to the theory’s predictions. According to the Copenhagen interpretation of quantum mechanics, the solution to this puzzle is that the quantum state should not be taken as a description of the physical system. Download Citation | REFLECTIVE METAPHYSICS: UNDERSTANDING QUANTUM MECHANICS FROM A KANTIAN STANDPOINT | Instead of either formulating new metaphysical images (as realists would do) or rejecting any metaphysical attempt (as empiricists would do), the | Find, read and cite all the research you need on ResearchGate. The only point that must be clarified is why it looks culturally so difficult to accept a reflective and non-ontological standpoint on physical theories. Discover the world's research. 19+ million members.