Classical and Non-Classical Representations in Physics II: Quantum Mechanics

Francis HEYLIGHEN*

PESP, Free University of Brussels, Pleinlaan 2, B-1050 Brussels, Belgium E-mail: Z09302 @ BBRBFU01.BITNET

ABSTRACT. The conceptual and formal structure of quantum mechanics is analysed from the point of view of the dynamics of distinctions, occcuring during the observation process. The Hilbert space formalism is simplified with the help of the concept of closure: closure of an eigenstate under an operator is generalized to the linear closure of a subset of states, and this is further simplified to orthogonal closure, meaning that a set of states can be distinguished by a single observation. Quantum states can be seen as (overlapping) subsets of unobservable infra-states, with the transition probability between two states proportional to the number of infra-states they have in common. This makes it possible to reconstruct the superposition principle. An analysis of the observation process leads to the interpretation of closed sets of infra-states as attractors of the dynamics induced by the interaction with the observation apparatus. This interaction is always partially indeterminate, because of the unobservable microstate of the apparatus.

KEYWORDS: quantum mechanics, distinction dynamics, closure, second order cybernetics, observation process.

Introduction

In a previous paper (Heylighen, 1990c), classical and non-classical theories in physics were in general analysed as distinction systems, to be modelled by a "dynamics of distinctions". Classical representations were characterized as systems where all distinctions are invariant, whereas non-classical representations were characterized by the fact that certain types (depending on the type of theory) of distinctions are not conserved. The dynamics of distinctions can in general be understood by the principle of "variation through recombination and selective retention of closed combinations". Closure is a mathematically defined concept which characterizes invariant distinctions.

This conceptual framework will now be applied to the epistemological and logical analysis of quantum mechanics, which in physics has always been the prototype of a non-classical theory. Moreover, quantum mechanics has inspired a lot of researchers in systems science as a model of non-classicality, especially with respect to the modelling of cognitive processes, and the concept of complementarity (see e.g. Kornwachs, 1988; von Lucadou, 1990; Löfgren, 1990). Quantum mechanics is also the only physical theory in which there is a formal representation of the

^{*} Senior Research Assistant NFWO (Belgian National Fund for Scientific Research)

observation process (though it is not very detailed). As such, the study of quantum mechanics is particularly relevant to the "second order" cybernetics, or the cybernetics of observing systems (von Foerster, 1981).

By studying the formalism of quantum mechanics, the specific ways in which distinctions change will be examined, with an emphasis on the non-trivial role of the closure concept. An interpretation of the formal structure will be given by looking at the observation process. But first we will review the historical origin of fundamental quantum mechanical concepts.

Conceptual structure of quantum mechanics

Origin of the quantum theory

The name of "quantum mechanics" or "quantum theory" comes from the assumed existence of a "quantum of action", i.e. a discrete unit of energy exchanged during microscopic interactions (interactions at the level of atoms or elementary particles). This assumption arose from the observation by Planck that—in order to derive the empirical laws of black body radiation—one had to assume that electromagnetic radiation is exchanged in discrete units, apparently in contradiction with the classical theory which considered electromagnetism to be carried by continuous fields or waves. This concept was elaborated by Einstein, who introduced the concept of a photon as a discrete, particle-like unit of electromagnetic radiation.

Shortly thereafter, an opposite phenomenon was observed: electrons, which were considered to be particles, appeared to undergo interference, which is typical of waves. This double phenomenon where waves behaved as particles and particles behaved as waves was called the *wave-particle duality*. The statement that something is at the same time a particle, hence discrete or discontinuous, and a wave, hence continuous, is a logical contradiction. Therefore we are confronted with a paradox which demands a radically new way of looking at things.

The prevailing attitude towards this problem (known as the "Copenhagen interpretation") was formulated most prominently by Bohr. According to Bohr, we cannot know physical reality as it is, independently of ourselves. We can only make certain representations of our interactions with physical systems. These representations are necessarily formulated in the language of classical physics. In the case of microscopic phenomena, however, there is no complete and consistent classical representation: there are only partial representations, such as the wave representation and the particle representation, which are complementary. Complementarity means that the representations are mutually exclusive, yet they are jointly necessary (they complement each other) for an exhaustive description of the physical phenomenon.

Bohr's reasoning in order to show the necessity of such complementarity is based on an analysis of the observation process. By thought experiments, he has shown that the observation set-up needed for measuring the position of the system (which is well-defined for particles) is incompatible with the set-up needed for measuring the momentum (which is well-defined for waves). Hence it is impossible to measure both properties at once, and since the measurement of the one perturbs the state of the system in an unpredictable way, it is impossible to measure the one after the other either. This impossibility of simultaneously determining two incompatible properties was expressed mathematically by Heisenberg, in the form of the "indeterminacy relations" (also called "uncertainty principle"), showing how a high precision in the determination of one property x (e.g. position) implies a low precision in the determination of a complementary or incompatible property p (e.g. momentum):

х.р ћ

x is here the uncertainty in the determination of *x*, \overline{h} is Planck's constant. The product *x*. *p* determines a volume in phase space. Hence the principle can be interpreted as saying that no measurement, however precise, can distinguish a part of phase space smaller than the volume \overline{h} .

The axioms of quantum mechanics

The initial mathematical formalisms for quantum mechanics (Schrödinger's wave mechanics and Heisenberg's matrix mechanics) were synthesized by von Neumann, who formulated quantum mechanics as an operator calculus in Hilbert space. We will summarize the axioms of this formal system, which is still accepted to be the basis of quantum theory (Jammer, 1974):

Axiom i: Each physical system can be represented by a Hilbertspace *S*. The state of the system is completely described by a vector of this Hilbertspace.

This axiom entails the so-called "superposition principle" which says that for every two different states of the system there exists a third state corresponding to the "superposition" (vector sum) of those states. There is no equivalent for this property in classical representations.

Axiom ii: Each observable (measurable quantity) can be represented in a unique way as a hermitian (self-adjoint) operator A in the Hilbertspace S. This means that A is a linear operator with real eigenvalues, corresponding to the possible measurement results. This axiom takes into account that a measurement must lead to a real measurement value.

In order to simplify the discussion we will assume the spectral theorem, which states that every self-adjoint linear operator can be reduced to a linear combination of projection operators. A projection operator projects a vector onto a subspace of the Hilbert space, and has only two possible eigenvalues: 1 (the vector already belongs to the subspace, and is hence projected onto itself) and 0 (the vector is orthogonal to the subspace and hence is projected onto the null vector). A projection operator corresponds to the observation of a property with two possible values : 1 (yes, the property is true for the system) or 0 (no, the property is not true for the system). This corresponds to a binary, Boolean variable, or proposition.

Axiom iii: For each system in a state $| \rangle$ the real number $| \langle | \rangle |^2$, where $\langle | \rangle$ denotes the scalar product of the vectors $| \rangle$ and $| \rangle$, gives the probability to obtain the measurement value a when the observable A is measured, where $| \rangle$ is the eigenvector of A corresponding to the eigenvalue a. This means that the result of an observation is in general unpredictable before the observation is carried out, even when the state is known, unless that state is an eigenstate of the observable, in which case the product is equal to 1.

Axiom iv: The time development of the state vector $|(t)\rangle$ is given by the Schrödinger equation:

$$H \mid (t) > = \hbar \frac{d}{dt} \mid (t) >$$
⁽¹⁾

Axiom v: Immediately after a measurement of the value a of an observable A the system is in the eigenstate $| \rangle$, corresponding to the eigenvalue a. This axiom is also called the "projection postulate", since it implies that the state $| \rangle$ before the measurement is projected onto the state $| \rangle$ after the measurement. Such a discontinuous state transition is sometimes called "the collapse of the wave function".

Distinction dynamics in quantum mechanics

If we look at these axioms from the point of view of the classical representation frame, we come to the conclusion that the only truly non-classical feature has to do with the observation process. The dynamical evolution (axiom iv) is a causal, deterministic and reversible process. The Hilbert space, apart from the superposition principle and the scalar product, which are connected to the observation process by axiom iii, behaves like a classical state space. Implicitly it is assumed that also representation structures such as objects, time, ..., behave classically (i.e. correspond to conserved distinctions, cf. Heylighen, 1990c). The only indeterminism arises during the observation process, which in classical physics is not represented within the formalism since it is supposed to be trivial.

The indeterminism of observation implies that the distinction between truth and falsity of a proposition is no longer invariant. Indeed suppose that a certain property a is observed on a system in a state which is not an eigenstate of the observable A. In that case the proposition "the system has property a" has no definite truth value before the measurement, in other words there is no invariant way to distinguish between the truth and the falsity of the proposition. It is impossible to predict which result (true or false) the measurement of the proposition will have.

The measurement itself will bring forth a new state, which is an eigenstate and which hence has a definite truth value for the measured property. Either the initial state $| \rangle$ is projected onto the eigenstate $| \rangle$, or it is projected onto some other eigenstate $| \rangle$ of A, with respective eigenvalues a = a'. The same initial state can have several distinct results, hence there is a distinction creation. But on the other hand it is possible that two initially distinct states $| \rangle$ and $| \rangle$ are projected onto the same eigenstate $| \rangle$, hence there is also distinction destruction.

We may conclude that in quantum mechanics there is conservation of all distinctions, except proposition distinctions. The non-conservation of proposition distinctions is equivalent—if the object is constant—to the non-conservation of predicate distinctions. In order to better understand this, we will look at closure in the quantum formalism.

Closure in quantum mechanics

Eigenstates as closed states

The equation which defines an eigenstate | > of an operator A is:

$$A \mid >= a \mid > \tag{2}$$

where *a* is the eigenvalue (a real number if *A* is hermitian). Often the set of eigenvalues $\{a\}$ of *A* (called the spectrum of *A*) is discrete. Since state vectors in quantum mechanics are determined up to a factor (the state is in fact defined by the ray in Hilbert space), the factor *a* before the vector does not correspond to an observable transformation of the state, and we might then as well write the equation in the following form:

$$A \mid > = \mid > \tag{3}$$

(this can be done either by taking *a* into the definition of the operator *A*, or by supposing that *A* is a projection operator, so that the eigenvalue a = 1). This equation can be interpreted in the following way: the state $| \rangle$ is invariant under the operator *A*, it is a fixpoint or attractor of the observation process represented by *A*. In other words, the state distinction determined by $| \rangle$ is closed within the larger distinction system determined by the operator *A* on the Hilbert space *S*. A similar meaning of the concept of *eigenstate*, generalized to *eigenbehavior*, is also used in second order cybernetics (see e.g. von Foerster, 1976; Vallée, 1988).

What is specific about quantum mechanics is that not all states are closed in this respect. In classical mechanics all states are implicitly supposed to be invariant under observation processes, hence closure is a trivial property. In quantum mechanics only the eigenvectors of the observable, which form just an infinitesimal fraction of the state space *S*, are invariant under the observation.

The same type of invariance under observation processes can also be seen in dynamical evolution. The Schrödinger equation (1) has the general solution:

$$| (t) > = \exp\left(\frac{-i}{\hbar}H(t - t_o)\right) | (t_o) >$$

The evolution operator $\exp\left(\frac{-i}{\hbar}H(t - t_o)\right)$ has the same eigenvectors as the Hamiltonian (energy) operator *H*. This means that if | > is an eigenvector with eigenvalue *E* of *H* then equation (1) reduces to the time independent Schrödinger equation :

$$H| > = E| > \tag{4}$$

This equation (of the same form as (2)) determines the stationary states of the quantum evolution, i.e. those states which do not change in time under normal dynamical evolution. For a bound system, such as an electron in an atom, the eigenvalues E are normally discrete (this is called the quantization of the energy levels). This means that an electron, in order to make a transition to another stationary state with a lower energy E', must emit a discrete amount of energy (quantum) under the form of a photon with energy h = E - E'. This phenomenon is called a "quantum jump": the electron suddenly jumps to a lower (or higher) level of energy.

Remark that this phenomenon cannot be described by the Schrödinger equation (1) or (4) describing the atom, since by definition the state $| \rangle$ is stable and cannot make a transition to another state, unless the dynamical operator *H* is suddenly changed by some external interference. This is one of the domains where the quantum formalism is not completely explicit about the processes it is supposed to describe (in order to describe this one must go to quantum electrodynamics). The dynamical or energy eigenstates are stable and hence correspond to point attractors of the dynamical evolution. However, the attractor basin (the states which evolve towards the stable state) is empty since the Schrödinger dynamics is reversible.

This is different for the observation process, though. The collapse of the wave function after a measurement is an irreversible process during which distinct states are projected onto the same eigenstate or attractor state | >. The eigenstate is not a real attractor of the observation process, because it is not a priori specified onto which of the eigenstates the initial state will be projected.

Linear closure

The problem with closure as it is defined now, is that it depends on the specific observation process: a state may be closed (be an eigenstate) of an operator A, but not of another operator B. This is the case when two observation operators are incompatible, which is specified mathematically by the fact that they do not commute: [A, B] = (AB - BA) 0. In that case they do not have common eigenvectors. This means physically that the measurement of A perturbs the measurement of B in a way which cannot be evaded.

We might define an operator-independent closure by requiring that a state be invariant under all possible hermitian operators, but the only vector which has this property is the null vector. We might also define a state to be closed if there exists an operator which leaves the state invariant, but then all states are closed. This type of closure is not trivial, however, if we look at *subsets* of states: $S' = \{ i \ S' | i = 1,..., n \}$ *S.* In classical mechanics, a subset of states corresponds to a proposition about the system, which can be represented as the disjunction of all atomic state propositions: "the system has state $_1$ " or "the system has state $_2$ " or "the system has state $_3$ " or ... Like all classical propositions this proposition has a definite truth value, which can be objectively determined by a set (conjunction) of observations of the system (because of the law of de Morgan, the disjunction of positive assertions can be reformulated as a conjunction of negative assertions: "the system is not in state $_1$ " ..., whose truth value is simply the negative of the truth value of the positive assertions).

In quantum mechanics, however, a state cannot be directly observed: the only "observable quantities" are the eigenvalues corresponding to the eigenvector onto which the state was projected after the observation. A state, as a proposition, can have a definite truth value only if it is an eigenstate corresponding to such an eigenvalue. A set of states can only have a definite truth value for a given observable if *all the states have the same eigenvalue*. It is always possible to define a hermitian operator which has the same eigenvalue for all the states of *S*': it suffices to consider the projection operator projecting on the subspace *T* generated by *S*'. However, it is not possible to find an observable allowing us to distinguish between *S*' and *T*. Indeed, any state | > in *T*, is by definition a linear combination or superposition of states $| _i >$ of *S*:

{
$$c_i | i = 1,...,n$$
 ; c_i complex number} such that: $| \rangle = \sum_{i=1}^{n} c_i | i \rangle$

Suppose that all $|_i>$ have the same eigenvalue *a* for a hermitian operator *A*, then because of the linearity of *A*, $|_>$ will have this same eigenvalue:

$$A \mid > = \prod_{i=1}^{n} c_i A \mid i > = \prod_{i=1}^{n} c_i a \mid i > = a \mid >$$

Hence every observable which produces the same value for the states of S', will also produce the same values for T, S' = T.

T can be viewed as the "linear" closure of S, which is generated by adding all superpositions of elements of S to S. Only sets of states which are "linearly" closed, (i.e. which correspond to subspaces of the Hilbertspace S), can be distinguished in an invariant way by a specific operator. We hence have defined a more general, operator-independent closure, which determines whether a subset of states is observable or not. In order to better understand the epistemological meaning of this formally defined property, we will now formulate quantum mechanics in a different way.

Orthogonal closure

The Hilbert space formalism with its vector space, complex numbers, hermitian operators, ..., is in practice quite complicated, so that its meaning remains rather obscure. It is however possible to reformulate quantum mechanics so that only the relations between propositions are left for study. This approach is known as *quantum logic* (Jammer, 1974). It is founded on the result of Birkhoff and von Neumann, showing that a Hilbert space is completely determined by the orthocomplemented lattice of its subspaces. A subspace corresponds to a projection operator, projecting on this space, and hence to a binary ("yes" - "no") observable or proposition about the system. We will further use the words "proposition" and "subspace" interchangeably.

In contrast to classical mechanics, the lattice of propositions in quantum mechanics is not Boolean (Piron, 1974; Aerts, 1983). This can already be understood by our previous argument,

7 QUANTUM MECHANICS

showing that a disjunction of state propositions does in general not correspond to an observable proposition, unless it is "closed", i.e. unless it corresponds to a subspace. A more fundamental way to see the difference between classical and quantum proposition systems consists in introducing the relation of orthogonality.

Definition: two propositions a and b are said to be *orthogonal*, written a = b, if a implies the negation of b (or equivalently, if b implies the negation of a).

The name of orthogonality can be understood by noticing that the subspaces corresponding to a and b are orthogonal (the product of vectors belonging to orthogonal subspaces is zero, hence according to axiom (iii) the transition probability between them is 0).

a b entails the existence of an observable *A*, such that *a* and *b* correspond to distinct eigenvalues of *A*. It suffices to take *A* as the projection onto the subspace *a*. The truth of *a* means that the state of the system is in that subspace, hence that its projection onto *a* has eigenvalue 1. The falsity of *b*, implied by the truth of *a*, means that the projection of the state onto the subspace *b* has eigenvalue 0, the state is in a subspace orthogonal to *b*. Conversely, the existence of an observable *A* such that *a* and *b* correspond to distinct eigenvalues of *A* entails the orthogonality of *a* and *b*. In conclusion, *a b* is equivalent to the statement that *a* and *b* can be distinguished by *a* single observable *A*. Non-orthogonal propositions can only be distinguished statistically, by calculating the different probabilities of observation results determined by axiom iii after repeated measurements.

Let us restrict our attention to state propositions (i.e. subspaces with dimension 1). In classical mechanics two different states 1 2 are always orthogonal, since the truth of 1 1 implies the falsity of 2. Hence orthogonality is a trivial relation in classical representations. In quantum mechanics, on the other hand, orthogonality of states is far from trivial. This can be understood from the superposition principle, which can be formulated as the following axiom:

Superposition principle: for every two distinct states $|_1 > |_2 > S$, there exists a third *superposition* state $|_1 > + |_2 > = |_3 > S$, which is orthogonal neither to $|_1 >$ nor to $|_2 >$ but which is orthogonal to all states orthogonal to both $|_1 >$ and $|_2 >$:

| > S such that $| > |_{1}$ and $| > |_{2}$, we have $| > |_{3}$.

The non-orthogonality of two states means that it is impossible to distinguish them by performing a single observation. This principle is another way of formulating the "irreducibility" or "atomic bisection" property of a quantum lattice (Ivert and Sjödin, 1978; Piron, 1974). It is not necessary to define a lattice (determined by a partial order relation with a greatest lower bound (g.l.b.) and a least upper bound (l.u.b.)) of propositions or subspaces. It can indeed be shown that the lattice of subspaces, which determines the Hilbert space, is itself determined by the set of states together with the orthogonality relation (Finkelstein, 1979). This allows us to reduce the complicated Hilbert space formalism to just a set with an orthogonality relation. The construction rests on the following definitions:

Definition: for a subset U = S of states, define the *orthogonal complement* of U to be $U = \{ | > S | | > U : | > U : | > | > \}$

Definition: define the *orthogonal closure* of U to be U.

Definition: U is called *orthogonally closed* if it is equal to its closure: U = U.

It is easily seen that has all the defining characteristics of a closure operation: monotonicity, idempotence and inclusion preservation (Heylighen, 1990a,c). It can be shown that the lattice formed by all closed subsets of the state space *S*, ordered by inclusion and with set intersection as g.l.b. and the closure of the set union as l.u.b., is equal to the quantum lattice of subspaces or propositions (cf. Finkelstein, 1979).

In fact the orthogonal closure of a set of states means that all superpositions of states in the set are added to the set. Hence orthogonal closure is equivalent to what we have called "linear closure" in the previous section. By introducing orthogonality, however, we have simplified the interpretation of the closure operation. The closure of the set U consists in the addition of all states which are indistinguishable from U by means of the orthogonality relation, in other words which are indistinguishable from U by a single observation.

The significance of closure

We have introduced first an "observation-dependent" closure, defined by the eigenstate equation (3), we have then generalized this operation by making it observable-independent, resulting in the concept of linear closure, and finally shown that linear closure can be reduced to orthogonal closure which is defined in a mathematically very simple framework, founded on a state space with the unique relation of orthogonality, which has a clear operational meaning. In this way we have reduced all the non-classical features of quantum mechanics to the single concept of orthogonal closure.

We must now interpret this formally defined concept epistemologically and physically. First we will restate the essential characteristics of any type of closure, which were introduced in a previous paper (Heylighen, 1990c). Closure of a substructure within a larger structure means that the external distinction (between the substructure and its complement or environment) is enhanced, whereas the internal distinctions (between the elements or components of the substructure) are reduced. In other words, closure diminishes the coherence between the substructure and its environment while amplifying the internal coherence of the substructure.

These two complementary effects of a non-trivial closure allow us to explain all the typical quantum effects. For example, the wave-particle duality is a typical example of a situation where on the one hand you have distinction, separation, discontinuity (particle character) and on the other hand connection, coherence, continuity (wave character). The enhancement of external distinction by closure accounts for the phenomenon of quantization, i.e. the appearance of discrete levels and hence of quantum jumps (entailing the existence of a quantum of action), as explained in section 3.1. The enhancement of internal coherence accounts for the phenomena of connectedness, implied by the superposition principle, which adds a certain coherence to classical state space, and which leads to the so-called "non-local correlations", where the measurement of one part of a distributed superposition state has an immediate "influence" on the other, spatially separate parts, like in the EPR-paradox (Einstein, Podolsky & Rosen, 1935; Jammer, 1974; Heylighen, 1990b). It seems as though everything which is mysterious about quantum mechanics can be explained simply with the help of the concept of closure.

We should perhaps also explain why these phenomena do not exist in classical mechanics, where there is also closure, but where closure is general. The fact that all elementary representation structures (objects, propositions, trajectories, ...) in classical mechanics are closed means in practice that closure becomes trivial. There is no longer any difference between (enhanced) external distinctions and (diminished) internal distinctions since all distinctions are simultaneously enhanced and diminished, so that all states, or subsets of states are equivalently coherent or incoherent. The result is that the structure of classical state space, on the level of observability of propositions, is just that of a set (cf. Piron, 1974), which is mathematically about the simplest structure we can imagine. If we want to better understand the origin of quantum phenomena we should not ask where this closure comes from (the argument of natural selection (Heylighen, 1990c) explains the existence of closed structures), but why there is only so much

closure, in other words which are the restrictions inhibiting general closure like in classical mechanics. In order to answer these questions, we must leave the framework of the quantum formalism, and analyse the processes of observation. But first we will propose an extension of that formalism, which may throw more light on the relation between formalism and process.

Probabilities and infrastates

Indeterminism and hidden variables

The fact that the result of an observation on a quantum system is generally unpredictable given the state of the system, is difficult to accept for many people. The classical world view is indeed deterministic, and assumes that any unpredictability is due to a lack of complete information about the system. This would mean that the state in quantum mechanics is not a complete determination of the system at a particular instant in time, but a partial determination lacking certain variables which are needed for an unambiguous computation of the further evolution of the system. Such an incomplete determination may be called a *macro-state*, since it only contains those properties which are *macroscopically distinguishable* by the observer. A macro-state can be viewed as an equivalence class of *micro-states*, determined by the equivalence relation "is macroscopically indistinguishable from". In order to really distinguish micro-states belonging to the same equivalence class, you would need to determine additional variables. Those variables, however, are "hidden" for the macroscopic observer.

Many attempts have been made to reduce quantum mechanics to a classical theory by introducing *hidden variables*, in such a way that the quantum state space would be reduced to a set of classes which is the result of the partition of another, much larger micro-state space (see Jammer, 1974). At first sight, such an approach would seem compatible with an analysis based on the closure concept: an equivalence class is indeed a very basic basic example of a closed structure (Heylighen, 1990a), with a clear factorization of coherence (equivalence of class members) and distinction (separation between one class and the other classes of the partition). The structure of a quantum proposition system is not that simple, however, and orthogonal closure cannot be reduced to equivalence closure. This is easily seen by remarking that the same state can belong to different orthogonally closed sets (subspaces) but not to different classes of the same equivalence relation. The impossibility of reducing a quantum state space description to a description consisting of macro-states as equivalence classes of micro-states, distinguished by additional, hidden variables, has been proven by von Neumann, and later by other authors (see Jammer, 1974; Aerts, 1986). A simple way to show why this is so, consists in comparing quantum probability with classical (hidden variable) probability.

Quantum and classical probability expressions

The probability of finding a certain measurement result *a* given a state $| \rangle$ is determined by axiom iii. This is basically a *conditional probability* P(|) of finding the eigenvalue *a* corresponding to the eigenstate $| \rangle$, knowing that the system is in state $| \rangle$, or a transition probability for the transition from the state $| \rangle$ to the state $| \rangle$ after the measurement. It can be shown immediately that this probability is not classical. Classical probability theory, which was axiomatized by Kolmogorov, presupposes the axiom of Bayes:

$$P(a \mid b) = \frac{P(a \& b)}{P(b)}$$
(5)

P(a & b) is here the probability of *a* and *b* both being true. This axiom is clearly not obeyed by quantum mechanics. For two non-orthogonal states | > | >, the proposition & is never true, since the system by definition cannot be in two different states at once, and hence we have:

$$P(||) = |\langle || \rangle|^2 - \frac{P(\&)}{P(|)} = 0$$
(6)

The rationale for the Bayes postulate comes from the idea that indeterminism is due to incomplete information about the state of the system. The probability of finding a certain outcome *a* is then equal to the number of "hidden" microstates for which *a* is true, divided by the total number of microstates. Call the micro-state space , and consider a measure *M* on . Define: $_{a} = \{$

| *a* is true for the state }. The probability of *a* is then given by:

$$P(a) = \frac{M(a)}{M(a)}$$
(7)

The probability of *a* and *b* both being true is equal to the "number" of micro-states implying both *a* and *b*, and hence to:

$$P(a \& b) = \frac{M(\underline{a} \ \underline{b})}{M(\underline{b})}$$
(8)

In the case of conditional probability the probability of finding *a*, knowing that *b* is true, P(a | b), is equal to the number of microstates in which *a* and *b* are true, divided by the number of states for which *b* is true:

$$P(a \mid b) = \frac{M(a \mid b)}{M(b)}$$
(9)

By applying (7) and (8), we now can derive the Bayes axiom (5). Formula (9) is more general than the original axiom (5), however. Indeed, it does not contradict the quantum probability expressed by axiom iii. It suffices to interpret the quantum states $| \rangle$ and $| \rangle$ as sets of microstates and , but such that those sets have a non-zero intersection, with M() proportional to $|\langle | \rangle|^2$.

Infra-states and the superposition principle

This interpretation is different from the traditional hidden variable interpretation mentioned in the previous subsection, since the latter approach sees the macro-states $| \rangle$ and $| \rangle$ as disjointed subsets of micro-states, corresponding to different equivalence classes of the same partition. Indeed, the hidden variable approach presupposes that properties which are distinguished in the quantum representation will also be distinguished in the lower-order micro-state representation (cf. Aerts, 1986), and hence distinct quantum states should correspond to distinct micro-states. In order to avoid confusion with the concept of micro-states as it is used in hidden variable theories, we will from now on use the term introduced by Finkelstein (1979), and call the elements of "infra-states".

If we replace equivalence closure by orthogonal closure, however, the non-disjointedness of and does not seem so strange any more. and can be interpreted as orthogonally closed subsets of the infra-state space , and as the smallest such sets. Indeed, within the quantum representation the states are the smallest directly distinguishable—and hence orthogonally closed—subsets. It is not possible to make finer distinctions, and this explains why the conjunction of two states, & , is a proposition which cannot be observed at all, and to which we have therefore given the probability 0 in (6).

In order to clarify this, we must redefine orthogonality within the infra-state framework. The orthogonality of quantum states (meaning that their transition probability is 0) corresponds, by relation (9), to the disjointedness of their corresponding sets of infra-states: implies

= . By extension, the same property may be assumed for propositions or subspaces containing several quantum states. The opposite implication is not true, however: the disjointedness of two infra-state sets is not sufficient to conclude that the macro-propositions they represent are orthogonal, in the sense of "distinguishable by a single observation". Indeed, it suffices to consider sets containing just one infra-state, and these are, by definition of the concept "infra-state", indistinguishable. This might be explained by assuming some "uncertainty principle", stating that it is impossible to distinguish a part of the infra-state space with a "volume" or "measure" smaller than some fixed constant $\hbar = M$ () where | > is a conventional quantum state.

By postulating the existence of such infra-states with an uncertainty principle we can now derive the superposition principle, instead of having to introduce it axiomatically. It suffices to define a quantum state as any subset of the infra-state space with volume equal to \hbar :

with $M() = \hbar$, corresponds to a quantum state | >.

Theorem: the set of all quantum states, defined in this way, satisfies the superposition principle.

Proof:

consider two quantum states $|_{1} > |_{2}$, characterized by the infra-state sets $|_{1} > |_{2}$ with $M(_{1}) = M(_{2}) = \overline{h}$. It is then always possible to find a third set $|_{3} = 1$, $|_{2}$, such that $M(_{3}) = \overline{h}$ with $|_{3} = 1$, and $|_{3} = 2$. The state $|_{3} >$ determined by this set is then orthogonal neither to $|_{1} > \operatorname{norto} |_{2} >$, since: $|_{3} = 1$ and $|_{3} = 2$.

Moreover $|_{3}$ is orthogonal to all states $|_{4}$ orthogonal to both $|_{1}$ and $|_{2}$: suppose $|_{4}$ $|_{1}$ and $|_{4}$ $|_{2}$, then $|_{4}$ $(_{1}$ $|_{2})$ = , but then also $|_{4}$ $|_{3}$ = and hence $|_{4}$ $|_{3}$.

₃ can hence be interpreted as a superposition of $|_1 >$ and $|_2 >$.

Though it is possible to reconstruct the quantum states and their superposition by assuming the existence of a minimally observable infra-space volume \hbar , and by reducing orthogonality to disjointedness of infra-sets, it is still not possible to reconstruct the Hilbert space structure, determined by the lattice of orthogonally closed subspaces. Disjointedness alone is clearly insufficient to define orthogonal closure, because it would imply that every set of infra-states, being disjointed from the other sets it is disjointed with, is orthogonally closed. The "uncertainty principle" only allows us to exclude infra-sets smaller than states, but it does not exclude unions of quantum sets which do not constitute subspaces. In order to better understand the physical and epistemological meaning of closure in quantum mechanics, we must analyse the observation process and thus try to better understand what an infra-state means.

An analysis of the observation process

Observation and perturbation

Quantum mechanics is a very good example of the problems to which a (more or less naive) realist epistemology leads. For example, the position of an electron bound to an atom is, according to the laws of quantum mechanics, indeterminate. This is something which is very difficult to accept for a realist, since he assumes that position is a really existing property which the electron must have, independently of the observer. In the present constructivist framework, on the other hand, position is a predicate, i.e. a component of a representation of the phenomenon constructed by the observer. The distinction or determination of the value of a predicate is an interaction process between the observer and the observed object. The result of this process, which may be a value for the position, is something which did not exist before the process started, in other words which has been "constructed" during the interaction process. Hence there is no paradox in the assertion that the position of an electron is indeterminate before an observation process is carried out. It does not make sense to ask *where* the electron was before the observation, since the question "where?" presupposes a system of distinctions (coordinates), dependent upon the observer, which must be filled in during an observation process.

The fact that in quantum mechanics the effect of such an observation process cannot be abstracted away is often expressed by remarking that the process necessarily perturbs the state of the object. Indeed, the existence of a quantum of action, as a minimal discrete amount of energy exchanged during microscopic processes, entails a discontinuous change of state after such a process. Since the energy of a typical quantum object (electron, atom, ...) is about the same order of magnitude as the quantum of action, this discrete change cannot be neglected. Hence it is impossible to perform an observation on a quantum system without perturbing it in a substantial way.

The observation problem in quantum mechanics is not due to the perturbation, however. In classical mechanics too, it often happens that a measurement perturbs the measured object. For example, you might measure the mass of a billiard ball by subjecting it to a known force during a known time interval and then computing, by means of Newton's law, the mass from the observed change of velocity of the ball. Evidently the ball has been perturbed by the observation. Yet we know what the perturbation was: we know that the mass has remained the same during the measurement, and that the velocity and the position have changed in accordance with Newton's law. Hence we can reconstruct the state of the ball before the measurement took place.

Indeterminacy of the perturbation

This is no longer true in quantum mechanics. Here the perturbation is indeterminate, and the result of the observation does not allow us to reconstruct the state before the observation took place. This can be illustrated by a classical thought experiment proposed by Bohr (1958; see also Jammer, 1974). Suppose we want to observe a particle by means of a set-up consisting of a diaphragm with a slit. If the diaphragm is rigidly connected to a frame, the position of the slit is fixed, and hence it is possible to determine the position of the particle (up to the arbitrary small width of the slit) by letting it pass through the slit. However the momentum exchanged between micro-object and diaphragm cannot be observed, because of the rigid connection between the diaphragm and the ideally unmovable (i.e. infinitely heavy) frame. The change of momentum of the frame caused by the interaction with the particle is, because of the heaviness of the frame, macroscopically indistinguishable.

Suppose that we would now want to measure the momentum of the particle using the same components. This could be done by using a light diaphragm connected to the frame by weak springs, so that it can move when hit by the particle. If the diaphragm is light enough, its change of momentum becomes macroscopically distinguishable, and hence we can determine the original momentum of the particle. However, the "movability" of the diaphragm signifies that its position can no longer be determined exactly. In fact, the easier the diaphragm can move, the more accurately we can measure the momentum, but the less accurately we can measure the position.

This complementarity of momentum and position observations can be expressed by a Heisenberg type of indeterminacy relation.

Let us try to understand what happens from the point of view of distinction-making. What we are doing when we let a microscopic particle interact with a macroscopic object (diaphragm) is to "magnify" (i.e. make macroscopic) a certain property of the particle. We somehow "force" the particle to make a choice between two alternatives (passing through the slit or not passing through the slit), which are distinguishable on a *macroscopic* scale (because the diaphragm is large enough to be seen by the observer) and in an unambiguous, *invariant* way (because the frame is rigid so that the position of the slit cannot vary). Bohr's reasoning shows that we cannot magnify all properties simultaneously.

We cannot magnify the second property (momentum) after the first one either, because the observation has perturbed the system: passing through the slit has changed the momentum of the particle. In order to know how the momentum has interacted with the apparatus, we should know the microscopic state of the observation apparatus before the experiment began. By definition the macroscopic observation apparatus is only macroscopically known, so that normally the observer does not know this state. He could of course magnify the state by using a second observation set-up for observing the first one, but the state of this one too would only be determined macroscopically. Hence it would perturb the state of the first one in an undetermined way, and be unable to completely magnify the state of the first apparatus. We could continue in this way adding more and more observation set-ups in which each new one would try to determine the microscopic state of the previous one, but this would just lead to an infinite regression, and we would not be able at any point to eliminate the remaining indeterminacy. This analysis may be summarized by remarking that *an observation apparatus (considered as an extension of the observer) cannot observe its own micro-state*, and this can be interpreted as an instantiation of the principle of the impossibility of complete self-reference (cf. Heylighen, 1990b,c; Löfgren, 1990).

Dynamics of the observation process

Let us now analyse the dynamics of the observation process in the infra-state space. Suppose the quantum system is in a state $| \rangle$ that corresponds to a set of infra-states, one of which (say

) is actual for the system. Suppose that a property is observed, such that \mid > is an eigenstate of the corresponding observable. In that case, according to the projection axiom (v), the state after the observation will still be \mid >, and hence the infra-state $\ '$ after the observation must still belong to $\$, though—since we assume some perturbation of the infra-state by the observation apparatus—it is not necessarily the same one: in general

Suppose now that another, incompatible property is observed, say an observation testing whether the system is in a quantum state | >, different from, but not orthogonal to | >. This state will correspond to another infra-state set , with , but . If the original , then the observation will give the result "yes", otherwise infra-state belongs to that set: the result will be "no". This is in accordance with the rationale, expressed by equation (9), for introducing the infra-state: the probability for finding given is proportional to the number of have in common. Again we may assume that the infra-state after the infra-states and : " measurement will have been perturbed, but will now belong to instead of

Let us finally assume that an observation is made of a state | > orthogonal to the original state | >: = , but In that case, according to (9), there is a non-zero probability to find "yes" for the observation, although the answer would have been "no" for the state | > we originally started with. This means that the infra-state has been perturbed once again, leading to a state ""

The dynamics generated by the coupling of the quantum system to the apparatus, resulting in an interaction, can thus be understood as having the following features: a specific observation set-up induces a specific partition of infra-state space, expressing the macroscopic distinctions between regions that observable by the set-up, with the classes of the partition corresponding to attractors. An *attractor* is a region of state space which is invariant under the dynamics (each state in the attractor is sent upon a state of the same attractor), but such that it does not contain subattractors. This means that there are no fixpoints inside the attractor: each point in the attractor is sent upon another point of the attractor. This accounts for a model in which there is both conservation of the distinctions between different attractors, and variation within each attractor. Since we have argued that the perturbation of the infra-state by the apparatus is macroscopically indeterminate, we may assume that the variation within the attractor is not observable. Even though performing the same observation repeatedly results in always the same macro-state, we may assume that the infra-state is perturbed in an indeterminate way during each such observation, remaining however within the boundaries of the attractor. On the other hand, if another, incompatible observation is performed, the same process of redistribution of infra-states within their attractors is executed, but this time the attractors are different, so that the variation may now move the infra-state out of the attractor it originally belonged to. If that happens, it is possible to find a third observation of a state | > which is orthogonal to the first one | >, yet which gives a positive result, since the infra-state has been moved to its attractor corresponding to the result "yes".

This analysis allows us to propose yet another, dynamical interpretation of the concept of orthogonal closure: a subset of infra-states is (orthogonally) closed if it forms an attractor of the dynamics induced by the interaction with an observation apparatus. The indeterminacy principle can be interpreted as specifying that such an attractor cannot be smaller than a specified discrete volume, since there always is an indeterminate perturbation of infra-states within the attractor. A quantum state then corresponds to an attractor having this minimal volume. Since different attractors, the same infra-state can belong to different macrostates.

This interpretation is more specific than the one proposed at the end of the last section, where the superposition principle was derived, since it is clear that not every subset with this given volume can be an attractor. Indeed, if we assume continuity for the dynamical process, then an attractor cannot consist out of disconnected subsets or points: it must be a topologically connected region. In order to be yet more specific, we would need a detailed dynamical model of the observation process.

A possible realization of such a model was proposed by Aerts (1986) on the example of a spin measurement. The model is characterized by two parameters, one representing the determinate micro-state (spin) of the quantum system, another one representing the indeterminate state of the measuring apparatus ("charge" or "force" influencing the micro-state). During the observation, the micro-state is attracted towards a fixed micro-state (point attractor), corresponding to the state where the observation is determinate. The force which attracts the micro-state to the attractor is indeterminate, however, so that different measurements of the same micro-state may end up in distinct point attractors, and hence give distinct results. Aerts was able to reconstruct the quantum expression for the probability of spin measurements on the basis of these assumptions. The peculiarity of this model is that the hidden variable, which would allow to predict the process if it were known, does not belong to the object to be observed but to the observing instrument. A rationale for this may be found in our analysis of Bohr's thought experiment.

This is not a real infra-state model, however. In Aerts's model, the result of the observation depends directly upon the indeterminate state of the apparatus. In our infra-state model, on the other hand, the result of the observation merely depends on the infra-state at the instant the observation begins: the apparatus state can only have an effect on the result of the subsequent

15 QUANTUM MECHANICS

observation. This apparent inconsistency may be disentangled by contrasting a realist interpretation of states with a constructivist one.

A constructivist interpretation of the observation process

In a realist interpretation a state is an intrinsic property of a system, which is independent of the observer. In a constructivist interpretation a state is a construct resulting from an interaction between the system and the observer, which is such that it allows the observer to find a maximum of coherence between the different results of observations (i.e. subject-object interactions). In general, no construction (representation) is optimal, and the advantages and disadvantages of a particular representation may be counterbalanced by different advantages and disadvantages of another representation.

In the present discussion we have considered four representations of quantum phenomena:

1) the traditional Hilbert space representation which has the advantage of being coherent with empirical observation results, but the disadvantage of being indeterminate (no distinction conservation) as far as observations are concerned, and of being formally complicated and conceptually vague;

2) a traditional "hidden variable" representation, which is conceptually clear, but inconsistent with the previous representation and hence incoherent with the empirical results;

3) an "infra-state" representation which is coherent with the Hilbert space representation, and which has the advantage of being formally and conceptually much simpler by proposing an "infra"-state space where observation processes are determinate (there is distinction conservation), but the disadvantage that the infra-state construct is not observable, (though the derived construct of an orthogonally closed subset of infra-states is);

4) the Aerts representation, which appears to be coherent with the Hilbert space representation, and which has advantages and disadvantages similar to those of the previous representation.

The difference between the latter two representations resides in the behavior of the infra- or micro-state. In the Aerts model the micro-state can be prepared: it is in a determinate point attractor after the observation, but it cannot be detected, since the micro-state is perturbed by an indeterminate apparatus state during the measurement. In the infra-state model the infra-state can neither be prepared nor detected: we can only determine that it belongs to an attractor set of the observation process, where it undergoes indeterminate transitions during the process. The advantage of the infra-state model is that its mathematical structure is very simple. The advantage of the Aerts model is that it proposes a more explicit analysis of the observation process.

We may further try to synthesize the two models. The infra-state can be reinterpreted in the Aerts model as the relation between the micro-state of the system and the micro-state of the apparatus. When determining a property such as the position or velocity of a particle, it is clear that there is no absolute position or velocity, but only a position or velocity with respect to the position or velocity of the observer and his observation set-up. Looked at in a relational way, it is no longer meaningful to ask whether the indeterminacy or impossibility to observe resides in the system or in the apparatus: the only thing that matters is that we cannot exactly determine (neither prepare nor detect) the relation between them.

During this whole discussion, we must not forget what we are trying to do. We are not trying to find the one true representation of what happens during the observation process at the microscopic level: this is by definition unobservable. The representation we try to find is one which is coherent with the empirical results, and which is moreover maximally simple and generalizable. The generalization we are interested in, is the one from a model of quantum observation to a model of the emergence of macroscopic distinctions out of (unobservable or less easily observable) "microscopic" distinctions.

On the basis of the microscopic causality principle (Heylighen, 1989a) we may trivially postulate that at the lowest, elementary level of existence there are conserved distinctions, which, however, are by definition unobservable since their domain of invariance does not extend beyond a single event. We may interpret these distinctions as infra-states. At the same time we must remark that macroscopic distinctions are *emergent* with respect to those microscopic distinctions, and hence that there is no a priori way to infer the conservation, creation or destruction properties of these observable distinctions from the behavior of the microscopic ones, apart from a general principle of "natural selection" (Heylighen, 1989b). What remains to be done is to develop a more detailed model of the mechanism of this "emergence", "creation" or "construction" of macroscopic distinctions during the observation process. The concept of orthogonal closure as defining an attractor of the observation process seems to provide a first step towards the development of such a model.

Conclusion

The conceptual and formal structure of quantum mechanics has been analysed from the point of view of the dynamics of distinctions. The only distinctions which vary in quantum mechanics are the distinctions between the truth and falsity of a proposition, consisting of a variable predicate attributed to an invariant object. This follows from the axioms of quantum mechanics, which entail the general indeterminacy of observation results. The complicated structure of the Hilbert space formalism, with the superposition principle, the scalar product of vectors, the self-adjoint operators and their eigenvalues and eigenstates, was simplified with the help of the concept of closure. It was shown that the concept of eigenstate is equivalent to the closure of a certain state under an operator. This closure was made operator-independent by demanding that there exist an operator leaving a set of states invariant. Such state sets were called linearly closed. Linear closure was further simplified to orthogonal closure, meaning that a set of states can be distinguished by a single observation. The complete Hilbert space structure can thus be reduced to a state set together with an orthogonality relation.

The non-classical probability entailed by the quantum axioms was reconstructed by introducing the concept of infra-state. Quantum states can be seen as (generally overlapping) sets of unobservable infra-states, with the transition probability between two states proportional to the number of infra-states they have in common. Orthogonality corresponds to the disjointedness of infra-state sets. It was hypothesized that only closed sets of infra-states can be distinguished, but this demands a concept of closure which is more detailed than that of orthogonal closure, since the infra-state interpretation of orthogonality entails trivially that all disjointed sets of infra-states are orthogonally closed, while by definition not all such sets are observable. It was proposed to add an uncertainty principle stating the impossibility to distinguish sets with a measure smaller than some fixed constant. This allowed us to reconstruct the superposition principle.

An analysis of the observation process led us to interpret closed sets of infra-states as attractors of the dynamics induced by the interaction with the observation apparatus. This interaction is always partially indeterminate, because of the unknown micro-state of the apparatus, and this leads to an indeterminate redistribution of infra-states within the attractor. Incompatible set-ups lead to different partitions of the infra-state space into attractors, and this explains how macro-states may change during sequences of incompatible observations.

This model of the emergence of macroscopic distinctions during the quantum observation process is much more explicit and understandable than the complicated and counterintuive Hilbert space model following from the axioms of quantum theory. In particular the concept of closure, with its complementarity of external distinction and internal coherence, proposes a simple interpretation for the basic quantum duality between the continuous (waves, superpositions) and the discrete (particles, quantum jumps). Moreover, the model seems more general and easier to apply to problems outside the domain of particle physics, since it only assumes distinctions and closure, instead of specialized concepts like wave functions, scalar products or self-adjoint operators. However, the model is still not sufficiently precise, lacking a clear understanding of the structure of the attractors representing the observation process and a complete interpretation of the concept of infra-state. One possible realization was proposed by reinterpreting a model presented by Aerts (1986), but there remain a lot of questions. What should be done now is to build a more detailed model of the dynamics of observation in quantum mechanics, as a special case of the more general dynamics of distinctions, with an emphasis on the "natural selection" of emergent, closed distinctions.

References

- Aerts D. (1983): "Classical Theories and Non Classical Theories as Special Cases of a More General Theory", J. Math. Phys. 24, p. 2441.
- Aerts D. (1986): "A Possible Explanation for the Probabilities of Quantum Mechanics", J. Math. Phys. 27, p. 202.
- Bohr N. (1958): Atomic physics and human knowledge, (Wiley, New York).
- Einstein A., Podolsky B. & Rosen N. (1935): "Can the Quantum Mechanical Description of Physical Reality Be Considered Complete?", Physical Review 47, p. 1804.
- Finkelstein D. (1979): "Holistic Methods in Quantum Logic", in: Quantum Theory and the Structures of Space and Time 3, Castell L., Drieschner M. & von Weizsäcker C.F. (eds.), (Hanser, Munich), p. 37-59.
- Heylighen F. (1989a): "Causality as Distinction Conservation: a theory of predictability, reversibility and time order", Cybernetics and Systems 20, p. 361-384.
- Heylighen F. (1989b) : "Self-Organization, Emergence and the Architecture of Complexity", Proc. of the 1st European Congress on System Science, (AFCET, Paris), p. 23-32.
- Heylighen F. (1990a): "Relational Closure: a mathematical concept for distinction-making and complexity analysis", in: Cybernetics and Systems '90, R. Trappl (ed.), (World Science Publishers, Singapore).
- Heylighen F. (1990b): Representation and Change. A Metarepresentational Framework for the Foundations of Physical and Cognitive Science, (Communication & Cognition, Gent).
- Heylighen F. (1990c): Classical and Non-classical Representations in Physics.
- Ivert P-A. & Sjödin T. (1978): "On the Impossibility of a Finite Propositional Lattice for Quantum Mechanics", Helvetica Physica Acta 51, p. 635-636.
- Jammer M. (1974): The Philosophy of Quantum Mechanics, (Wiley, London).
- Kornwachs K. (1988): "Cognition and Complementarity", in: Nature, Cognition and System, M.E. Carvallo (ed.), (Kluwer, Dordrecht).
- Löfgren L. (1990): "On the Partiality of Self-Reference", in: Self-Steering and Cognition in Complex Systems, Heylighen F.. Rosseel E. & Demeyere F. (eds.), (Gordon and Breach, New York), p. 47-64.
- Piron C. (1976): Foundations of Quantum Physics, (W.A. Benjamin, Menlo Park, California).
- Vallée R. (1987): "Sur les "éléments propres" de Heinz von Foerster", Revue Internationale de Systémique 1, p. 37-46.