Philosophical Problems

of

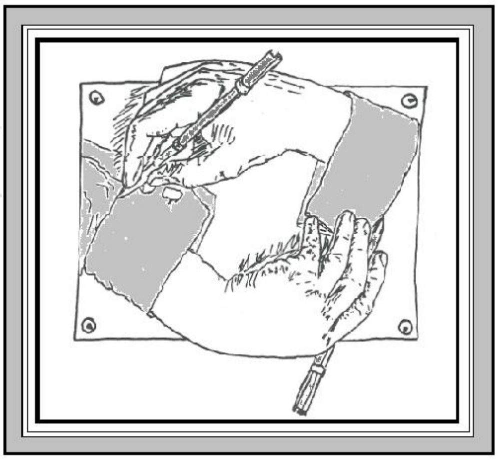
QUANTUM ONTOLOGY

by

GRAEME DONALD ROBERTSON
B.Sc. Physics, (Edin.) Part III Maths, (Cantab.)

A dissertation submitted to the University of Cambridge for the degree of Master of Letters.

Trinity College 1976



Who is as the wise man? and who
knoweth the interpretation of a thing?
Ecclesiastes 8:1

Contents

Preface

This dissertation develops the question; what is a physical object according to the theory of quantum mechanics? The first answer to be considered is that given by Bohr in terms of the concept of complementarity. According to this interpretation of quantum theory (QT) a quantum object in certain experimental circumstances behaves like a classical particle and other complementary circumstances behaves like a classical wave. To ask what an object in itself *is* is then considered to be meaningless since a quantum object cannot be conceptually dissociated from, and is only defined in relation to, a given experimental arrangement by means of which the object is observed. This interpretation is illustrated by way of an example, the two slit experiment, which highlights some of the associated problems of ontology.

One such problem is the so-called problem of measurement or observation. Various interpretations of measurement in QT, including those of Heisenberg, von Neumann, Everett and Bohr, are compared and contrasted. In this context the Schrödinger cat paradox arises.

A second problem concerns the question of whether or not QT can be considered complete and therefore satisfactory as a basis for physics. According to Einstein, Podolsky and Rosen (EPR), given what they consider a satisfactory criterion of physical reality, QT does not provide a complete description of physical reality. Einstein's interpretation of QT is discussed and Bell's theorem, which sets limits to the ways in which QT might conceivably be completed, is proved. Bohr's reply to the EPR paper is analysed and some consequences as regards the classical conception of physical reality are indicated.

A recent approach to QT has been the search for another new theory which is complete in the sense of EPR. Various attempts to complete QT by means of the addition of 'hidden variables' to the quantum mechanical state function are considered and their aims and achievements assessed. For the most part these theories are found to be unacceptable and are in any case unable to restore all the necessary ingredients of the classical conception of reality. A quantum theory is then outlined which would seem to fulfil the role of a successful hidden variables theory (hvt) except in so far as it accepts the orthodox interpretation of QT.

Up until Chapter Five only non-relativistic QT has been considered. Although it is well known that relativistic QT (RQT) contains mathematical incongruities, it does nevertheless make predictions which can be tested and are in extremely close agreement with observation. I therefore investigate some of characteristic ontological problems for the orthodox interpretation of RQT; creation and destruction of matter and the existence of 'virtual' matter. The basis of the S-matrix approach, proposed by Heisenberg to avoid the mathematical difficulties of conventional RQT is explained, and finally, philosophical issues relating to the resulting possibility of 'bootstrapping' matter are examined.

Since this dissertation is primarily concerned with ontological questions, the logico-algebraic approach, initiated by Birkhoff and von Neumann (1936) and which has recently received increasing attention from philosophers of science (see Reichenbach (1944); Putnam (1969); Hooker ed. (1975) etc.) is here almost entirely omitted.

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration.

Acknowledgements

I am greatly indebted to Professor Mary Hesse for her continual help and encouragement throughout the course of this research and for her valuable comments and criticisms of an earlier draft of this dissertation.

Part of the work was done while the author was visiting the Physics Department of the University of the West Indies, Jamaica. I would like to thank Professor Geoffrey Kent for his hospitality and kindness during my stay. I would also like to thank Dr Neville McMorris for a number of extremely interesting discussions.

Thanks are also due to my friends Mr Roderick Macdonald and the late Dr Gary Rodgers who have provided the opportunity for many stimulating conversations on important issues and especially my friend Miss Leslie Patterson, the secretary.

Finally, for material and financial support, I have pleasure in thanking Trinity College, the University of the West Indies and, above all, my parents.

Introduction

"The mathematical equipment of the [new quantum] theory," writes Heisenberg, "was complete in its most important parts by the middle of 1926, but the physical significance was still extremely unclear."

Now, half a century later, the discussion of the interpretation and solution of fundamental problems raised by the theory continues unabated. Out of a variety of early interpretations, each one has turned out to be unsatisfactory. Only the radical and penetrating interpretation of Niels Bohr has remained substantially intact. From about 1935 to 1952 practically all physicists, with the notable exceptions of Einstein and Schrödinger, accepted the 'Copenhagen Interpretation' (CI) which is not a single comprehensive doctrine as the name seems to imply, but whose basic ideas are set down in the crucial papers by Bohr (1925, 1927, 1929), Heisenberg (1930), Born and Jordan (1930), Dirac (1930), von Neumann (1932; English translation published in 1955) and Pauli (1933). These have been, with some justice, together described as an 'orthodoxy' and regarded as being of decisive importance to the whole subsequent development of the discussion of quantum principles. Apart from the early objections to the orthodox interpretation of QT from physicists such as Einstein, de Broglie and Schrödinger, which can be seen from the Fifth Solvay Congress (1927) proceedings, there was also early dissatisfaction with certain basic features of the orthodox interpretation from the more philosophically oriented side, eg from Margenau (1931, 1932) and Popper (1935). This initial criticism of the orthodox Copenhagen Interpretation was, however, unsuccessful, and for the first Quarter of a century the orthodox view largely held the field, in particular it was accepted in the textbooks. Complementarity is a fundamental principle at the heart of both Bohr's distinctive interpretation of QT and the CI as a whole.

In the last Quarter of a century philosophical debate has resumed with renewed vigour, primarily due to two papers by Bohm (1952a & b) which for the first time, propound a reasonable theory of classical-type hidden parameters for QT. This can be seen as a follow up of the ideas of Einstein (1935, 1949) and of the early ideas of de Broglie (1927). This new line of research has made rapid progress and has led to the proposal of a variety of hidden variable theories which seek to restore a causal interpretation to QT. Although none of these theories has yet led to any definite physical results, they have had the effect of showing that the CI is not the only possible interpretation of QT as was believed by many writers. For example, Jordan wrote: "There is only one interpretation (ie the CI) which is capable of conceptually ordering the totality of experimental results in the field of atomic physics." This belief was mainly a result of the work of von Neumann who, it was thought, had proved conclusively in 1932 that it is impossible to restore causality to QT by means of the addition of heretofore unknown dynamical variables to the quantum description of the state of a system. Bohm's great achievement lies in his demonstration that von Neumann's definition of an hvt is not the only one but rather an inappropriate one - but it does seem causality cannot be restored (see Ch 4, II). The irrelevance of von Neumann's proof with regard to Bohm's definition of hidden variables has thus reopened the debate on the interpretation of QT including the celebrated Bohr-Einstein debate on the completeness of QT.

In addition to the appearance of more or less direct opposition to the orthodox school from hidden variable theorists there is now a second critical approach which, together with the first, forms a much more lively stage in the development of the philosophical problems raised by QT. Authors of this approach, which include Feyerabend (1962, 1968), Landé (1965), Putnam (1965), Popper (1967), Bub (1973) and Hooker (1973), do not seek to return to the concepts of classical physics. They agree with the orthodox view to the extent that QT is an essentially new type of physics, and they are therefore closer to orthodoxy than are the hidden variable theorists. The predominant trend in this approach has been an increasing stress on and analysis of the probabilistic character of QT, together with a firm rejection of the subjective interpretations which offer themselves for the probabilities as an expression of lack of knowledge. Also these writers are more willing than the original adherents of the CI to admit that none of the arguments behind the CI is powerful enough to guarantee the absolute validity of this interpretation and to justify the claim that the theories of the microscopic domain will forever have to conform to a certain pattern. "Such restrictions are possible," writes Feyerabend, "only if certain philosophical ideas are used as well. The result of the (philosophical) criticism is, of course, that now as ever we are free to consider whatever theories we want when attempting to explain the structure of matter." Many physicists are still opposed to this result however. For example, Rosenfeld (1957) writes, "Recent criticism of the foundations of QT originates from a number of physical and epistemological misconceptions." Feyerabend (1962 pp192-4) cites three reasons for the persistence of CI in the face of objections:

(i) Partial empirical success has a soporific effect on many physicists who are not, in general, very fond of philosophy and to whom the ideas of the CI have come to appear to be the expression of physical common sense.

(ii) The main principles of the CI are vague. This vagueness has allowed complementarity to be reinterpreted in the light of objections and still give the impression that these objections had really missed the point.

(iii) Many of the objections against complementarity are irrelevant since they see complementarity as the direct result of a positivistic epistemology which they reject and neglect its physical content.

These two principal directions by no means exhaust the range of critical activity relating to the orthodox interpretation of QT during the last Quarter of a century. More extensive reviews are to be found in Heisenberg (1955), Bunge (1956), d'Espagnat (1971) and Jammer (l974) on the whole subject of fundamental problems in QT. Since the appearance of Bohm's rival interpretation, a number of distinct interpretations of QT have appeared in the literature. These include Everett (1957), Witmer (1967), Jauch (1968), Ballentine (1970), Stapp (1971), Weingarten (1971), Bedford & Wang (1975) and others. Since there are now so many varieties and shades of interpretation of QT and since, as Groenewold (19714 p353) points out, even "those who might be called adherents (to the CI) have quite divergent opinions about foundations and interpretations of quantum mechanics", it is, perhaps, both reasonable and timely (see Feyerabend (1969 p103)) that this dissertation should begin with a critical reappraisal of the concept of complementarity, revealing its physical basis and philosophical assumptions.

Chapter One: Complementarity

I

Bohr has often been criticised for his vagueness of expression and lack of terminological definiteness in his publications relating to the foundations of QT: de Broglie referred to his "obscure clarity" and called him the "Rembrandt of contemporary physics". One could even be led to suppose, on an initial reading of Bohr, that his writings on the subject are no more than popular versions which spare an uninitiated public all technical details. The way in which Bohr expressed his thoughts, however, is undoubtedly the one that he considered appropriate to the subject. Two comments by Heisenberg may help to explain Bohr's unusual style of argument on matters of physics. Heisenberg (1967 p98) writes, "I noticed that mathematical clarity had in itself no virtue for Bohr. He feared that the formal mathematical structure would obscure the physical core of the problem, and in any case, he was convinced that a complete physical explanation should absolutely precede the mathematical formulation." Heisenberg (1967) also notes that "Bohr was primarily a philosopher, not a physicist, but he understood that natural philosophy ... carries weight only if its every detail can be subject to the ... test of experiment."

It seems that Bohr chose to present his ideas in the way that he did because of his belief that the concept of complementarity has very general application and can even be applied in fields outside the realm of pure physics. A particularly striking example which corroborates this can be seen from a study of Bohr's Como lecture at the 1927 International Congress of Physics which was convened in the Italian city of Como, where Volta was born and died and in commemoration of the centenary of his death. Here, for the first time, Bohr presented in public his ideas on complementarity. But nowhere in this lecture did Bohr refer to position and momentum as complementary quantities, although on various occasions he could easily have done so. This might seem surprising since an earlier definition of complementarity had been given by Pauli who defined it as the relation between non-commuting variables (eg position and momentum). Bohr, however, attempted to give a more general exposition of the concept although in later writings he explicitly accepts that position and momentum are complementary quantities.

The following two points should be made relating to subjects outside the realm of pure physics. According to Jammer (1966 p176), Bohr obtained the concept from William James (1890 Vol I, p206). James uses the word complementarity in a psychological context which Bohr appropriates for use in the interpretation of QT. James writes: (just as) "in certain persons, at least, the total possible consciousness may be split into parts which coexist but mutually ignore each other, and share the objects of knowledge between them," so in physics certain cognitions which coexist and harmonise in the classical theory are split, in QT, into mutually exclusive views complementary to each other. Developing James' use of the term in psychology, Bohr writes (193); p96): "A complete elucidation of one and the same object may require diverse points of view which defy a unique description. Indeed, strictly speaking, the conscious analysis of any concept stands in a relation of exclusion to its immediate application." The first point, then, is that the very concept of complementarity has appeared in the first place from outside the realm of physics, from psychology. The second point relates to the application of the concept to other subjects since its success in physics. Bohr himself was the first to extend the use of complementarity to biology. In an address delivered to the

International Congress on Light Therapy held in Copenhagen on August I5, 1932, Bohr explained that the revision of the foundation of physics, "extending to the very question of what may be meant by a physical explanation, has not only been essential for the elucidation of the situation in atomic theory, but has also created a new background for the discussion of the relation of physics to the problems of biology ... We should doubtless kill an animal if we tried to carry the investigation of its organs so far that we could describe the role played by single atoms in vital functions ... On this view, the existence of life must be considered as an elementary fact that cannot be explained, but must be taken as a starting point in biology, in a similar way as the quantum of action, which appears as an irrational element from the point of view of classical mechanical physics, taken together with the existence of elementary particles, forms the foundation of atomic physics." Since then complementarity has been applied in such fields as biology, psychology (see Jung (1960 p287)), ethics and theology (see Feyerabend (1962 p191)) and linguistics and economics.

While in Bohr's writing there is no single formulation of QT based entirely and consistently on the principles proposed by him, the reason for this is not simply that he wished to allow for further developments of these principles in physics as well as other fields, although it is probably true that he felt a more formal presentation of his ideas might inhibit creative thought and thereby prevent progress in the understanding of QT. The reason is rather that Bohr was unable to find a unified and logical formulation of QT. He suspected that a more general formulation of complementarity would have to be found before this could be done. As it stands, Bohr's account of complementarity seems to cover a whole spectrum of doctrines; it has been conceived, for example, from a purely ontological viewpoint, as many popularisations of modern physics suggest, or from an epistemological point of view, as Bohr advocated, or from a logical viewpoint, as his pupil C F von Weizacker proposed. On the matter of expressing formally the content of Bohr's principle of complementarity, Einstein (1949 p674) complained that "despite much effort which I have expended on it, I have been unable to achieve the sharp formulation of Bohr's principle of complementarity." Notwithstanding these objections, I shall argue that Bohr's analysis of QT still remains the most profound and penetrating yet conceived. His View forms the basis of all varieties of the CI and remains more perceptive and radical than any subsequent reinterpretation of QT.

II

Feyerabend (1958) identifies two philosophical ideas by which Bohr is guided and upon which the validity of Bohr's approach entirely depends. The first idea is that "no content can be grasped without a form" (Bohr (1949 p240)) and in particular that "any experience makes its appearance within the frame of our customary points of view and forms of perception" (Bohr (1934 p1)). This idea is very much in the spirit of Kant (1787) who claimed that the possibility of principles of science which serve as presuppositions is based upon the use of a priori forms of intuition together with the categories of the understanding. Thus, while Bohr is often regarded as a positivist and physicists following the orthodox interpretation of QT often accept positivism, it follows from this idea that Bohr's view is different from positivism in an important respect. Positivism maintains that experience is ordered by means of laws but the experiences themselves have no order, they are unorganised simple elements. According to Bohr however, even our experiences are organised by "categories" (1949 p239) or "forms of perception" (1934 p1,8,17,22,94,etc) and they cannot exist without these forms. Bohr's second idea is that "however far the new phenomena transcend the scope of classical physical explanation, the account of all evidence must be expressed in classical terms" (1949 p209; 1934 p53,77,94,114; 1948 p313; 1934 p702). This, according to Feyerabend, implies that the "forms of perception" are, and always will be, those of classical physics. In this conclusion, Bohr seems to be retaining an important element of positivism. He writes (1934 p5): "We can by no means dispense with those forms of perception which colour our whole language and in terms of which all experience must ultimately be expressed." Thus there is an aspect of experience which is "given" and not capable of further analysis. Feyerabend (1958 p82) therefore refers to Bohr's point of view as a "positivism of a higher order".

We are now in a position to state in the form of a postulate one of the essential features of Bohr's interpretation of QT:

Postulate (a). The description of the classical apparatus and of the results of observation, which form part of the description of a quantum phenomenon, must he expressed in the concepts of classical physics (including those of everyday life), eliminating consistently the quantum of least action.

Definition. According to QT, the quantum of least action has a magnitude equal to one 'Planck' where 1 Planck = h = 6.625 x 10^{-27} erg.sec.

This postulate puts QT in an unusual logical relationship with classical physics. While QT is taken to be a more accurate theory for describing micro- phenomena than classical theory, the apparently only approximately valid classical concepts cannot be dispensed with when investigating micro-phenomena. Heisenberg (Syllabus of the Gifford Lectures, 1956 p8) writes, "... it would seem plausible that QT has disproved and replaced classical mechanics. This view, however, cannot be maintained if one remembers that the concepts of classical physics are necessary for the definition of any experiment relevant for QT." Bohm (1951 p625) expresses this point clearly when he writes: "QT presupposes the classical level and the general correctness of classical concepts in describing this level; it does not deduce classical concepts as limiting cases of quantum concepts." These conclusions follow directly from the writings of Bohr. For example, he writes (1949 p39), "... however far the phenomena transcend the scope of classical physical explanation, the account of all evidence must be expressed in classical terms. The argument is simply that by the word 'experiment' we refer to a situation where we can tell others what we have done and what we have learnt and that, therefore, the account of the experimental arrangement and of the results of the observations must be expressed in unambiguous language with suitable application of the terminology of classical physics" (See also 1934 p53,94; 1935; 1937 p293; 1948 p313; etc).

We have attempted above to follow Bohr's most general and most penetrating account of this argument in terms of Kantian epistemology. It should be noted that Bohr is not using QT as a proof of this philosophical view; he says, rather, that QT is an example of this view. (A more sophisticated approach is to use QT to support a philosophical view.)

Let us now take a closer look at this argument of Bohr's in order to point out what might seem to be a rather obvious loophole in the reasoning. The most striking feature of the argument is Bohr's contention that we can never go beyond the classical framework; that we can never avoid the use of classical concepts even in the description of micro-phenomena where classical concepts are strictly invalid (because of the existence of the quantum of action). Although we can accept Bohr's first philosophical idea, that a physical theory has no content without a form, the question is: why need this form be the form of classical physics? Bohr seems to be implying that the human mind is incapable of inventing a new and different conceptual scheme. Following Feyerabend's analysis (1962 p230), Bohr‘s premises leading to this conclusion are:

(i) We invent (or should use) only such ideas, concepts, theories, as are suggested by observation; "only by observation itself," writes Bohr (1934 p1), "do we come to recognise those laws which grant us a comprehensive view of the diversity of phenomena."

(ii) Because of the formation of appropriate habits any conceptual scheme employed for the explanation and prediction of facts will imprint itself upon our language, our experimental procedures, our expectations, as well as our experiences.

(iii) Classical physics is a universal conceptual scheme, ie, it is so general that no conceivable fact falls outside the domain of its application.

(iv) Classical physics has been used long enough for the formation of habits, referred to under (ii), to become operative.

The argument itself runs as follows: if classical physics is a universal theory (premise iii) and has been used long enough (premise iv), then all our experiences will be classical (premise ii) and we shall therefore be unable to conceive any concepts which fall outside the classical scheme (premise i). It does not seem unreasonable at this stage to ask such questions as: does quantum physics itself form a conceptual scheme distinct from classical physics employed for the explanation and prediction of facts (containing, for example, the concept of complementarity) and if so how long will it take for the formation of habits appropriate to this conceptual scheme? Was the appearance of QT (and even classical theory itself) impossible before a sufficiently long habitual use of classical concepts? Answers to these questions are outside the scope of this essay.

But consider Aristotelian physics. The great generality of the Aristotelian theory of motion would seem to make it a much stronger candidate for the argument than classical physics could ever be. It is, nevertheless, a fact that Aristotelian physics has been superseded by the physics of Galileo and Newton which involves a very different conceptual apparatus. "Clearly," writes Feyerabend (1962 p231), "this new conceptual apparatus was then not suggested by experience as interpreted in the Aristotelian manner and it was therefore a 'free mental creation'. This refutes (i). That (ii) needs modifying becomes clear when we consider that a scientist should always keep an open mind and that he should therefore always consider possible alternatives along with the theory he is favouring at a certain moment. If this demand is satisfied, then the habits cannot form, or at least they will not any longer completely determine the actions of the scientist. Furthermore, it cannot be admitted that the classical scheme is universally valid. It is not applicable to such phenomena as the behaviour of living organisms (which the Aristotelian scheme did cover), to personal consciousness, to the formation and behaviour of social groups, and to many other phenomena."

From this argument Feyerabend concludes that Bohr's arguments against the possibility of alternatives to the point of view of postulate (a), and hence of complementarity whose validity rests heavily on the validity of postulate (a) are all inconclusive. His argument does not question whether Bohr's View is a possible point of view but only whether Bohr's view is the only possible view. Feyerabend concludes that it is not; his argument consists of a refutation of the absolute validity of the second philosophical assumption made by Bohr, that the form of a physical theory must be the form of classical physics and everyday language.

Is it, however, really this easy to refute Bohr's clear contention that "it would be a misconception to believe that the difficulties of the atomic theory may be evaded by eventually replacing the concepts of classical physics by new conceptual forms" (1934 p16) and that there exist "general limits of man's capacity to create concepts" (1934 p96)? Bohr was well aware that, in the past, physical theories (eg

Aristotelian or Newtonian theories) had been superseded by new theories which brought with them new concepts. Bohr seems to be arguing that there is something intrinsically new and different about QT which makes its logical status with regard to previous theories totally different from any other theory. After all, "QT ... does not deduce classical concepts as limiting cases of quantum concepts" (Bohm, 1951 p625). If this is so then QT is set apart from older physical theories in this respect and therefore simple comparison with historical examples of theories which supersede previous theories might be inadequate as an argument to show that the unusual case of atomic theory can follow the same course and eventually replace classical concepts by some new and totally different concepts. Feyerabend's argument fails to take into account the unique attribute of QT which Bohr defines in what he calls the quantum postulate:

Postulate (b). Every quantum phenomenon has a feature of wholeness or individuality which never occurs in classical physics and which is symbolised by the quantum of least action, one Planck.

In order to understand what is meant by wholeness or individuality (indivisibility) of a quantum phenomenon, Bohr resolves a phenomenon into three elements: object, observing apparatus and the interaction between these two. He then asserts that these form an inseparable whole in a quantum phenomenon, but not in a classical phenomenon. This does not imply that further details of each of these three elements cannot he given in the case of a quantum phenomenon but only that the treatment of each element must pay significant regard to the others. In particular, although a formal quantum-mechanical treatment can be given of an isolated object, this is not something which could have any physical significance. Here the wholeness of a phenomenon comes into play, and demands that the corresponding apparatus should also be taken into account, for only by means of the apparatus can the properties of the object be meaningfully and quantitatively defined: "a measurement can mean nothing else than the unambiguous comparison of some property of the object under investigation with a corresponding property of another system, serving as a measuring instrument" (1939 p19; see also 1935 p699; 1937 p291). Thus, in accordance with postulate (a), reference is made to the object purely in classical terms since its properties are determined by being correlated with a property of the apparatus (which is, of course, described in classical terms). Prompted by the considerations of Einstein et al (1935) (see Ch 3), Bohr (1935) was led to clarify a point here which had been until then rather unclear in his writings: In a quantum phenomenon, measurement does not so much determine a property of the object as essentially define it in the first place, as far as possible. These considerations led d'Espagnat (1971, P395) to give an alternative version of postulate (b) in his analysis of Bohr's thesis. He writes:

Postulate (b'). "In general, quantum systems should not even be thought of as possessing individual properties, independently of the experimental arrangement."

(It is rather unfortunate that the very similar words 'individuality' in (b) and 'individual' in (b') have totally different meanings in the two contexts.)

We can now begin to see how this postulate offers further justification for the view that any descriptive conceptual scheme must remain restricted to the classical repertoire, postulate (a). Postulate (b) implies that the description of a quantum phenomenon must, in practice, be given solely in terms of a description of the experimental arrangement and of the experimental results. Bohr (1958) writes: "by the word experiment we can only mean a procedure regarding which we are able to communicate to others what we have done and what we have learnt." Description of experimental arrangement and results, moreover, must be given, at least for lack of an alternative, "in plain language, suitably refined by the usual physical terminology." Since this language is precisely the language of classical physics, what Bohr asserts is that ultimately it is in the language of classical physics that our statements about physical phenomena of any kind receive an unambiguous meaning. Thus, if, as Feyerabend argues, the classical forms of perception are not a priori and therefore can change, since postulate (b) implies that every observable quantum mechanical process must have a classical aspect, any newly created quantum forms of perception would necessitate a redescription of objects hitherto regarded as essentially classical. That is, if classical forms of perception are inadequate to describe quantum phenomena then since all quantum phenomena make their appearance through macroscopic instruments, Feyerabend's argument seems to become that classical forms of perception are inadequate to describe the classical level. This, however, is not the view of Bohr who accepts that classical concepts are adequate at the familiar classical level. Thus postulates (a) and (b) are not entirely independent, one supports the other in the claim that classical concepts are acceptable for the classical level and Feyerabend's critique of the finality of (a) is weakened by the supportive nature of (b).

It has been shown how postulate (a) is closely related to Kantian epistemology. The same is true of postulate (b). The world in itself (noumenon) is, according to Kant, unknowable; only the world as conceptualised by us in our perception (phenomenon) is knowable. He thus tied together conception and perception. In particular he argued that one can't distinguish between a phenomenon and its observation; this is the essential content of (b). Following Kent's lead, Bohr reasoned that the very circumstances which make a measurement possible also serve to define the concept measured. And the general limits of man's capacity to create concepts have their roots in our differentiation between subject and object; this differentiation is just what is impossible in QT. There is here the problem as to whether this subject- object distinction refers to an interaction between mind and world or between experimental apparatus and measured object only. This will be discussed in detail in Chapter Two.

Returning again to Bohr's arguments for the retention of classical concepts in the description of quantum phenomena, Hooker (1972 Sec 14) presents the following version of this argument:

A. A conceptual framework for unambiguous communicable knowledge is only possible where a sharp separation between subject and object is possible.

B. Physically, an effective separation between instrument and system is only possible at the classical macro-level where Planck's constant, h, is negligible.

C. Since instruments are the means used to gain knowledge of a system, the instrument-system relation is a subrelation of the subject-object relation which is taken in the sense of relating the human subject to the system.

Therefore,

D. Only concepts formulated under classical conditions will be suitable for the formulation of unambiguous communicable knowledge.

Notice that in this formulation postulate B is essentially the same as (b) above and is the crucial assertion in the argument, as opposed to Feyerabend's formulation where (b) is totally omitted from the argument. Is this presentation of the argument, then, immune from Feyerabend's objection? It does explain in a less dogmatic way why, in a discussion of QT, classical concepts should be regarded as indispensable. Hooker, however, does point out one reason why "one (should) not be too hasty in assuming the description of the micro-world must forever remain classical in nature". The reason is the recent discovery of macro-non-classical (ie quantum) effects such as superconductivity and coherent laser generation. Ferromagnetism is another macro-non-classical effect which has been familiar to men for centuries and has eluded classical description. Nevertheless, even if the argument is accepted that quantum phenomena can and must be described in classical terms, it would be most unwise to assume that this limitation must apply to all future phenomena. A priori arguments limiting the future of scientific development seldom, if ever, stand the test of time. In this respect we are in agreement with Feyerabend. Thus, for example, the absolute and eternal validity of postulate A may be questioned on the grounds that, in the future, communication of knowledge may take a wholly different form (cf music and art) , enabling the conceptual limitation inherent in A to be circumvented.

Continuing our analysis of Bohr, we now turn to the concept of interaction in QT. According to Bohr the interaction between object and measuring apparatus is 'finite', 'inevitable' , 'not negligible' and 'not separately accountable' and these conclusions are viewed as a consequence of postulates (a) and (b) above. He writes (1961 P60): "The element of wholeness, symbolised by the quantum of action and completely foreign to classical physical principles has, however, the consequence that in the study of quantum processes any experimental inquiry implies an interaction between the atomic object and the measuring tools which, although essential for the characterisation of the phenomena, evades a separate account if the experiment is to serve its purpose of yielding unambiguous answers to our questions," and further, "... the interaction between measuring instruments and the object under observation cannot be neglected or compensated for" in the way that it can be in classical physics. The argument which Bohr uses to show that the interaction cannot be neglected seems straightforward. He writes (1927 p54): "The quantum postulate [(b)] implies that any observation of atomic phenomena will involve an interaction with the agency of observation not to be neglected." In the sense that there is a lower bound to any quantum interaction, represented by the quantum of action (h), it is in principle impossible to neglect the interaction, although this does not of course exclude its being neglected in practice in the classical limit. In certain cases, however, the interaction cannot be neglected; then the only other possibility would be that of keeping a close watch on the interaction so that it may be allowed for. Postulate (a) alone cannot have any consequences which contradict classical physics: for its essential function is to establish classical physics as an indispensible part of quantum theory. Therefore, alone, it cannot imply the impossibility of neglecting or of allowing for the interaction, since both these are classically permissible. Only when postulate (a) and (b) are taken together can the impossibility of allowing for the interaction in a quantum phenomenon be demonstrated. The argument runs: Having shown that the interaction cannot be neglected, by using postulate (b), it follows from postulate (a) that it cannot be allowed for or controlled either since "the necessity of basing the description of the properties and manipulation of the measuring instruments on purely classical ideas implies the neglect of all quantum effects in that description, and in particular the renunciation of a control of the reaction of the object on the instruments more accurate than is compatible with the (Heisenberg uncertainty) relation."

In other words, since apparatus and results must eventually be described in classical terms and since this involves an unavoidable error (the quantum of action being neglected), it is impossible to follow with arbitrary accuracy the details of the interaction between the object and apparatus and therefore also impossible to allow for the effect of the apparatus on the observed object which inevitably accompanies every observation. There is a temptation at this stage to carry classical presuppositions into the quantum domain. For example, to speak of the existence of the quantum of action as providing an uncontrollable disturbance of the observed system or to claim that merely our knowledge of atomic systems is limited in the manner indicated by the uncertainty relations (as if there existed an information barrier in the world). Bohr rejected this point of view. For Bohr there is no atomic world, other than that of experimental experience. This is not a subjectivist position since the results of experiments are objective, but the world is knowable and describable only through experimental contexts. Heisenberg, on the other hand, continually uses 'disturbance language'. He also tends to positivism, unlike Bohr. Thus the CI which is founded principally on the writings of Bohr and Heisenberg is not a single doctrine. It is, in fact, usually closer to Heisenberg than Bohr.

It will be helpful, at this stage, to clarify a confusion which has arisen from the use of 'disturbance language' when discussing the results of certain thought-experiments. Following Heisenberg, a number of writers seem to imply that the uncertainty principle can be deduced from the results of these thought-experiments. Gamow (1958), for example, writes: "But in the atomic world we can never overlook the disturbance caused by the introduction of the measuring apparatus." He then proceeds to argue that, from the fact that high energy quanta can be used to determine accurately the position of an electron, say, and low energy quanta can be used to determine accurately the velocity of an electron, "Heisenberg showed that the combined uncertainty in position and in velocity (ie the product of the two uncertainties) can never be smaller than Planck's constant divided by the mass of the particle" where these uncertainties are caused in the following way: Low energy quanta have "long wavelength; they the disturb trajectory of the particle very little but only hazily indicate its path", while high energy quanta have "short wavelength; each quantum sharply locates the particle but completely disturbs its trajectory. Thus the trajectory can at the most be only roughly approximated." We are here, however, using macroscopic concepts in a situation where the actions involved are of the order of one Planck and therefore where these concepts are simply not applicable - see (a). Also, it is not clear from this account why such disturbances cannot be allowed for as is the case in corresponding classical situations. "The reason why such allowance cannot be made in QT," writes Hesse (1965 p267), "is that the uncertainty principle entails that the results of the thought-experiment are relatively uncertain, but is not entailed by them." The uncertainty principle is not proved by considering a few thought-experiments, it is only illustrated. The evidence for it is of two kinds. First, no one has yet found any experimental way to defeat the limitations in measurements which it implies (but see Note). Second, the laws of QT seem to require it if their consistency is to be maintained, and the predictions of these laws have been confirmed over and over again with great precision. The UP should be regarded as the quantitative logical backbone of the QT.

Returning to our analysis of Bohr we find that, as a consequence of the impossibility of neglecting or allowing for the interaction in a quantum phenomenon, the interaction must be an integral constituent of the phenomenon. "The fundamental difference with respect to the analysis of phenomena in classical and in quantum physics", writes Bohr (1954 p72), "is that in the former the interaction between the objects and the measuring instruments may be neglected or compensated for, while in the latter this interaction forms an integral part of the phenomena." The interaction is thus seen as an essential part of the phenomenon itself which is in accord with postulate (b) showing once again the wholeness and impossibility of arbitrary subdivision of a quantum phenomenon. A phenomenon is destroyed by a complete surveillance of the interaction. Had the influence on the interaction incurred by such surveillance been an influence on the object then the resulting destruction of the phenomenon would not be surprising. Here, however, the interaction which is normally thought of as lying outside the phenomenon itself can be seen to be an integral part of the phenomenon. Bohr writes: "The individuality of the typical quantum effects finds its proper expression in the circumstance that any attempt of subdividing the phenomena will demand a change in the experimental arrangement introducing new possibilities of interaction between objects and measuring instruments which in principle cannot be controlled" (1949 p40). And later: "The essential wholeness of a proper quantum phenomenon finds indeed logical expression in the circumstance that any attempt at its well-defined subdivision would require a change in the experimental arrangement incompatible with the appearance of the phenomenon itself" (1954 p72).

From the above view of the logical status of the interaction in a quantum phenomenon it becomes clear that the object can no longer be described as having any behaviour distinct from the interaction in a measurement. We are not "any longer in a position to speak of the autonomous behaviour of a physical object, due to the unavoidable interaction between the object and the measuring instruments which in principle cannot be taken into account, if these instruments according to their purpose shall allow the unambiguous use of the concepts necessary for the description of experience" (Bohr 1937 p293). More particularly, "the unavoidable interaction between the objects and the measuring instruments sets an absolute limit to the possibility of speaking of a behaviour of atomic objects which is independent of the means of observation" (1938 p25); "above all, we must realize that this interaction cannot be sharply separated from an undisturbed behaviour of the object, since the necessity of basing the description of the properties and manipulation of the measuring instruments on purely classical ideas implies ... the renunciation of a control of the reaction of the object on the instruments" (1939 p19).

Thus the classical ideal enabling one to describe unambiguously the behaviour and properties of objects independently from the instruments of observation must be forfeited in the description of quantum objects. Bohr (1937 p293) writes: "From the above considerations it should be clear that the whole situation in atomic physics deprives of all meaning such inherent attributes as the idealizations of classical physics would ascribe to the object." Because of the non-independence of a quantum object from the measuring instrument no unambiguous description of the object can be given without explicit reference to the measuring instrument being used. This implies a limited applicability of the concepts of classical physics; only those concepts which are relevant to the given experimental context can be assigned a definite meaning. Bohr writes (1948 p313): "In this situation, an inherent element of ambiguity is involved in assigning conventional physical attributes to atomic objects."

Some writers have detected an instrumentalist tendency in Bohr's analysis at this point (eg Hooker (1972) and Gardner (1972b p107)) since he seems to be indicating that a complete description is only possible at the macro-level. Bohr, however, can be given a totally different interpretation on this point of the origin and significance of 'ambiguity' in this matter. Jammer (1966) has pointed out that one of Bohr's teachers was a student of Kierkegaard and there was a "continuing Kierkegaardian tradition in the Copenhagen of Bohr's youth" (Hesse 197h p65). Kierkegaard's existentialism holds that an observer in a sense creates the external world, and further that an individual's decision makes the world as it is. Following this line of thought, one can understand a possible reason for Bohr's reluctance to formulate an atomic ontology as distinct from the epistemological consequences of his interpretation of QT as well as a possible reason for the appearance of an ambiguity in assigning conventional physical attributes to atomic objects. If an observation in some sense creates the object being observed then clearly one cannot discuss the existence of the object independently from a consideration of what is known about the object as a result of the observation. Ambiguity results from an attempt to do this in QT because classical concepts, which have to be used in the description of the apparatus, presuppose an independently existing objective world and yet the quantum world cannot be described as having well-defined existence independent of its observation (or at least of the means of its observation; ie a description of the particular experimental arrangement involved). This is not instrumentalism, since Bohr does attempt to give some meaningful interpretation of the quantum world although he supposes that this cannot properly be done without reference to the macro-world.

Although Bohr himself seems to avoid discussion of the ontological implications of QT and for the most part confines himself to epistemological considerations, it is important to be clear that just because he believes that the form of the content of a physical theory (in particular QT) must be the form of classical physics, this does not imply that the world is in any sense classical. "Classical forms are themselves idealizations" (1934). This is merely the way which we have to capture the object; that is, the concepts which we use to describe the familiar world must be classical, but the world itself need not be classical. Hooker (1972 sec 14), in his analysis of Bohr's ontological position, abstracts four opposing tendencies in Bohr's writing:

I) The realist position where both macro and micro ontologies are taken seriously;

II) The macro-phenomenalist position where the classical level is given a privileged status and the existence of a micro-ontology is denied;

III) The Kantian position where any immediate knowledge of the world (in itself) is denied; and

IV) The view that the discussion of ontology is improper.

Some other followers of the CI however, have made more definite claims about quantum ontology.

The von Neumann proof of the non-existence of hidden variables was believed to imply that the uncertainty relations impose not merely an epistemological but also an ontological prohibition on quantum objects having actual definite simultaneous positions and momenta either during a measurement or between measurements (but see more recent theorems weakening this; Bell, Fine, etc in Ch4, II). According to von Neumann, it is inconsistent with QT to assume that, before a measurement is made, a. quantum object is in some definite classical ontological state which is destroyed during a measurement. This is contrary to the positivist position in claiming that assertions can be made about what exists even when it is admitted to be beyond the possibility of anything that can be observed. While one should acknowledge that, contrary to III, Bohr implies that the world-in-itself can be described insofar as QT itself describes it in terms of superpositions of states and these superpositions are beyond the possibility of anything that can be observed, the relation between the point of view of Bohr and of von Neumann is by no means very close. For example, Bohr argues that the measuring device must be described in classical terms while it is essential for von Neumann's theory of measurement that both object and measuring device be described with the help of a wave function. The latter procedure leads to difficulties which do not arise in Bohr's treatment. Hence, when dealing with von Neumann's approach, we are not dealing with a refinement of Bohr's arguments - we are dealing with a conflicting approach (see Ch 2).

We should mention here an important aspect of any ontology which one may attempt to construct from Bohr's ideas. The crucial difference between a classical and a quantum object, according to Bohr, is not a question of size as is sometimes assumed (see eg Gardner 1972b p103). "Within the frame of classical physics," writes Bohr (1955 p89), "there is no difference in principle between the description of the measuring instruments and the objects under investigation." A distinction has to be made in classical physics between objects and measuring instruments both in the sense that they are distinguishable entities and in the sense that they play different roles in an experiment, but this distinction is not a fundamental one since a theoretical analysis of the experiment treats both entities with the methods of classical physics. The all important step taken by Bohr in his analysis of QT rests on his supposition that the theoretical treatment of the object is to be given in terms of quantum physics while that of the measuring instrument is to be given in terms of classical physics. This introduces a fundamental distinction between objects and measuring instruments which does not exist in classical physics. As a consequence of this distinction it appears impossible to reduce the whole of physics (including measuring instruments) to a single unified formalism. Once again we find Bohr's position to be in conflict with positivism which assumes that it is possible, at least in principle, to construct a unified formalism in terms of which all observable data may be expressed. Indeed Bohr's approach is contrary to the whole methodological tradition of looking for universalisable theories. According to Bohr, then, the decision whether an entity is to be regarded as quantum or classical in nature depends, not on the size of the entity, but on whether it plays the role of an object or a measuring instrument. (A number of times in his analysis of thought-experiments Bohr treats relatively large entities such as a 'double-aperture screen’ or a 'photon box' as essentially quantum mechanical systems since in that context they appear in the role of the object being observed.) Many followers of the CI argue that not only is there non-reducibility of macro to micro laws but also non-reducibility of macro to micro ontologies. There is no single universal ontology; instead there is a duality between an ordinary classical ontology and superpositions of states. This view is, however, closer to von Neumann than Bohr where von Neumann adopts a more epistemological position of a duality between consciousness and the world (see Ch 2).

With respect to physical ontology, Hooker (1972 sec2) maintains that Bohr rejects two assumptions which are basic to classical physical ontology:

a) Physical reality is divisible into conceptually distinguishable elements and all elements have equal ontological status.

b) All complex objects consist in definite structures of the fundamental elements which are their constituents.

Without these assumptions ontological reductionism is impossible. But Bohr himself was never this explicit. Perhaps the reason for this can be found from a remark by Jammer (1971; p207) who writes: "... Bohr, recognising the insufficiency of the phenomenalistic position, regarded the measuring instrument as being describable both classically and quantum mechanically. By concluding that the macrophysical object has objective existence and intrinsic properties in one set of circumstances (eg, when used for the purpose of measuring) and has properties relative to the observer in another set of circumstances, or, in other words, by extending complementarity on a new level to macrophysics, Bohr avoided committing himself either to idealism or to realism. Summarising, we may say that for Bohr the very issue between realism and positivism (or between realism and idealism) was a matter subject to complementarity. In any case, according to Bohr, the ontological presuppositions in quantum mechanics differed from those in relativity." In order to understand this remark let us continue with our analysis of the concept of complementarity.

So far we have abstracted two fundamental postulates of Bohr's interpretation of QT. The first of these postulates, (a), stated that all experimental outcomes are to be described from within the classical conceptual framework while the second, (b), stated that there exists a finite minimum quantum of action (one Planck) associated with all microprocesses which implies that a microsystem and macro-measuring instrument are indivisibly connected through the non-negligible interaction. From these two postulates we deduced that the measurements used to define a quantum phenomenon do not lead to a description of the object which is classically complete. The completeness demanded by classical physics can be imposed only at the cost of destroying the phenomenon concerned since the different experimental arrangements that are necessary belong to different phenomena. Hooker (1972 sec9, B2), in his analysis of Bohr, has taken this to be a separate postulate not derivable from the first two.

Postulate (c). The applicability of classical (and all other) concepts to a particular situation is dependent on the relevant conditions obtaining in that situation.

That it was Bohr's opinion that (c) should follow as a consequence of (a) and (b) can be seen from his analysis of the interaction in QT (see p15-19) together with the statement (1935 p699), "the renunciation in each experimental arrangement of the one or the other of two aspects of the description of physical phenomena, - the combination of which characterizes the method of classical physics, ... ~ depends essentially on the impossibility, in the field of QT, of accurately controlling the reaction of the object on the measuring instruments, ie, the transfer of momentum in case of position measurements, and the displacements in case of momentum measurements."

From postulates (a), (b) & (c), Hooker concludes:

Postulate (d). There exists an inherent limitation on the simultaneous use of classical concepts to the same situation under the same conditions.

The example which is most often given by Bohr of this limitation is of the simultaneous use of space-time coordination and dynamical conservation laws. He writes, (1959 p22; etc..): "Any phenomenon in which we are concerned with tracing a displacement of some atomic object in space and time necessitates the establishment of several coincidences between the object and the rigidly connected bodies and movable devices which, in serving as scales and clocks respectively, define the space-time frame of reference to which the phenomenon in question is referred. Just this situation implies, however, a renunciation of any sharp control of the amount of momentum or energy exchanged during each coincidence between the object and the separate bodies entering into the experimental arrangement. Inversely, every phenomenon in which we are essentially concerned with momentum and energy exchanges - and which therefore necessitates an experimental arrangement allowing at least two successive determinations of momentum and energy quantities - will, in principle, imply a renunciation of the control of any precise space-time coordination of the objects in the time intervals between these measurements."

The concept of complementarity was devised in order to describe such situations where there is a limitation on the simultaneous use of different classical concepts. We shall here assume complementarity to stand for a binary relationship between two contrasted phenomena. The use of complementarity as a relationship between phenomena is not entirely consistent with Bohr earlier writings where he often uses expressions in which complementarity is seen as a relationship between 'descriptions' or 'information' or 'experience' or ‘pictures’ or 'evidence' or 'concepts'. For example, Bohr writes (1927 p56): "In fact, here ... we are not dealing with contradictory but with complementary pictures of the phenomena, which only together offer a natural generalization of the classical mode of description." Or (1957 p291): "The apparently incompatible sorts of information about the behaviour of the object under examination which we get by different experimental arrangements can clearly not he brought into connection with each other in the usual way, but may, as equally essential for an exhaustive account of all experience, be regarded as 'complementary' to each other." And again (1938 p26): "Information regarding the behaviour of an atomic object obtained under definite experimental conditions may, however, according to a terminology often used in atomic physics, be adequately characterized as complementary to any information about the same object obtained in some other experimental arrangement excluding the fulfilment of the first conditions. Although such kinds of information cannot be combined into a single picture by means of ordinary concepts they represent indeed equally essential aspects of any knowledge of the object in question which cannot be obtained in this domain."

In his later writings, however, Bohr comes to consistently use the term phenomena to describe the elements in the relation of complementarity. For example (1962 p25): "The impossibility of combining phenomena observed under different experimental arrangements into a single classical picture implies that such apparently contradictory phenomena must he regarded as complementary in the sense that, taken together, they exhaust all (classically) well defined knowledge about the atomic objects.” The nearest Bohr came to a definition of complementarity was probably in 1929 when he declared that the quantum postulate "forces us to adopt a new mode of description designated as complementary in the sense that any given application of classical concepts precludes the simultaneous use of other classical concepts which in a different connection are equally necessary for the elucidation of phenomena." More recently, d'Espagnat (1971 p399) has given the following formulation:

Principle of Complementarity. The nonseparable whole constituted by the system and a definite instrument can be described by using a simplification of our language according to which some of the properties that the system and the instrument share with one another are conventionally attributed to the system. However, other properties, which our classical experience leads us to think of, cannot then be attributed to the system. They are said to be complementary to the first ones. They can also be attributed to a quantum system of a similar type as the one considered so far but this is only possible if that system builds up an indivisible whole with some new instrument which is appropriate for a measurement of the new quantities. This account of complementarity leads one to suppose that properties of a quantum system are the elements in the complementarity relation. In Bohr's view, microsystems need not have any intrinsic properties (that is properties that would not depend on the instruments) therefore d'Espagnat's account is not strictly in accord with Bohr. It is closer to the first introduction of the word complementarity into the physicists' vocabulary by Pauli who described precise position and simultaneous precise momentum as complementary (properties a quantum system).

We define a complementary relation thus:

Definition. A quantitative phenomenon is complementary to another if certainty of the theoretically expected observable value of one quantity necessarily implies uncertainty in the expected value of the other.

III

The uncertainty principle (UP) is the logical basis of QT and may be stated as follows for the case of position and momentum:

Uncertainty Principle. The simultaneous measurement of position, q, and momentum, p, (or certain other pairs of variables, called complementary variables) of a quantum object is subject to an inherent limitation that such that if Δq is the uncertainty of position, and Δp of momentum, then the product $\Delta q \Delta p$ is necessarily greater than a quantity of order one Planck.

Thus if q could be determined exactly, p would be wholly uncertain and vice versa. Bohr's most frequent example of complementarity, namely that between space-time localization and dynamical conservation laws, can readily be used to show the application of complementarity to the UP. In the non-relativistic case, time and energy can be ignored, The example above then becomes that between a spatial localization at some instant (determined from measuring instruments which create a fixed frame of reference in space and have purely classical description) and a determination of momentum (by applying the appropriate conservation law). The uncertainties Δq and Δp can then be formally allocated to the respective constituents in the above complementarity relation.

Scheibe (1975 P55) has, however, pointed out a difficulty as regards the significance of this allocation. He writes: "Bohr regarded the complementarity between space-time coordination and the dynamical conservation laws (which we have now converted into a complementarity between the determination of position and that of momentum) as being presented by two complementary phenomena, with one phenomenon involving a position determination and the other a momentum determination, depending on the experimental arrangements used, and the two determinations are mutually exclusive." Yet there is no doubt that the uncertainty relations must be interpretable in terms of one and the same phenomenon only."

Thus, while the inherent limitation stated in the UP is assumed to refer or apply to a single quantum phenomenon, the relationship of complementarity, according to the discussion above, holds between two mutually exclusive phenomena. The solution to this problem may be found in Bohr's interpretation of the UP. He warns that ambiguities and misunderstandings are likely to arise if one attempts to objectify in a classical way physical attributes. One must be particularly careful in the interpretation of UP. He writes (1937 p292): "It would in particular not be out of place in this connection to warn against a misunderstanding likely to arise when one tries to express the content of Heisenberg's well known indeterminacy relations ... by such a statement as: 'The position and momentum of a particle cannot be simultaneously measured with arbitrary accuracy'. According to such a formulation it would appear as though we had to do with some arbitrary renunciation of the measurement of either the one or the other of two well-defined attributes of the object, which would not preclude the possibility of a future theory taking both attributes into account on the lines of the classical physics."

Apparently, according to Bohr, the UP is saying something more than merely that it is impossible simultaneously to measure two complementary attributes of an object. He implies that the limitation has something to do with the nature of the object (or more correctly, phenomenon) itself rather than a limitation of the measuring instruments only. In further clarification of this Bohr says (1949 p40) it must here be remembered that even in the indeterminacy relation ... we are dealing with an implication of the formalism which defies unambiguous expression in words suited to describe classical physical pictures. Thus, a sentence like 'we cannot know both the momentum and the position of an atomic object' raises at once questions as to the (simultaneous) physical reality of two such attributes of the object, which can be answered only by referring to the conditions for the unambiguous use of space-time concepts, on the one hand, and dynamical conservation laws, on the other hand."

In this case the emphasis has been made on the limitation of the UP to knowledge, not simply to measurement. But Bohr wishes to reject as misleading both these interpretations of UP because, he claims, it is not merely our knowledge or measurements which are limited according to the UP but the very possibility of knowledge or measurement has limitations. We normally think of a measurement as determining a property of an object which the object already has even before the measurement was made, and similarly that knowledge is knowledge of something that already exists, but Bohr wishes to modify the usual concepts of measurement and knowledge in QT. In certain circumstances, he suggests, we should not even speak about measurement or knowledge. The crucial point for Bohr is not that, for example, position and momentum cannot be simultaneously measured or known but rather that the two classical concepts cannot be simultaneously defined. (This sounds like a vestige of operationalism, ie if they cannot be defined by instruments, they do not exist - but we do not have to accept this view.)

The UP gives a quantitative expression of this limitation. So we find Bohr stating that "these circumstances find quantitative expression in Heisenberg's indeterminacy relations which specify the reciprocal latitude for the fixation, in quantum mechanics, of kinematical and dynamical variables required for the definition of the state of a system in classical mechanics. In fact, the limited commutability of the symbols by which such variables are represented in the quantal formalism corresponds to the mutual exclusion of the experimental arrangements required for their unambiguous definition. In this context, we are of course not concerned with a restriction as to the accuracy of measurements, but with a limitation of the well- defined application of space-time concepts and dynamical conservation laws, entailed by the necessary distinction between measuring instruments and atomic objects" (1953 p5).

In order to emphasise that the limitation implied by the UP is not merely of knowledge but of definition, Bohr writes (1955 p699): "Indeed we have in each experimental arrangement suited for the study of proper quantum phenomena not merely to do with an ignorance of the value of certain physical quantities, but with the impossibility of defining these quantities in an unambiguous way." We may therefore interpret Bohr (who is not always consistent on this point) as saying that the UP applies to a single phenomenon in the sense that it specifies the extent to which, in any given phenomenon, the physical quantities may simultaneously be well defined. He writes (1957 p293): "the proper role of the indeterminacy relations consists in assuring quantitatively the logical compatibility of apparently contradictory laws which appear when we use two different experimental "arrangements, of which only one permits an unambiguous use of the concept of position, while only the other permits the application of the concept of momentum defined as it is, solely by the law of conservation."

If we are to take it, then, that it is not simply the particular experimental arrangement but the UP which forms the logical basis for the definition of a given quantity, once again Bohr's view can be seen to differ from that of positivism, for here we are not only concerned with what is observably 'given' since it is ultimately the theory and not observation which determines that. Thus, when seen as a logical criterion, the UP can be regarded as applying to one phenomenon. Complementarity should, however, still be regarded as a relation holding between two phenomena. On this point Bohr is not explicit, he writes (1927 p60): the content of UP "may be summarized in the statement that according to the QT a general reciprocal relation exists between the maximum sharpness of definition of the space-time and energy-momentum vectors associated with the individuals. This circumstance may be regarded as a simple symbolical expression for the complementary nature of the space-time description and the claims of causality. At the same time, however, the general character of this relation makes it possible to a certain extent to reconcile the conservation laws with the space-time coordination of observations, the idea of a coincidence of well-defined events in a space-time point being replaced by that of unsharply defined individuals within finite space-time regions."

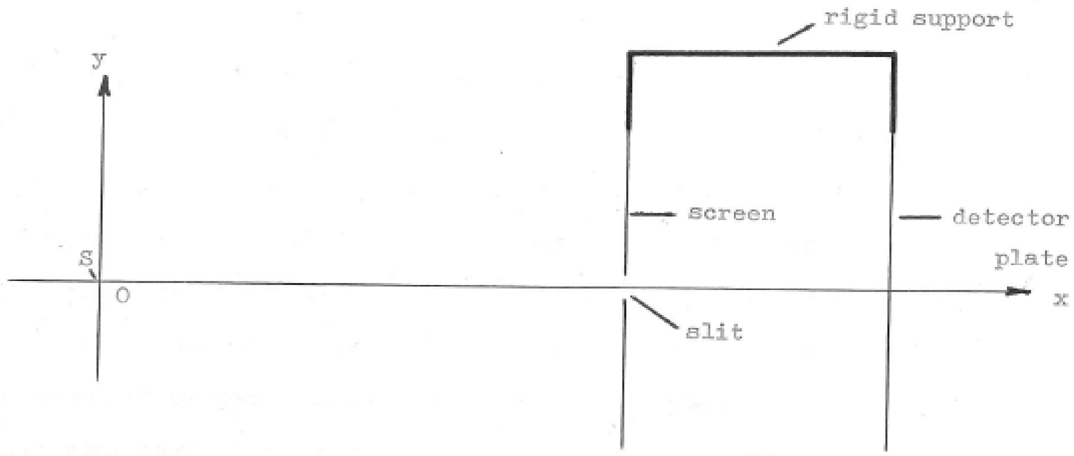
He regards the UP as a symbolical expression of complementarity which enables two complementary phenomena to be reconciled. It is important to stress the fact that complementarity and the UP are not synonymous. This follows from the recognition that the UP is an immediate mathematical consequence of the formalism of QT whereas complementarity is an extraneous interpretative addition to it. In fact, the quantum mechanical formalism with the inclusion of the Heisenberg UP can be, and has been, interpreted in a logically consistent way without any recourse to complementarity. Foch (1951) was therefore in error when he wrote: "At first the term complementarity signified that situation which arose directly from the uncertainty relations. Complementarity concerned the uncertainty in coordinate measurement and in the amount of motion ... and the term ‘principle of complementarity’ was understood as a synonym for the Heisenberg relations." Jammer (1966) moreover notes that Bohr had already conceived of the concept of complementarity and of its application to QT before Heisenberg formulated his UP.

Another aspect of Bohr's interpretation of the UP should now be analysed. This concerns the role of the interaction. Bohr considers that the UP "directly reflects the interaction between the system under observation and the tools of measurement" (1962 p91). Thus the statistical nature of the UP is not to be regarded as resulting from an inability of the measuring instruments to determine states which are in principle classically describable. Instead, the impossibility of surveying the interaction and hence the inherent limitation on the simultaneous use of classical concepts (postulate (d)) is the origin of the statistical content of the UP. The impossibility of surveying the interaction in measurement "clearly shows," says Bohr (1959 p19), "that the statistical character of the uncertainty relations in no way originates from any failure of measurements to discriminate within a certain latitude between classically describable states of the object, but rather expresses an essential limitation of the applicability of classical ideas to the analysis of quantum phenomena. The significance of the uncertainty relations is just to secure the absence, in such an analysis, of any contradiction between any imaginable measurements."

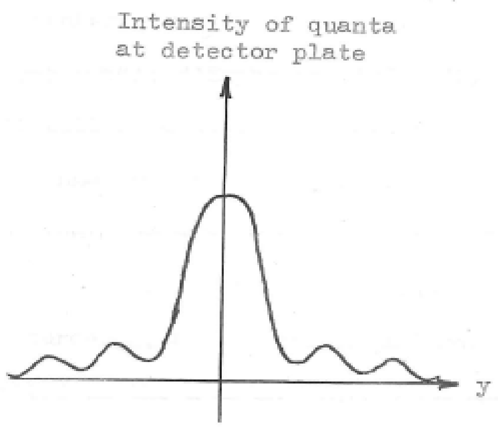
The UP is a consequence of the impossibility of surveying the interaction and once more is regarded as the logical foundation of the theory since it plays the role of avoiding contradictions. The UP "amounts to a proof of the consistency of the point of view of complementarity" (Feyerabend 1958 p96).

The most significant philosophical consequence of the UP together with the complementarity interpretation is the appearance of irreducible probabilities, ie the unpredictability of individual quantum phenomena. Heisenberg himself was the first to indicate this philosophical conclusion of the UP. Taking the law of causality in the strong (Laplacian) formulation as "the exact knowledge of the present allows the future to be calculated," Heisenberg (1927 p197) pointed out that "it is not the conclusion but the hypothesis that is false," because the UP implies that the exact initial values cannot be discovered which excludes the possibility of strict predictability of future events. "Since all experiments obey the quantum laws and, consequently, the indeterminacy relations, the incorrectness of the law of causality is a definitely established consequence of quantum mechanics itself," Heisenberg concluded. Prior to QT some philosophers seem to have claimed that the law of causality is indispensible for science. For example, Wittgenstein (1921 6.562) writes: "... what the law of causality is meant to exclude cannot even be described." Thus it would seem impossible, according to Wittgenstein, to describe anything which does not conform to the law of causality, ie classical determinism. It might appear at first sight that Wittgenstein is not in conflict with Bohr since micro-objects which, according to QT, do not obey the law of causality, cannot be described except in relation to some macroscopic measuring device. And Nagel (1961 p516) writes: "indeterminacy is not exhibited in any experimentally detectable behaviour of macroscopic objects. Indeed the theoretical indeterminism as calculated from quantum mechanics ... is far smaller than the limits of experimental accuracy" (see also de Broglie 1959 p250). This, however, is mistaken. While most macroscopic objects obey the law of causality to the degree that classical physics is causal, many do not. In particular, measuring instruments when measuring quantum quantities do not necessarily obey the law of causality. Wittgenstein's claim must therefore be rejected. This gives rise to irreducible

Figure 1



Single Slit Experimental Arrangement



Diffraction pattern

probability and thence the problem of measurement discussed in Chapter Two. Another who believed in the indispensible nature of the law of causality, although he did not think that it represented a necessary relationship between events, was Hume. He claimed (1759 Bk.I, Pt.III sec.3) that science is impossible unless one particular cause gives rise to one and only one particular effect. If Bohr is accepted, however, then Hume too was mistaken.

IV

In the history of physics it has usually been found that when a new fundamental principle (eg UP) is proposed its validity is established in terms of a conceptually simple thought-experiment. For a thought-experiment the experimenter is allowed an ideal workshop in which to mentally construct any kind of instrument, the only proviso being that its design and functioning do not contradict basic laws of physics. For example, the principle of conservation of momentum in classical mechanics made use of a thought- experiment involving the collision of two ideally elastic spheres in empty space while the principle of equivalence in general relativity was explained in terms of a laboratory in an elevator. In the case of the principle of complementarity a number of different thought-experiments have been devised including Heisenberg's γ-ray microscope and Einstein's photon box. We shall here discuss two related thought-experiments, the single and double slit experiments, which have been performed long ago for light and recently by Jönsson for electrons and which have been frequently discussed in the literature, see eg Heisenberg (1930 p25); Bohr (1955), (1949); Feyerabend (1962 p199); Hesse (1965 p265); d'Espagnat (1971 p10); Fine (1972 p4); Scheibe (1975 p42); as well as a number of textbooks on physics eg Bohm (1951), Messiah (1961 p142); Feynman & Hibbs (1965 p2).

Let us first consider the single slit experiment which involves the arrangement indicated in Fig.1. Electrons (or photons) each having equal momentum, p, leave a source, S, at the origin and move towards a screen which has a single slit located on the x-axis. Electrons passing through the slit then strike a detector plate (eg a photographic plate) leaving a dot at the point at which the electron is incident. The screen and detector are firmly secured to a rigid support. (The assumptions of the thought-experiment are non-relativistic.) If the slit width is sufficiently narrow and the detector a sufficient distance from the screen then the detector plate will show an interference pattern or diffraction pattern in which the width of the principal peak is considerably greater than that of the slit and which has very narrow interference fringes around this peak. This phenomenon is reminiscent of those encountered for classical waves in a medium, and indeed, with the aid of the de Broglie relation $p = h/\lambda$, one can show that the assumption that electrons of momentum p are classical waves of wavelength λ leads to a correct prediction of the proportions and structure of the observed pattern. However, although the assumption that electrons are waves is consistent with the above phenomenon, this assumption is not consistent with all the observable phenomena relating to electrons. In particular, the above experiment may be conducted with a source of very much reduced intensity in which case one observes that the pattern on the plate is built up of a number of discrete dots whose individual positions at any given time cannot be predicted but whose statistical distribution can be predicted from the Schrödinger wave equation and turns out to be a diffraction pattern. The dots cannot be explained (even qualitatively) in terms of classical wave theory which would predict for a low intensity source a diffraction pattern of correspondingly low intensity which gradually builds up into the final intense pattern. The dot pattern, however, can easily be explained in terms of classical particles. If one assumes that electrons are classical particles which leave the source, pass through the slit and arrive at the plate then the dot pattern is easily understood. But this particle model of electrons is not consistent with the final diffraction pattern.

In order to obtain a harmonic plane wave in the x direction and hence a diffraction pattern on the detector plate, the particle model would require particles to the left of the screen to have well-defined velocities in the x direction. In this case the particles would either be absorbed by the screen or else pass through the slit and arrive on the plate in a region exactly opposite the slit, giving neither interference fringes nor widening of the central peak. Neither is one able to explain the result in terms of interactions between large numbers of particles for, as we have seen, the experiment can be conducted at very low intensities such that, if the dots are interpreted as the arrival of single particles, there is only one particle in the apparatus at a time and the diffraction pattern is still obtained. Further, if particles are involved and one attempts to explain the interference pattern in terms of these particles then one would expect that the shape of the pattern should remain unaltered (only its size should change) if the position of the detector plate on the x-axis is altered. This is found not to be the case, in particular when the plate is moved close to the screen such that the distance between screen and plate is of the same order of magnitude as the slit width; at this position the interference pattern becomes totally changed. The simple particle model is unable to explain this non-equivalence of the patterns. Finally, one may argue that the interference pattern could conceivably be explained in terms of particles if one assumes that there is a statistical effect at the slit causing the pattern to be produced. We shall see, however, that when the case of two slits is considered the resulting pattern cannot be explained in this way without the introduction of further difficulties. Thus we have the situation where a particle model of electrons can (and indeed must for lack of an alternative) be used to account for certain of the observed phenomena associated with electrons but is unable to account for all observed phenomena while a similar circumstance holds for the wave model. Also it turns out that between them these two models are able to account for all observed phenomena associated with electrons (or photons etc). This very general property of microscopic objects of appearing under one of two contradictory aspects, that of waves or that of particles as the case may he, has come to be known as wave-particle duality.

We must now examine how, exactly, this duality appears as a result of the complementarity of space-time coordination and dynamical conservation in laws. Following the particle model, it is clear that at the moment at which the particle passes through the slit its position is known to lie in the range -y' to +y' where -y' and +y' are the y-coordinates of the lower and upper edges of the slit. Since the screen and detector are rigidly attached together and since the particle to the left of the screen has momentum p which is in the positive x direction, the single slit experiment shows that the particle reaches points on the detector plate which could not have been reached classically. According to classical theory particles should pass through the slit undeviated and arrive at the detector at a location directly opposite the slit. Since some change in the momentum of the particle must occur in order to explain how the particle could arrive at points not opposite the slit, one must assume that there has been an interaction between the particle and the screen. However, because the screen is rigidly secured it is impossible to determine accurately how much momentum is exchanged between screen and particle, that is, it is operationally impossible to apply the law of conservation of momentum. Bohr writes (1935 p697): "Let us first assume that, corresponding to usual experiments on the remarkable phenomena of electron diffraction, the diaphragm, like the other parts of the apparatus ... is rigidly fixed to a support which defines the space frame of reference. Then the momentum exchanged between the particle and the diaphragm will, together with the reaction of the particle on the other bodies, pass into this common support, and we have thus voluntarily cut ourselves off from any possibility of taking these reactions separately into account in predictions regarding the final result of the experiment, - say the position of the spot produced by the particle on the photographic plate."

In the particle model, the interaction between particle and screen, in so far as it can be understood at all, can be regarded as resulting from an exchange of momentum between screen and particle (with negligible energy exchange). This then explains why the particle momentum which was originally well defined on the left hand side of the screen becomes poorly defined on the right hand side; it is due to the interaction between screen and particle. According to Bohr, moreover, this interaction necessarily occurs in a quantum phenomenon, it cannot he surveyed, it is a necessary constituent of the phenomenon and its properties are a consequence of the classical description of measuring instruments and the individuality of quantum phenomena (postulates (a) & (b)). The reason why the particle's momentum, which is originally well defined, becomes undefined (or defined only to within certain limits) is because of the occurrence and unsurveyability of the interaction of the particle with the screen; the occurrence of the interaction is due to the wholeness of object-apparatus phenomenon, while the unsurveyability is due to the necessity for classical description of the phenomenon, together with the 'experimental fact' of the UP.

In order to make a rough quantitative calculation from the qualitative results of this experiment, let us make the usual (though largely arbitrary) assumption that the uncertainty in the y-component of momentum Δp_y is equal to the y-component of the momentum of a particle which arrives on the detector plate at the position of the first minimum in the diffraction pattern, ie

$$\Delta p_y = p \sin \alpha \qquad \dots 1$$

where p is the definite momentum before passage through the slit and α is the angle between the x-axis and the first diffraction minimum. This assumption has the advantage of relying entirely on the use of macroscopic experimental results. If the slit width d (= 2y') in the y direction is taken as the uncertainty in the position of a particle immediately after passage through the slit then:

$$\Delta y = d \qquad \dots 2$$

Taking the wave model which correctly predicts the overall diffraction pattern we may use the equation from diffraction theory,

$$\sin \alpha = \lambda/2d \qquad \dots 3$$

where λ is the wavelength of the wave which is related to the original momentum p through the de Broglie relation

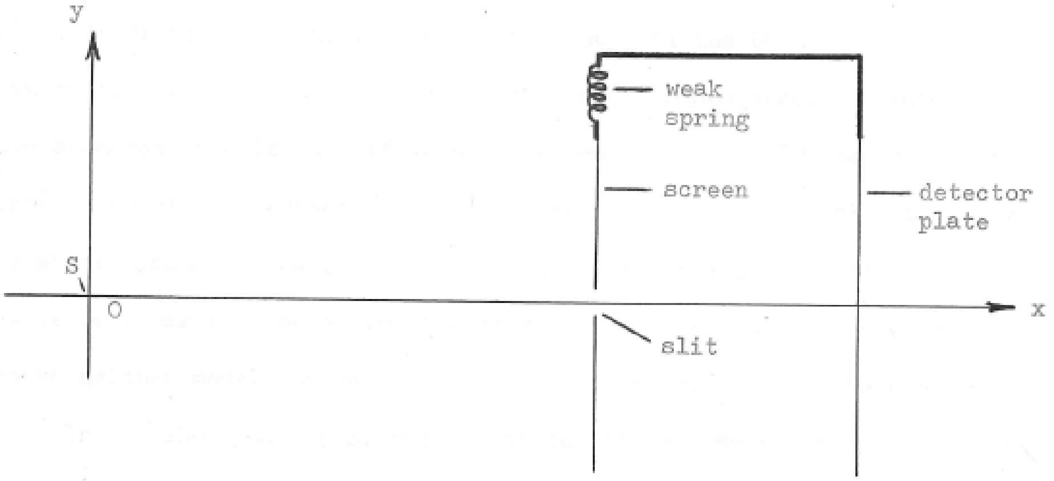
$$p = h/\lambda \qquad \dots 4$$

Combining the above four equations gives immediately

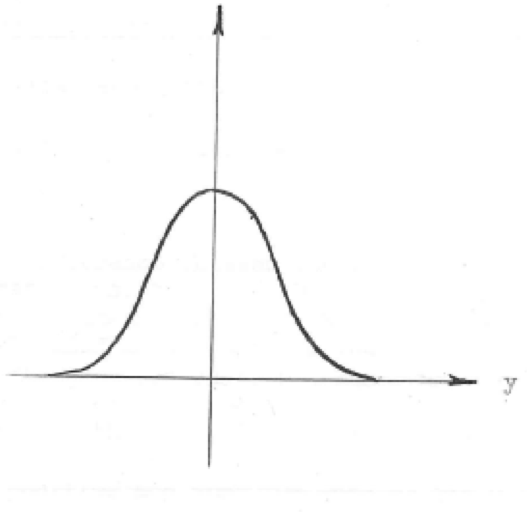
$$\Delta y . \Delta p_y = \tfrac{1}{2}h$$

and hence the UP since Δy and Δp_y are the minimum possible uncertainties in position and momentum in the y direction of an electron as it emerges from the slit. It should be emphasised that the UP has not been proved, nor is it logically implied by the results of the above experiment. It is only illustrated. It does, however, show how, if the UP is accepted, the results of the one slit experiment follow as a consequence. It might therefore have appeared more logical if we had proceeded from the UP and predicted as a consequence of this principle the observed duality between waves and particles. The above discussion has, however, given a justification for the use of either a wave or a particle model for the description of appropriate phenomena even though neither model can be used to describe or explain all the phenomena.

Figure 2



Single Slit Experimental Arrangement
(with movable screen)



Non-interference pattern

In the above analysis, the reason for the impossibility of surveillance of the interaction producing an uncertainty in the y component of momentum rested on the assumption that the screen remained strictly at rest during the flight of the particle through the slit. In this case the uncertainty relations for the screen have been neglected. It may, however, be objected that the momentum transferred between screen and particle could be discovered by making the screen movable in the y direction and observing the recoil of the screen as the particle passes through the slit - see Fig.2. In this way, by using the law of conservation of momentum and thus calculating the final to momentum; the particle, surveillance of the interaction may be possible and the absolute validity of the UP would be falsified. Concerning this objection, Bohr (1955 p698) points out that: "The principal difference between the two experimental arrangements under consideration is, however, that in the arrangement suited for the control of the momentum of the first diaphragm, this body can no longer be used as a measuring instrument for the same purpose as in the previous case, but must, as regards its position as regards the rest of the apparatus, be treated, like the particle traversing the slit, as an object of investigation, in the sense that the quantum-mechanical uncertainty relations regarding its position and momentum must be taken into account."

In other words, in a situation where the screen, which was previously in the role of a measuring instrument giving the position of the particle (y=0) together with uncertainty relating to this measurement, adopts the role of an object whose momentum has to be determined (in order to determine the momentum change of the particle), the uncertainty relations applying to this object cannot be neglected (because of the non-negligible, uncontrollable nature of the interaction between object and measuring instrument). Writing the accuracy with which the momentum of the screen in the y direction can be measured as δp_y, then it follows from the law of conservation of momentum that the minimum uncertainty of the particle after interaction Δp_y must be δp_y, ie

$$\delta p_y = \Delta p_y \qquad \ldots 5$$

Since the purpose of the new experimental arrangement is to obtain an accurate measurement of the y component of momentum for the particle, this implies that:

$$\Delta p_y << p \sin \alpha \qquad \ldots 6$$

Combining the relations 3, 4, 5 and 6 gives:

$$\delta p_y << h/2d \qquad \ldots 7$$

Taking 7 together with the uncertainty relations for the screen,

$$\delta y \, \delta p_y \geq \tfrac{1}{2}h \qquad \ldots 8$$

yields the final result that

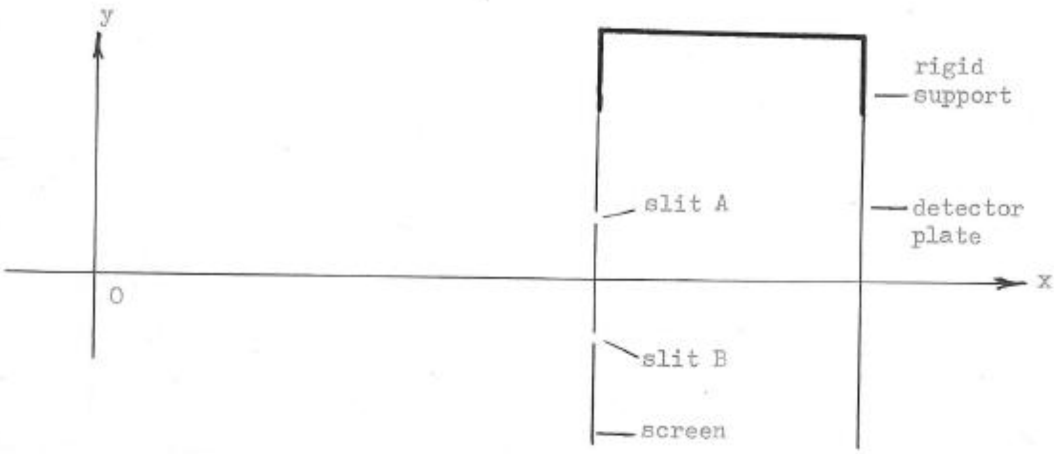
$$\delta y >> d \qquad \ldots 9$$

From relation 9 it can be seen that the screen in this new experimental arrangement can no longer be used as a fixed frame of reference with which to determine the particle position to the accuracy of 2, and further, the final pattern on the plate will no longer be a diffraction pattern since the condition for its production, namely a narrow slit of well-defined position on the y-axis acting as a diffracting aperture for waves, no longer holds. Thus the attempt to survey the momentum destroys the very phenomenon being investigated. The situation may be summarised as follows: If an arrangement is used in which the screen is rigidly connected with the frame of the local coordinate system (Fig 1) then the position of the electron can be determined to an accuracy of the width of the slit, whereas any information concerning the exact momentum exchange between electron and screen is forfeited because of the rigid connection between the screen and frame. If, on the other hand, the screen is suspended by weak springs (Fig 2), for example, then the momentum transfer may be determined, whereas any information concerning the exact position of the electron is forfeited owing to the indeterminate location of the screen. Similar results can be obtained for time and energy measurements as was shown by Bohr, for example in his analysis of Einstein's photon box thought-experiment at the Sixth Solvay Congress in 1950.

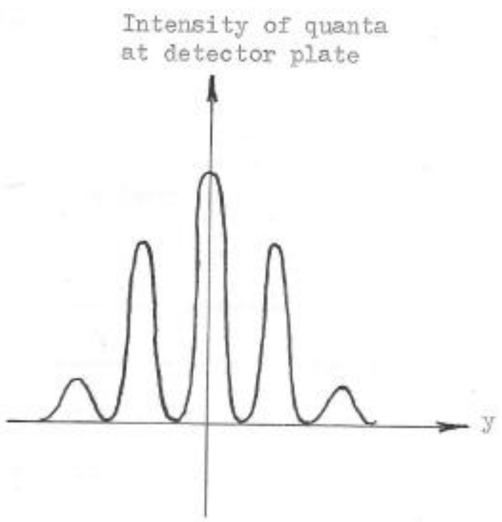
Bohr generalised these results by contending that descriptions in terms of space-time coordinates and descriptions in terms of energy-momentum transfers cannot both be operationally significant at the same time since they require mutually exclusive experimental arrangements. In the case of the above thought-experiment mutual exclusiveness of the experimental arrangements is ensured by the fact that the existence or absence of a rigid connection between the screen and the coordinating frame are logically contradictory to each other. The two experimental arrangements he called complementary for although mutually exclusive, they are jointly necessary (or complement one another) for an exhaustive description of the physical situation. From the above logically contradictory hypothesis it becomes clear why the simultaneous use of complementary descriptions lead to contradictions while a non-simultaneous use of such descriptions leads only to complementarity. Strictly speaking we should, perhaps, have confined the use of the term complementarity to that relationship holding between phenomena. The two complementary phenomena in the thought-experiment just considered are the diffraction pattern obtained in the situation where the screen is rigidly attached to the coordinating frame (Fig 1) and no attempt is made to determine the momentum transferred between screen and electron and the non-interference pattern obtained when the screen is made movable in the y direction in an attempt to determine the momentum transferred between screen and electron (Fig 2). There is, however, a natural tendency (to be found in Bohr and subsequent authorities) to transfer the term complementarity to the experimental arrangements and thence to the descriptions associated with these experimental arrangements which produce the complementary phenomena. Thus spatio-temporal and conservational descriptions are regarded as complementary. Finally, the term complementarity may be carried over to the parameters or variables themselves in terms of which complementary descriptions are formulated so that a position coordinate and a momentum variable are called complementary to each other. But this last use of the term is justified only if the variables are used in descriptions corresponding to complementary experimental arrangements.

We shall now turn to the double slit experiment which offers a particularly striking example of the complementarity between two phenomena. The only difference in the experimental arrangement between the single and double slit experiments is that in the double slit experiment the screen contains two slits rather than one (see Fig 3). The resulting interference pattern on the detector plate is again independent of the number of electrons which are present in the system (ie independent of the intensity of the source). This pattern can again be accounted for in terms of the wave model. Electrons leaving the source have constant momentum, p, and therefore (from the de Broglie relation $p = h/\lambda$) constant wavelength. Thus to the left of the screen the electrons appear as plane waves moving along the positive x direction. On arrival at the screen the wave passes through the two slits, A and B, which act as secondary sources of semi-circular waves. These waves expand and interfere coherently thus giving rise to an interference pattern on the detector plate. If, however, the intensity of the source is so weak that no more than one electron is on the way between the source and detector at any given time then each electron can be seen to leave one and only one localized dot on the detector, although after a sufficient time large numbers of these dots gradually build up into the interference pattern. The appearance of discrete dots can in no way be explained in terms of the wave model.

Figure 3



Double Slit Experimental Arrangement



Interference pattern

In order to explain the discrete dot structure on the detector it is natural to think in terms of the classical particle model. The screen is then regarded as an apparatus for determining the position of each incoming particle. Immediately after interaction with the screen each particle is assumed to lie in one or other of two quite separate localized regions just to the right of the two slits. But the experimental arrangement is unable to determine in which of these two regions the particle is located and hence through which slit the particle passed. Notice, however, that in this simple assumption (ie that each electron passes through one or other slit but not both) which is natural in the particle model (although not in the wave model), we have made an attempt to theoretically subdivide the phenomenon and have thus violated the assumption of the essential wholeness of a quantum phenomenon (postulate (b)). Bohr (1954 p72) maintains: "The essential wholeness of a proper quantum phenomenon finds indeed logical expression in the circumstance that any attempt at its well-defined subdivision would require a change in the experimental arrangement incompatible with the appearance of the phenomenon itself." We shall now see how, in accordance with Bohr's view, any attempt to determine through which slit the particle passes destroys the interference pattern, and thus the phenomenon under investigation.

The most drastic way of determining through which slit the electron passes is to close one of the holes, all dots appearing on the detector are then known to have been produced by electrons passing through the open slit. This slit may then be closed and the other opened. Again the resulting dots are known to have been caused by electrons passing through the open slit. If the assumption that each electron acts like a classical particle and passes through one slit or the other but not both is correct then one would expect that the final pattern obtained with only one slit open at a time would be exactly the same as the interference pattern obtained with both slits open. It turns out, however, that the pattern obtained when only slit A is open, the A-pattern, superimposed in this way on the B-pattern gives a pattern totally different from the interference pattern. We shall call this new pattern the additive pattern.

The paradoxical nature of this result, if electrons are regarded as particles, can be emphasised by the following consideration: a point X can be found on the detector plate which is such that in both the A-pattern and the B-pattern (and hence in the additive pattern) dots are found to appear at X while when both slits are opened no dots appear at X. Since each pattern can be produced by allowing only one electron to be in the apparatus at a time, the mutual interactions (if any would occur) between particles may be neglected. Now, if it is correct that each particle always possesses a well-defined trajectory, then the appearance of dots at X with only A open must be due to the fact that some particles arrived there along some path SAX. Assume that E is such a particle which is about to enter slit A. If at this moment slit B is opened then since no particle will arrive at X when both A and B are open the conditions are such that E will not arrive at X. Thus, if one wishes to retain the particle interpretation by assuming some statistical effect takes place at the slit (causing for example the interference pattern in the single slit experiment) one must introduce a new type of influence unknown to classical physics whereby the process of opening slit B must lead to a change in the path of E and hence to a change in the former statistical effect occurring at slit A. Further, this influence does not take place by means of some change propagated through the medium between A and B for, in so far as the wave interpretation is correct in accounting for the interference pattern, these effects must be instantaneous since this is the only assumption consistent with the wave interpretation. The reason for this is as follows: In the particle model the effect of opening B on slit A must be simultaneous because in the wave model the effect on the pattern at the detector is caused by the wave passing through B which travels at the same velocity as that through A, interferes with the wave through A and arrives on the detector at the same time as the wave through A. Thus the two waves producing the interference must pass through A and B at exactly the same instant t. Therefore, in the particle model the effect of influencing the path of the particle at A also occurs at time t, the time at which B is opened.

Also, this influence between A and B cannot be accounted for by assuming action at a distance since the conservation laws (which are also valid in QT) make no allowances for energies deriving from such actions. And this sort of explanation would have to explain why these actions do not work everywhere in space but only along surfaces which are surfaces of equal phase in the wave model. The wave model is, however, able to explain the origin of both the interference pattern (in terms of the coherent interference between the waves from slits A and B) and the additive pattern which results since no coherent interference between waves from A and waves from B can take place. But in attempting to explain the discrete dots appearing on the detector the wave model fails entirely. The wave model is also in conflict with the quantum postulate which demands that energy comes in indivisible units proportional to the frequency of the wave ν (E = h.ν) rather than being continuously distributed over the wavefront as is assumed in classical theory.

Thus we see again how a change in the experimental arrangement designed to subdivide the phenomenon and hence to survey the interaction leads to a destruction of the original phenomenon. Other attempts, besides the drastic one above, have been imagined to determine through which slit each electron passes but in each case it can be shown by means of the UP that this determination of the electron path necessarily destroys the interference pattern. Consider, for example, the case where the path of each electron is determined by means of a light source situated behind the screen and directed vertically such that, if an electron should pass through slit A, say, then light will he reflected from the electron. If the resolution of the instrument analysing the reflected light enables one to determine that the light was reflected from a region in the vicinity of slit A rather than B then it is possible to say that the electron has passed through slit A (see Feynman & Hibbs 1965 p7). According to Bohr such an arrangement must destroy the interference pattern if it is to determine through which slit each electron passes. One may object:

(i) If one uses a very weak source of light then it should be possible to determine through which slit each electron passes without significantly interfering with each electron. And a negligible interference cannot produce a large change in pattern on the detector plate. But weak light does not mean a weaker interaction between electron and light. Light is quantised and comes in photons of momentum p = h/λ where λ is its wavelength. Weak light just means fewer photons in which case we may miss seeing some of the electrons passing through A or B. The ones which we do see however will be scattered by the same amount (of order h/λ) as when using intense light. Those electrons which are observed form an additive pattern while those which are not form an interference pattern. The act of observation can therefore not be made to have a negligible effect in this way.

(ii) If light of longer wavelength is used then the effect of the light on the electron from which it is scattered (which is of order h/λ) can be indefinitely reduced. In this way it should be possible to make the interaction between light and electron negligible and therefore the effect on the final interference pattern negligible but still determine through which slit each electron has passed. However, if light of too long wavelength is used, we shall be unable to tell whether it was scattered from behind slit A or B; for the source of light of wavelength λ cannot be located in space with precision greater than order λ. lt turns out, in fact, that consistent application of the UP always yields the result that it is impossible to determine through which slit each electron passes without at same time destroying the interference pattern.

The above modification of the experimental arrangement is designed in order to survey the interaction between screen and the particle to the extent that the slit through which each particle passes can be determined. An alternative arrangement would be one similar to the modification described for the single slit experiment in which the screen is detached from its rigid support and suspended by a weak spring so that it becomes movable in the y direction. The change in the momentum of the screen (and hence of the particle by conservation of momentum) due to the interaction between screen and particle can then be determined. However, in this case, the screen which before modification of the experimental arrangement acted as a measuring instrument and was accordingly described in classical terms (postulate (a)) is now converted into the object and must therefore be described in quantum mechanical terms; in particular the UP must be applied explicitly to the screen. Therefore, if the final momentum of the screen after interaction is determined accurately then, from the UP, the positions of the slits become relatively uncertain and thus the position of the particle becomes uncertain. If the momentum of the particle is determined to sufficient accuracy and the point at which the particle arrives on the detector plate is known then the slit through which the particle passed can be determined. Scheibe (1975 p49), Feynman & Hibbs (1965 p10) etc have proved that, if the UP is accepted then it follows that the slit through which each electron passes cannot be determined in this way without at the same time destroying the interference pattern. Or, alternatively, the fact that the interference pattern is destroyed in situations where the slit through which each electron passes is determined is supportive evidence for the validity of the UP. Attempts have been made, most notably by Einstein (Solvay Congress 1930) and Popper (1955), to devise thought-experiments which manage to get round the limitation imposed by the UP but all these attempts have since been shown to fail. Double slit experiments have now been performed by Dempster & Bathe (1927); Faget & Fert (1957); Pfleegor & Mandel (1967) which confirm the theoretical prediction of the UP.

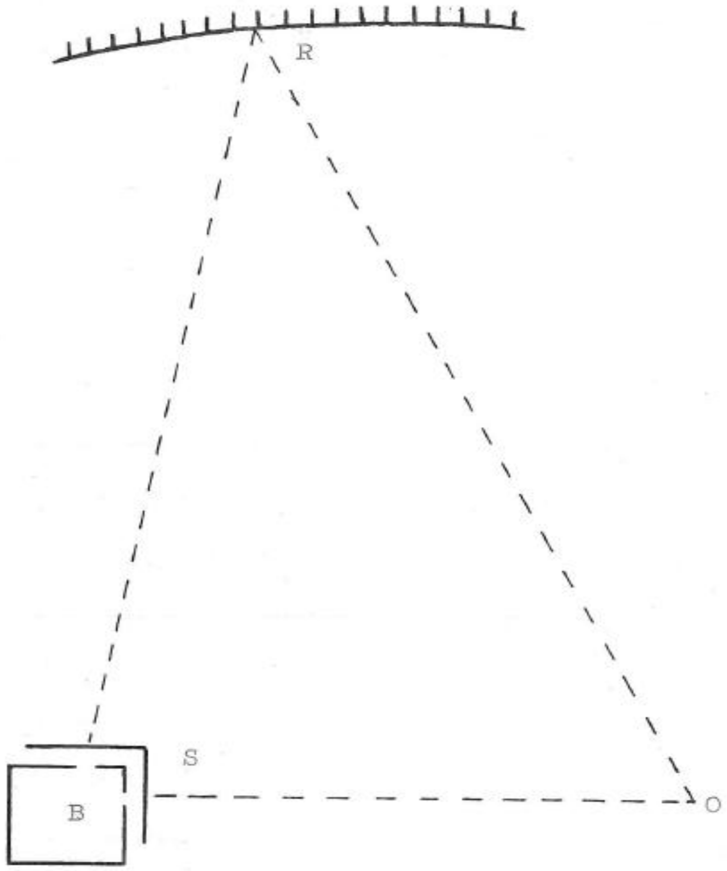
A more recent proposal on the simultaneous measurement of incompatible (complementary) observables has been made by Margenau (1950 p375) who suggests using two microscopes, one involving gamma rays and the other waves of long wavelength, one for locating the electron and the other for determining its momentum. Since Margenau adopts the statistical interpretation of QT which interprets the UP as a correlation between the spread in an ensemble of, say, position measurements with the spread in an ensemble of momentum measurements it nowhere suggests "with precision what may be expected in a single measurement of any kind." Thus, according to Margenau, there is no law in QT which basically prohibits a double measurement from succeeding. These ideas have been pursued since 1965 by a number of other writers but little agreement has been reached on basic issues of theory because, it seems, of diverging definitions of 'simultaneous measurements'. We shall discuss the statistical interpretation (SI) further in Chapters Three and Four. If, however, one assumes that QT is a theory which is able to describe individual entities, for example, electrons as is assumed in the CI, rather than ensembles of similarly prepared entities as is assumed in the SI then, according to Nagel (1961 p304-5): "It is a blunder to suppose ... that by improving our experimental techniques we may perhaps be able to ascertain the precise simultaneous values of the positions and momenta of electrons - in the senses of 'position' and 'momentum' fixed by current QT. Such a conjecture is on a par with the conjecture that by more intense study we may eventually be able to discover whether or not the ratio 2/3 is odd."

The conjecture that '2/5 is odd' is, of course, meaningless because the terms odd and even have not been defined for fractions. In exactly the same way, Nagel, following Bohr, argues that it is the theory which determines the meanings of theoretical terms, and the UP sets a limit not only to the precise simultaneous predictability of positions and momenta but also to their exact simultaneous definability. Thus he continues: "The supposition overlooks the crucial point that in virtue of the uncertainty relations the expression 'the precise simultaneous values of the position and momentum of an electron' has no defined sense in quantum mechanics."

An apparent difficulty relating to the CI of the UP concerns the apparent retrodictability of exact joint values of position and momentum variables of a micro-object. Bohr seems to have accepted this possibility but denied that it has any predictive significance. "Indeed," he said in his Como lecture in 1927, "the position of an individual at two given moments can be measured with any desired degree of accuracy; but if, from such a measurement, we would calculate the velocity of the individual in the ordinary way, it must he clearly realized that we are dealing with an abstraction, from which no unambiguous information can be obtained." Similarly, Heisenberg (1930 p20) accepts that "the uncertainty relation does not refer to the past." He gives the following example: suppose the velocity of an electron is accurately known and then its position is accurately measured at time t. Although this measurement will change the momentum of the electron in an unpredictable way, knowledge of the past momentum together with knowledge of position at time t enables one to calculate the position of the electron for times prior to t. For these past times both position and momentum of the electron will be known to an accuracy greater than the usual limit $\Delta p \Delta q$. Thus Heisenberg admits the possibility of retrodiction within the limits of the UP. He concludes (1950 p20), however, that this knowledge "can never (because of the unknown change in momentum caused by the position measurement) be used as an initial condition in any calculation of the future progress of the electron and thus cannot be subjected to experimental verification." He therefore says "it is a matter of personal belief whether such a calculation concerning the past history of the electron can be ascribed any physical reality or not." Following this Popper (1967 p25) argues that not only are precise retrodictions of complementary variables possible but also they are necessary for testing the theory. Clearly, if retrodictions are ascribed physical reality then this would have grave repercussions for, at least, the von Neumann version of the CI because, according to von Neumann, QT implies that micro-objects do not have simultaneous precise values of complementary variables. Therefore, according to the CI, retrodictions of this sort are not usually ascribed physical reality (see eg d'Espagnat (1971a p398)).

On this question of retrodictions Einstein adopts an interpretation of the UP which is even more radical than that of Bohr or Heisenberg or the CI in general. In a short paper entitled 'Knowledge of Past and Future in QM' Einstein, Tolman, and Podolsky (ETP) (1931 p780) argue that "the principles of QM actually involve an uncertainty in the description of past events which is analogous to the uncertainty in the prediction of future events." ETP consider a thought-experiment wherein a small box B, in Fig 4, contains identical particles in thermal agitation such that, when the shutter S is opened automatically for a short time, it can sometimes happen that a particle leaves the box through the top opening and travels over the path SRO through elastic reflection at the ellipsoidal reflector R and another particle leaves the box through the side opening and travels the direct path SO to an observer at O. Now, according to Heisenberg, the momentum and time of arrival of the particle on path SO can be determined at O and this will enable one to retrodict the time at which the shutter S opened (knowing the distance SO, the mass of the particle and hence its velocity). However, if the box B is accurately weighed before and after the shutter has opened then it is possible to determine the total energy of the particles which have left. Knowing the velocity of the first particle it is then possible to calculate its energy and hence the energy and velocity of the particle on path SRO. Having measured accurately the distances SRO and SO, and assuming SRO to be much greater than SO, it is then possible to accurately predict, from knowledge of the time at which the shutter opened, the time of arrival and the energy of the second particle. This result is in conflict with the energy-time UP as applied to the future.

Figure 4



ETP thought-experiment

In order to resolve this apparent paradox, ETP (1951 p781) argue that "the past motion of the first particle cannot be accurately determined as was assumed" by, for example, Heisenberg. They continue: "Indeed, we are forced to conclude that there can be no method for measuring the momentum of a particle without changing its value. For example, an analysis of the method of observing the Doppler effect in the reflected infrared light from an approaching particle shows that, although it permits a determination of the momentum of the particle both before and after collision with the light quantum used, it leaves an uncertainty as to the time at which the collision with the quantum takes place. Thus in our example, although the velocity of the first particle could be determined both before and after interaction with the infrared light, it would not be possible to determine the exact position along the path SO at which the change in velocity occurred as would be necessary to obtain the exact time at which the shutter was open." Thus ETP show that by assuming the possibility of retrodictions, similar to those assumed by Bohr and Heisenberg, it is possible to derive a contradiction with the usual predictive UP. The ETP argument thus avoids the difficulties for the CI of retrodiction.

V

As we have said (on p25) the nearest Bohr came to a definition of complementarity was when he declared that the quantum postulate (b) “forces us to adopt a new mode of description designated as complementary in the sense that any given application of classical concepts precludes the simultaneous use of other classical concepts which in a different connection are equally necessary for the elucidation of phenomena." According to this statement modes of description or descriptions are complementary. Of more direct relevance to thought-experiments such as the double slit he writes (1958): "Indeed, every experimental arrangement permitting the registration of an atomic particle in a limited space-time domain demands fixed measuring rods and synchronized clocks which, from their very definition, exclude the control of momentum and energy transmitted to them. Conversely, any unambiguous application of the dynamical conservation laws in quantum physics requires that the description of the phenomena involve a renunciation in principle of detailed space-time coordination."

From considerations such as this Bohr argued that space-time and conservational descriptions cannot both be operationally significant at the same time since they require mutually exclusive experimental arrangements. In the case of the double slit the mutually exclusive experimental arrangements are, for example, that in which the screen is rigidly attached to the support and that in which it is suspended by a weak spring. In the first arrangement a description of the final interference pattern can he given in terms of classical wave concepts the use of which precludes the simultaneous use of classical particle concepts that are equally necessary in the second arrangement to explain the final additive pattern which suggests that each electron passes through one slit or the other but not both.

Heisenberg agreed with Bohr to the extent that any interpretation of an experimental situation should make use of the concepts of classical physics but whereas Heisenberg was satisfied that either particle language or wave language could be used subject to the limitations imposed by the UP, Bohr insisted on the use of both. For Heisenberg the point of departure was the UP which he regarded as imposing a limitation on the applicability of classical concepts, such as position or momentum, to quantum objects. For Bohr, however, the UP was an indication, not of the inapplicability of the particle language or wave language, but rather of the impossibility of using both modes of expression simultaneously although only their combined use gives an exhaustive description of physical phenomena. For Bohr the starting point was wave-particle duality. He argued that any derivation of UP from thought-experiments must have recourse to the Einstein-de Broglie relations $E = h.\nu$ or $p = h/\lambda$. These relations relate wave and particle attributes and thus express wave-particle duality. The UP indicated to him that the classical conception of explanation rather than the classical concepts has to be revised. Thus he could write (1934 p96): "We must, in general, be prepared to accept the fact that a complete elucidation of one and the same object may require diverse points of view which defy a unique description." In considering the possibility of a breakdown of the classical ideal of explanation, Bohr can once more be seen to be in sympathy with the philosophy of Kierkegaard (see p.20). Denying grades of perfection, Kierkegaard claimed that perfect cognition is impossible. This being so, and since perfect cognition is at the basis of the classical ideal of explanation, a breakdown of the classical ideal must occur.

Although the UP can be derived from the general formalism of QT and an independent demonstration of the UP can be given by analysis of a thought-experiment together with the Einstein-de Broglie relations, for Bohr this was not evidence that the UP should be regarded as the starting point for an explanation of QT. He regarded it rather as a proof of the consistency of his own complementarity interpretation of microphysics with the mathematical formulation of QT. Wave-particle dualism and hence ultimately the use of different models (or pictures) for the complementary modes of description was, for Bohr, the reason for indeterminacy in quantum physics. "The measurement of the positional coordinates of a particle," he said in his Como lecture, "is accompanied not only by a finite change in the dynamical variables, but also the fixation of its position means a complete rupture in the causal description of its dynamical behaviour, while the determination of its momentum always implies a gap in the knowledge of its spatial propagation. Just this situation brings out most strikingly the complementary character of the description of atomic phenomena." This unavoidable rupture of description or the 'switching over from one mode of description to its complementary mode' (what Heisenberg later called the 'reduction of the wave packet'), is, according to Bohr, the origin of indeterminacy in QT. However, Bohr is not consistent in his explanation of the origin of indeterminacy for on other occasions he has explained the lack of determinacy in QT as resulting from an operational physical feature, namely the role of the interaction in a measurement. He writes (1934 p93) that it is the "interference with the course of the phenomena (due to the uncontrollable, non-negligible interaction between object and instrument of observation), which is of such a nature that it deprives us of the foundation underlying the causal mode of description." We shall discuss the problem of the reduction of the wave packet in Chapter Two, for the moment we shall continue with the discussion of duality.

We have shown how the results of the experiments with slits can be explained by means of the complementary use of classical particle and classical wave modes of description. There is, however, a radical difference between this use of the concept of complementarity and that discussed previously as a relation between phenomena. The phenomena are defined by Bohr as the sum total of the "observations obtained in specific circumstances, including an account of the whole experiment," (1948 p317; 1949 p257). In case of the slit experiments a phenomenon is the pattern obtained on the detector plate as well as an account of the experimental arrangement. At no stage, however, is a particle or a wave observed directly, their existence is only inferred because of the observation of phenomena which are interpreted to be the result of particle-like or wave-like behaviour. The particle and wave models are contrived to explain the origin of the phenomena, ie to explain what happens between phenomena. Reichenbach (1944 p17) calls these waves and particles which, for a given phenomenon, are unobservable in principle (since to observe them would involve additional experimental apparatus thus changing the experimental arrangement and hence the phenomena) and which are assumed to exist between the observable phenomena, interphenomena. Complementarity between waves and particles is here thus a complementarity between interphenomena rather than phenomena.

There is, of course, a close connection between phenomena and interphenomena. The interphenomenon in any given phenomenon is taken to be that object or entity which in classical physics is able to explain the origin of the phenomenon. For example, in the double slit experiment the interphenomenon involved in the situation where an interference pattern forms part of the phenomenon is taken to be a wave since classically a wave is able to explain the origin of the interference pattern. Now in Bohr's analysis of the complementarity between phenomena, this complementarity arises as a result of the nature of interaction between an object and the apparatus used to observe the object. The object in the double slit experiment is the interphenomenon, conceptualized in terms of waves or particles as the situation requires. Thus the phenomenon presupposes an interphenomenon which acts as the object of observation. We have seen above how, for example, when dots appeared on the detector plate, or when a determination of the slit through which an electron passes was made it was appropriate to regard the interphenomenon as being made up of particles. However, we have also seen how the attempt to give a unique description of the interphenomenon which is applicable in all experimental circumstances has failed; waves, while useful as a description of the interphenomenon in one set of circumstances, are inappropriate in other circumstances and similarly for particles. The reason for this failure is to be found from postulate (d) above which states that as a result of the inevitable use of classical concepts to describe quantum phenomena (postulate (a)) and of the inherent wholeness of a quantum phenomenon (postulate (b)), there exists an inherent limitation on the simultaneous use of classical concepts to the same situation under the same conditions. This limitation is not present in the case of particle or wave interpretations of the interphenomenon because:

1. Particle and wave descriptions are entirely classical and do not therefore exhibit those features peculiar to quantum descriptions of phenomena.

2. The particle and wave descriptions are intended to describe what happens between interactions while the peculiarities of QT appear precisely because of the difficulties involved in the description during interaction.

As a consequence of postulate (d) there must therefore be a limitation to the use of particle or wave descriptions in a quantum phenomenon. One demonstration of this limitation has been given by Reichenbach (1948 p337) who shows that in any given phenomenon, interpretation of the interphenomenon in terms of either waves or particles involves what he calls a 'causal anomaly'. He writes: "If there exists a causality 'behind' the observables of quantum mechanics, it is not of the normal type, but contradicts the principle of action by contact." For example, an attempt to explain the interference pattern in terms of particles involves a causal anomaly at the slits because of the instantaneous effect on particles passing through one slit produced by opening or closing the other. Similarly, an attempt to explain the results in terms of waves produces a causal anomaly at the detector plate when the wave front collapses instantaneously to produce a dot at some unpredictable point on the plate. This violation of normal causality, or causal anomaly in the "world of interphenomena" is raised by Reichenbach to the status of a principle, the "principle of anomaly", according to which no causal description of the interphenomena in QT can be given. For this reason Reichenbach seems to favour "restrictive interpretations" of QT which "restrict the assertion of quantum mechanics to statements about phenomena." One is then unable to say anything about the interphenomena, for example whether an electron goes through one slit or other or both in the double slit experiment.

Strictly speaking, however, the restrictive interpretation is not really an interpretation at all, because it asserts that one can only speak about phenomena while nothing can be said about interphenomena, or events in nature which come between the observable phenomena. This view seems to have been supported by Heisenberg (1958 p52) who maintains that the double slit experiment "shows clearly that the concept of the probability function does not allow a description of what happens between two observations. Any attempt to find such a description would lead to contradictions." But the restrictive interpretation implies that no model can be found to describe the intervening events between observations (other than the 'mathematical model' resulting from the formalism of QT which enables one to describe the relation between initial and final observable states of an interacting system). The wave and particle pictures of interphenomena are then regarded as partial representations of the mathematics on those occasions when they appear to apply, not as descriptions of actually existing entities (see also Hooker (1973)). According to the CI, however, although no single model can be found which is able to account for the functioning of microscopic entities in all circumstances, wave and particle models can be used in a complementary way to account for all quantum phenomena. Bohr (1958 p25) reasoned that "the unavoidable interaction between the objects and the measuring instruments sets an absolute limit to the possibility of speaking of a behaviour of atomic objects which is independent of the means of observation." Accordingly (1958 p40), it is impossible to make "any sharp separation between the behaviour of atomic objects and the interaction with the measuring instruments." Many writers have concluded that for Bohr and the CI in general, "an atom has no properties at all when it is not observed" - see Bohm (1957 p92); Gardner (1972b p91). In order to be able to assign properties to a microscopic object, so the argument goes, one must incorporate the object in an appropriate experimental arrangement which "serves to define" those properties of the object which are observables in this arrangement. It is, of course, natural to ask at this stage: What are electrons or photons then, according to the CI? Or more generally: What are quanta? This is essentially a request for a model in terms of which one can visualise a quantum. According to some writers, quanta are to be regarded as particles while according to others, quanta are to be regarded as waves, but this is not the view of Bohr. According to Bohr QT deals not with the properties of microscopic objects as classically modelled but rather with nothing more than the relationships among the observable large-scale phenomena. In certain experimental situations the observed phenomena are consistent with the assumption that quanta are particles in the classical sense while in other situations the observed phenomena are consistent with the view that quanta are waves in the classical sense, but according to this interpretation there is no model which can be found which is able to describe consistently the nature of quanta in all possible experimental circumstances.

One reason for Bohr's reluctance to search for a single generalised classical model for micro-objects is that he regarded phenomena as indivisible wholes, which it would be wrong to analyse, even abstractly and conceptually, as made up of parts consisting of various micro-objects; this is Bohr's demonstration that the laws of QT permit one consistently to renounce the notion of unique and precisely defined conceptual models in favour of that of complementary pairs of imprecisely defined models. Further, his assumption that the basic properties of matter can never be explained in terms of single unambiguous models implies that complementary pairs of imprecisely defined concepts will be necessary for the treatment of every physical domain ever to be investigated. Thus, according to Bohr, the limitations expressed in the principle of complementarity are final. It should be stated, however, that single generalised models have sometimes been found useful in the development of QT in seeming opposition to Bohr's contention that this would give rise to contradictions. In developing a model for identical particles in QT Schrödinger argued that with each FRS the nominee is given a medal; each medal is exactly the same as any other but they are not identical in the way that the titles 'FRS' are. It is the title 'FPS' which Schrödinger considers as the model for the identity of particles in QT rather than the medal which fails as a model (see also Hesse 1966 p50).

In justification of complementarity between models it should be remembered that the words 'wave' and 'particle' (as well as the words 'position’ and 'momentum') are borrowings from classical physics. In the new context of QT they must be understood in terms of the restrictions placed on their use by the principles of QT and not in their senses established in classical physics. Since the rules governing their use in the two contexts are different, what they mean in QT cannot be the same as their original meaning. Further, according to Nagel (1961 p144): "The fact that a visualizable model embodying the laws of classical physics cannot be given for QT ... is not an adequate ground for denying that the QT does formulate the structural properties of subatomic processes. It is doubtless desirable to have a satisfactory model for the theory. But the type of model that is regarded as satisfactory at any given time is a function of the prevailing intellectual climate." This can clearly be seen in the conflicting views of Duhem and Kelvin on the desirability and necessity of finding a mechanical model for physical theories. Duhem regards the desire to find such a model as a peculiarly English characteristic. But Nagel continues: "Even though current models for QT may strike us as strange and even 'unintelligible', there are no compelling reasons for assuming that the strangeness will not wear away with increasing familiarity, or that a more satisfactory interpretation for the theory will not eventually found." The 'unintelligible' nature of the complementarity between models is, in fact, largely due to a failure to note that the words 'wave' and 'particle' are being employed analogically. Even in Newtonian physics where the particle or 'corpuscular' model is sometimes regarded as a factual description it should be remembered that use of the concept of a particle is still in the form of an analogy because certain features of all commonly observed entities (eg billiard balls) which are taken to be like atomic particles in many respects other than size are believed to be missing in actual atomic particles (eg colour). Those respects in which for example billiard balls may be said to be like classical atoms are called by Hesse (1966) the positive analogy while those which are not like classical atoms are called the negative analogy. Since, following Hesse (1965 p27), " a model is intended as a factual description if it exhibits a positive analogy and no negative analogy in all respects hitherto tested, and if it has surplus content which is in principle capable of test," it is clear that even classical use of the word 'particle' is really as a model rather than a factual description since all observed entities which could be regarded as being like particles exhibit negative analogies to the way a classical particle is usually conceived (ie indestructible, frictionless, hard, etc). Thus, already in classical physics particles and waves should not be taken as factual descriptions of microscopic objects.

The surprising thing in QT is that when the language of 'waves' and 'particles' is understood in terms of the way these words are actually used in the quantum mechanical characterisation of, for example, electrons, the wave and particle models both succeed and both fail in a way totally unfamiliar to classical physics. In certain circumstances classical particles exhibit some positive analogy with electrons while in different circumstances classical waves exhibit some positive analogy with electrons. Since much of the positive analogy of the particle model is the negative analogy of the wave model and vice versa, the two models appear to be contradictory. However, in the context when each of these models is found to be applicable in QT no contradiction actually arises. The experimental contexts in which the particle model is applicable and those in which the wave model is applicable are mutually exclusive. Complementary use of both models can therefore made without fear of direct contradiction. However, because direct contradiction would arise if the positive analogies of both models were regarded as factual descriptions of electrons, Bohr maintains that neither model should be taken as explanations of underlying mechanisms of phenomena but rather they should be taken as different (complementary) aspects of observable events. Since both these models, in their appropriate contexts, have proved, and continue to prove useful in predicting natural expectations in new, and heretofore unexplored, domains they should not be regarded as dispensable simply because of their strange and unfamiliar dual usage. Apart from the question of whether they are essential to the understanding of QT as a whole, they are both without doubt of essential importance in physical research where they can be generalised, extended and tested in a way in which the purely formal deductive mathematics cannot.

Because of the unfamiliar stance of wave-particle duality and partly because of the tendency to regard models as factual descriptions, it is often asked why a totally new model cannot be found for microscopic entities which is able to dispense with duality and enable electrons, for example, to be pictured in a more realistic fashion. The honest answer would seem to be that there is, in fact, no a priori reason why this cannot be done and indeed such an attempt has been made by Bohm (1957

Ch4) which will be discussed in Chapter Four. However, the request for a totally new model does not appreciate the venerable character of the ontologies associated with particles and waves - with the atom (in its generic sense) and with the plenum. Much of the debate concerning the foundations of Western science has taken the form of a debate between the proponents of these two great ontologies, for example the debate between Democritean atomists and Parmenidean monists, or that between Newton and Descartes-Huygens, or that between Dalton and Faraday-Maxwell, or the shorter debate between Einstein and believers in ether. Thus the atomic and plenum ontologies cannot easily be dispensed with and it is even less easy to find some totally new alternative (in fact Bohm's 'new' model is really founded on the atomic ontology). In order to demonstrate how deeply ingrained the atomic and plenum ontologies are in classical physics, Hooker (1975) has analysed the mathematical structures of classical particle mechanics (CPM) and of classical field theory (CFT). From this analysis Hooker (1975 p203) concludes that "the set-theoretic structure of CPM ... is mirrored exactly by the specific kind of individuality of the objects in the classical (atomic) ontology," and (1975 p236) that "the peculiarities of the plenum ontology find a direct reflection in the mathematical structure of CFT." Thus, he continues, "there is the most intimate connection between the view of the world the theory expresses, the fundamental conceptual apparatus required to express physical descriptions from that point of view and the mathematical structure in which the view finds expression as a theory in exact science."

Since, according to Hooker (1973 p209), "quantum mechanics is a fusion of the characteristic mathematical structures of a field theory and a particle theory," and since there is clearly a radical difference between the plenum and atomic ontologies (exhibited, for example, by the ‘maximal contrast' between the two ontologies in their account of the relationship between fundamental and reducible properties, their analysis of the reality of change, discontinuity, creation and destruction, etc.), Hooker (1975 p260) concludes that the complementarity approach which maintains a duality between waves and particles (this duality being exactly one of field (plenum) and particle (atom) conceptual structures) is therefore "doomed at the outset to be no more than a makeshift, more or less ad hoc, adjustment pressing various fragments of the particle and field conceptual schemes into service blindly as the interpretational occasion demands." Thus Hooker argues for the view that "no consistent, plausible, realist interpretation of QM is possible within our present conceptual resources," and therefore that "QM demands either a new conceptual-ontological scheme ... or the abandonment of QM as a hopelessly bastard offspring of an attempted marriage of the two great classical theoretical structures, doomed forever to a gerrymandered interpretation in terms of one of them." (1973 p270) Bohr, however, was well aware of the ‘depth of the cleavage’ between the wave and particle conceptual schemes and also of the intimate mathematical connection between wave and particle concepts in QT for he regarded the Einstein-de Broglie relations as the mathematical foundations of the concept of wave-particle duality, and these relations go deeply into the whole mathematical structure of QT. Clearly, however, for Bohr the fact that "ordinary physical pictures fail in accounting for atomic phenomena" (1948), does not imply that QT is in any way ‘doomed’. On the contrary, he writes (1934), "only by a conscious resignation of our usual demands for visualisation and causality was it possible to make Planck's discovery fruitful in explaining the properties of the elements on the basis of our knowledge of the building stones of atoms." The reason for this resignation of the demand for visualisation (as well as Bohr's claim that the dual use of two radically divergent ontological schemes in QT does not lead to contradictions) follows directly from Bohr's analysis of the concept of complementarity. According to this analysis (1949), "any attempt to subdivide a phenomenon demands a change in the experimental arrangement introducing new possibilities of interaction between object and instrument which in principle can't be controlled." Since these experimental arrangements are mutually exclusive, Bohr concludes that "evidence from different experimental conditions can't be comprehended within a single picture" (cf postulate (d)). The atomic and plenum ontologies are two different attempts to incorporate in a single picture all physical phenomena (in particular the phenomena relating to microphysics) but this is what Bohr argues cannot be done according to QT. Also, since, as Hooker.(1973 p211) correctly points out, the radical difference between the two great ontological traditions in western science "is not perhaps more apparent to everyone only because elements of both views are built into ordinary language", and since Bohr demands that "the account of all evidence must be expressed in classical terms" (1949), including the terminology of ordinary language, it is not surprising that aspects of both ontological schemes are to be found in an exhaustive account of micro-objects regarding which "only the totality of phenomena exhausts the possible information about the objects" (1949). It is perhaps only surprising that the two opposing ontologies can be differentiated into a complementary rather than a contradictory relationship.

Summarising this analysis of duality and models we conclude that, according to Bohr, no monistic model can be given of micro-objects. The cry that this is unsatisfactory fails to appreciate that micro-objects cannot logically be discussed in isolation from measuring instruments (any attempted subdivision being a fundamental change) and also that classical models which seem able to explain the complete causality of classical physics were in fact never factual descriptions of micro-objects since they always contained negative analogies. In the second chapter we shall examine some of the various views as to how QT is able to cope with the fundamental appearance of non-causality (probability) in physics and finally see if complementarity, which Bohr regarded as "a rational generalisation of the classical ideal of causality" (1948; 1949), is able to account satisfactorily for this problem which is one of the principal difficulties in the CI. We shall see that the problem centres round one of the greatest differences between the atomic and plenum ontologies, superposition. According to the atomic ontology no two fundamental individuals can occupy the same spatio-temporal location whereas in the plenum case there is nothing to prevent the mutual interpenetration, or superposition, of ‘quality complexes‘. However, while superposition must be admitted in QT, the atomic ontology cannot be discarded in favour of a plenum ontology without the appearance of new and difficult problems.

Chapter Two: Superstates and Consciousness

I

One of the most difficult and sensitive problems raised by QT is referred to as the 'problem of measurement'. That this problem is indeed difficult and sensitive can be seen from the divergence of views, even within the CI, as to the nature and ultimate solution of the problem. Since measurement (or quantitative observation) provides the link between theory and experience, it has always played a central role in physics. Indeed, Campbell (1928) has defined physics as the science of measurement. However, in classical physics (CP) it is possible to ignore the problems raised by an analysis of the measurement process itself. According to CP a measurement involves an interaction between an object under observation and the measuring instrument used to make the observation. As well as this physical interaction there is also a psychophysical interaction between the measuring instrument and the observer (his sense organs and, ultimately, his consciousness). Thus measurement seems to involve ontological and epistemological problems concerning the relationship between human consciousness and physical objects, but these problems may be ignored in CP for the following reason. The interaction between object and measuring instrument implies an action of the object on the instrument together with an action of the instrument on the object. Since the action of the object on the instrument must produce large scale physical effect (eg the movement of a pointer) if the instrument is to serve as a measurement device, the order of magnitude of this action can be considered to be much greater than that of the instrument on the object, the latter being regarded as negligible or, at least, eliminable in principle. Of course, the action of object on instrument cannot be neglected because the resulting state of the instrument (eg pointer position) must depend on the state of the object if the instrument is to serve as a measuring device. Further, any psychophysical problems appearing as a result of the interaction between the instrument and observer were regarded classically as being outside the realm of physical theory. Thus CP was able to treat physical processes as being independent of observation, thereby eliminating the need to consider explicitly the observer. According to QT, however, one Planck represents a minimum and "therefore in principle non-negligible action between physical entities. Thus, in certain cases the action of object on instrument may have the same order of magnitude as the action of instrument on object so that the condition for the consistency of the classical conception of measurement fails to be satisfied and consequently the above mentioned problems of measurement can no longer be ignored.

The following are some tacit assumptions made in CP concerning the nature of measurement and the logical status of measuring instruments:

1. The influence of a measuring instrument on an observable quantity being measured can in principle be reduced indefinitely in such a way that the value of the observable immediately after the measurement has been completed has almost the same value as the observable had the moment before the measurement is made.

2. The concept of measurement plays no fundamental role in physical theory since it is assumed exactly the same laws are obeyed in interactions which correspond to measurements as are obeyed by all physical interactions. Thus measurement-type interactions are no more than a subclass of physical

interactions (although an important subclass) and are therefore reducible in principle to physical interactions which are not themselves measurements (see eg Einstein (194 p674)).

3. Microscopic physical processes (which include at least the final part of measurement processes) proceed in a manner which is independent of the presence or absence of a conscious observer.

We shall see how each of these assumptions is questioned in various interpretations of QT. Perhaps it should be mentioned that certain philosophies of science have found it necessary to question certain of the above assumptions even without explicit reference to QT. For example positivism and operationalism have denied assumption 2: positivism on the grounds that science is concerned only with what is observably given (and with immediate law-like generalisations from observation and experiment), therefore only what is observable is real; operationalism on the grounds that every physical concept only becomes useful and meaningful when defined by reference to physical phenomena and operations of measurement. Thus, for both these theories, measurement plays a fundamental role in physical theory. One could say that for these theories all physical interactions are regarded as a subclass of measurement type interactions rather than the other way round. Apart from such objections the above assumptions concerning measurement have usually been taken for granted in CP which generally adopts a realist philosophy.

It turns out, as a matter of convenience, that it is easier to discuss the peculiarities of measurement in QT in terms of the concepts of Wave Mechanics (WM) than in terms of those of Matrix Mechanics (MM). We shall therefore embark on a short digression concerning the equivalence of these two theories before beginning the analysis of measurement in QT.

The Matrix Mechanics of Heisenberg, Born and Jordon grew out of the work of Bohr and Kramers on atomic energy levels and spectral line frequencies and intensities. The theory associates with each observable quantity a matrix; these matrices obey a non-commutative algebra in contrast to the ordinary algebra obeyed by the quantities found in CP. This is the essential mathematical difference between MM and CP. If, for example, a and b are two observable quantities then, according to MM, there correspond to these quantities two matrices, A and B, such that, in general, $A.B - B.A \neq 0$ or, equivalently, $[A,B] \neq 0$. This is the fundamental property of a non-commutative algebra. Interpreting this we find that A.B means: first make an observation of quantity b and then an observation of quantity a and multiply the two results together. Similarly, B.A means perform these observations in the opposite order and multiply. Thus $[A,B] \neq 0$ or $A.B \neq B.A$ means that an observation of b and then a does not give the same result (multiplying the two observed values together) as an observation of a and then b. We therefore see immediately that the usual classical assumption that measurements can be performed without significantly affecting the value of an observable is not in general true in MM. It turns out that $[A,B] \neq 0$ in the case where a and b are complementary quantities while for all other sets of quantities $[A,B] = 0$. Further, it has been shown from purely mathematical considerations by Robertson (1929) that the product of the standard deviations of two self-adjoint matrices A and B is never less than $\sqrt{-1}$ times half the absolute value of the mean of their commutator, ie

$$\Delta a\, \Delta b \geq i\, ½|\langle A.B - B.A \rangle|$$

Thus the UP can be seen to follow logically from the non-commutative property of the matrices representing complementary observables in MM together with the physical result that [A,B] in the case where a and b are complementary is equal to $i\, h/2\pi$ ($\equiv i\, \hbar$).

The Wave Mechanics of Schrödinger adopts an entirely different point of view. Starting from de Broglie's hypothesis that wave-particle duality is a general property of microscopic objects applying equally to matter and light, Schrödinger generalised the notion of matter waves by producing a wave equation which describes the propagation of a wave function, ψ, which represents the state of a quantum system. According to Dirac, "Schrödinger got this equation by pure thought, looking for some beautiful generalisation of de Broglie's ideas, and not by keeping close to the experimental development of the subject in the way Heisenberg did." It is, however, interesting to note one similarity between both these forms of QT in their historical development. Both were constructed by formal mathematical methods before the appearance of any substantial physical interpretation. Born's probabilistic interpretation of the wave function and the development of the quantum theory of measurement by Bohr and by Heisenberg came only after the acceptance of the mathematical theory.

It is generally accepted that in 1926 Eckart and Schrödinger proved the theoretical equivalence of WM and MM. However, Hanson (1965 Ch8) has cast serious doubt on such claims. He argues that, at most, Schrödinger and Eckart prove the mathematical identity of WM and MM but this is not the same as proving that WM and MM are equivalent physical theories. As physical theories, WM and MM were, according to Hanson, born with different interpretations built into them, they were therefore fundamentally different physical theories which could never be proven equivalent. Born, however, by means of his probabilistic interpretation of $|\psi|^2$, managed to give both mathematical algorithms the same physical interpretation. In this way, and by assuming both algorithms to be mathematically identical, he was able to render them equivalent as physical theories. But by thus changing the original physical interpretations of WM and MM, Hanson argues, Born did not prove the equivalence of WM and MM but of two new theories 'WM' and 'MM' which were quite different from the original theories. "By finding the Matrix mechanical analogue of $|\psi|^2$ and by interpreting both that matrix and the ψ function as giving the probability of finding a microparticle within a given volume element," Hanson (1963 p129) points out, "Born alone deserves the credit for establishing the equivalence as physical theories of WM (ie 'WM') and MM (ie 'MM'). Historically speaking, it would be erroneous to suppose that anything approaching Born's achievement is to be found in the Eckart and Schrödinger papers."

Hanson, however, continues his critique by demonstrating that the physical equivalence arguments of Born as well as the mathematical identity arguments of Eckart and Schrödinger are entirely inductive in nature. Born shows that for every problem so far considered every Born-interpretable observation statement in 'WM' has a corresponding statement in 'MM'. While it is tempting to conclude from this that the correspondence will hold for any problem which can be formulated in 'WM', this is an inductive generalisation which has not itself been proved. Indeed, it is hard to see how such a generalisation could be proved for these physical theories since they are each individually 'open-ended with regard to discussion of future possibilities. Similarly, the 'proofs' of Schrödinger and Eckart involve an inductive generalisation. They do not prove generally that every MM formulation and solution to any problem has some WM analogue, but rather they prove that this is so for certain selected typical types of microphysical problems. To then assert that this constitutes an identity proof for the two mathematical systems is therefore suspect since an inductive generalisation is involved. Further, it is sometimes claimed that Schrödinger and Eckart have shown that the mathematical algorithms of WM and MM both follow from a more general operator calculus (O) and, therefore, that WM and MM are mathematically identical. This, however, involves a logically invalid inference:

If O implies WM and
If O implies MM
Then WM is equivalent to MM,

which is obviously invalid. There is thus no strict mathematical proof of the equivalence of WM and MM either as mathematical algorithms or as physical theories. Despite this fact, Hanson concludes that the arguments of Schrödinger, Eckart and Born amount to one of the strongest demonstrations of equivalence that one could ever have between physical theories. Thus there is a high degree of justification, particularly since Born has managed to give the same physical interpretation to both theories, for being able to discuss the problems relating to measurement in QT in terms of whichever theory, WM or MM, is most convenient. We shall talk primarily in terms of the ψ function of WM and of its interpretation in this chapter.

According to WM the state a quantum system is specified by means of the function ψ which is the solution of a wave equation, Schrödinger's equation, which can be written

$$i\,\tfrac{h}{2\pi}\,{}^{d}/_{dt}\;\psi(\mathbf{r},t) = \left(-\,{}^{h^2}/_{8m\pi^2}\;\Delta + V\right)\psi(\mathbf{r},t) \qquad \ldots 10$$

where m is the mass of the system which is in a potential V and where

$$-\,{}^{h^2}/_{4\pi^2}\;\Delta = ({}^{h}/_{2\pi i})^2\,({}^{d^2}/_{dx^2} + {}^{d^2}/_{dy^2} + {}^{d^2}/_{dz^2})$$

The system is thus described by a wave whose amplitude varies in space and time and which is expressed mathematically by the function ψ(**r**,t). The first interpretation given to this wave by de Broglie and by Schrödinger was simply that physical systems are these waves. This interpretation, however, is too simple and faces grave objections. First of all these waves, as represented by ψ, are much more complicated than ordinary waves eg sound waves or water waves. The amplitude of these ψ waves is complex (ie involving the square root of -1) rather than real as is the case with ordinary waves. Also these waves are not waves in ordinary space but in an abstract mathematical space (ie Hilbert space in the Dirac and von Neumann formulations of QT) which may have many more than three dimensions, the waves are therefore impossible to visualise in terms of ordinary waves in physical three dimensional space. A more serious objection to this interpretation is the fact that according to WM the shape or structure of the wave is dependent on the observable being measured. For example, if position is the observable quantity under examination then ψ will be considered to lie in a space the dimensions of which correspond to physical spatial dimensions, one dimension for each degree of freedom of the system. In the case of a quantum system containing two particles having coordinates (x_1, y_1, z_1) and (x_2, y_2, z_2), for example, the ψ function for the system would have six dimensions, one for each position coordinate. However, if momentum were the observable being measured then the wave would be considered to lie in a space the dimensions of which correspond to the six momentum coordinates of the system. The ψ function expressed in terms of the position coordinates is called the position representation of the system while that in terms of the momentum coordinates is called the momentum representation. It is possible to transform from one representation to the other by means of the rules:

$$\psi(\boldsymbol{r}) = h^{-\frac{3}{2}} \int \emptyset(\boldsymbol{p})\, e^{i\boldsymbol{p}.\boldsymbol{r}/\hbar}\, d\boldsymbol{p}$$

$$\emptyset(\boldsymbol{p}) = h^{-\frac{3}{2}} \int \Psi(\boldsymbol{r})\, e^{-i\boldsymbol{p}.\boldsymbol{r}/\hbar}\, d\boldsymbol{r}$$

The functions ψ and ϕ are equivalent representations of the same dynamical state but both correspond to different waves and it is therefore difficult to think of these waves as physically existing entities. Finally, the wave cannot be regarded as a physically existing wave because in almost every interpretation of QT the evolution of the wave is not always continuous as suggested by equation 10. The wave does not always obey the Schrödinger equation but sometimes changes discontinuously in a way which is apparently without cause and not governed by any mathematical equation. For example, in the two slit experiment, the electron is described by a wave, ψ, which divides into two parts on passing through the screen, ψ_A the part passing through the slit A and ψ_B the part passing through the slit B. These two waves then interfere giving the expectation of an interference pattern at the detector. As we have seen however, each electron produces one localised dot on the detector rather than a spread out interference pattern as suggested by this interpretation of the wave function in the Schrödinger equation. Thus the wave, if regarded as physical, must collapse to a point before hitting the detector. From the arguments in Chapter One we see further that the collapse must take place instantaneously and at the surface of the detector plate. In order to obtain the final interference pattern after a large number of electrons have passed through the apparatus, each wave must collapse at a point whose position on the detector although unpredictable in individual events must correspond to a certain probability distribution determined by the predicted interference pattern. Thus the wave must undergo not only a continuous expansion as prescribed by the Schrödinger equation but also a discontinuous reduction or 'jump' not controlled by any known equation or physical process.

Schrödinger had thought that his interpretation of WM as opposed to the Bohr-Heisenberg interpretation of MM eliminated all need for talk in terms of quantum jumps and discontinuous changes which were evident in discussions of MM. In September 1926 Schrödinger was invited by Bohr to Copenhagen to discuss the interpretation of QT but he was unable to convince Bohr and Heisenberg that a continuous description of quantum phenomena could be given if ψ is interpreted as a physically existing wave. The depth of the gulf separating the interpretation of WM by Schrödinger and the interpretation of MM by Bohr and Heisenberg, even though Schrödinger's 'proof'' of the formal equivalence between WM and MM had been known for six months, can best be seen from an exchange between Bohr and Schrödinger during his visit as reported by Heisenberg (1955). At the end of the discussion Schrödinger exclaimed: "If all this damned quantum jumping were really to stay, I should be sorry I ever got involved with QT," whereupon Bohr replied: "But we others are very grateful to you that you did, since your work did so much to promote the theory." There must be few examples in the history of science where the interpretation of a remarkably successful fundamental theory proposed by the innovator of the theory is so rapidly rejected by the scientific community to be replaced by a much more radical interpretation. The appearance of this alternative interpretation finds its explanation in another historically unusual circumstance, the almost simultaneous appearance of a rival theory of equally successful predictive power yet of entirely different formal character, MM. If Schrödinger's interpretation of WM is less physically penetrating than that of Bohr and Heisenberg, this may well be explained by Schrödinger's more mathematical than physical approach in constructing his theory. Yet there is a further reason too, for without Born's probabilistic interpretation of the wave function WM is generally regarded as incomplete. Under this circumstance Schrödinger could hardly be expected to give a satisfactory interpretation of his equation.

Born, who was deeply involved in the development of MM with Heisenberg and Jordan, was greatly impressed by the new approach of Schrödinger's WM, but he could not accept Schrödinger's interpretation of it in terms of a 'causal continuum theory in the classical sense'. As Born (1943 p23) explained: "On this point I could not follow him. This was connected with the fact that my institute and that of James Franck were housed in the same building of the Gottingen University. Every experiment by Franck and his assistants on electron collisions ... appeared to me as proof of the corpuscular nature of the electron." Very soon after the publication in 1926 of Schrödinger's WM Born published his original probabilistic interpretation of the ψ function; according to this interpretation the modulus of ψ(**r**) squared gives the probability of a particle being at position **r**, or more accurately $|\psi(\mathbf{r})|^2 d\mathbf{r}$ measures the probability density P(**r**)d**r** of particles within the volume element d**r**, ie between **r** and **r** + d**r**, the particle being conceived in the classical sense as a point mass possessing at each instant both a definite position and a definite momentum. Thus ψ does not represent the physical system as was Schrödinger's view but only our knowledge of the system. WM therefore, according to this interpretation, cannot in general answer questions concerning the definite state of a quantum particle after interaction other than questions concerning the probability of finding the particle in a given definite state. WM is no longer seen as a causal theory of particle interactions but is a quantitative probabilistic theory. "The motion of particles," writes Born (1926 p304), "conforms to the laws of probability, but the probability itself is propagated in accordance with the law of causality." Thus Schrödinger's equation 10 describes the propagation of probability rather than of the particle itself.

We can now begin to see how this probabilistic interpretation of WM reveals the basic physical equivalence of WM and MM. For example, consider making a simultaneous prediction of both position and momentum of a quantum object. According to the above interpretation of ψ(**r**), an exact prediction of the position of the object can only be given if the wave is such that it occupies a very small volume in position space (ie the mathematical space of ψ in the position representation). Then the probability that the particle is to be found in the corresponding small volume of physical space will be very high indeed. However, a glance at the Fourier transformation rules of equations 11 reveals that if the wave ψ(**r**) occupies a small volume of position space then the wave ϕ(**p**) will occupy a correspondingly large volume of momentum space. The momentum of the particle will therefore be relatively uncertain showing that precise simultaneous knowledge of the position and momentum of a quantum object is impossible according to WM. This is the wave mechanical analogue of the UP in MM.

It is not hard to see how this interpretation is able to deal with the objections to Schrödinger's interpretation. Since ψ is not itself regarded as a physically real wave, the fact that it is multidimensional and complex valued represents no serious obstacle; only its squared absolute value which is a non-negative real number has physical significance. Further, since the wave function is viewed as an expression of our state of knowledge of the system, the instantaneous reduction of the wave packet on measurement is easily accounted for, since this only reflects a sudden change in our knowledge of the system from a set of possibilities to a single well defined certainty when the result of a measurement becomes known. Finally, the representation-dependency of the wave function is only to be expected since knowledge about position gained from the position representation is naturally different from the knowledge about momentum gained from the momentum representation.

Despite this success, and in particular the initial success of this interpretation in the field of atomic scattering, Born's original interpretation does, nevertheless, have difficulties of its own. According to Born, electrons are ordinary classical-type particles. However their specific motions are not governed by any deterministic laws. Only the probability of electron position, momentum etc are determined by the quantum laws. Thus, for example, in the double slit experiment, Born supposes that each electron must pass through either slit A or B but not both since we are dealing with a classical-type particle. However if quantum mechanical probability is applied in this way to electrons which are assumed to conform to a classical ontology then it seems probability must be interpreted as an expression of lack of knowledge only (ie probability has no ontological significance). This view of quantum probability has been called the ignorance interpretation. But this interpretation cannot account for interference in the double slit experiment which involves interaction between the two components, ψ_A and ψ_B of the electron wave function. Unless ψ has ontological force quantum probability cannot be construed as indicating lack of knowledge only.

This is a problem for Born's original interpretation. However, in his later interpretation this problem is avoided. Instead of interpreting $|\psi_e(\mathbf{x})|^2$ as the probability that a particular classical-type electron exists at position x, Born later came to interpret $|\psi_e(\mathbf{x})|^2$ as the probability that one would find an electron at x if one were to make a position measurement. Thus, in this later interpretation, it is not necessary to suppose that a particle exists at a particular place until one makes a measurement. This interpretation avoids the difficulties encountered by the ignorance interpretation.

In the analysis of the double slit experiment it became clear that exactly the same results are obtained if the intensity of the source is reduced until only one electron is in the apparatus at a time. The ψ function associated with each particle must therefore divide into two components, each passing through a different slit, which then interfere with one another. This mathematical interference is manifested physically by the final interference pattern on the detector. Dirac (1930 sec3) makes this same point with regard to photons: "The new theory which connects the wave function with probabilities for one photon gets over the difficulty (of explaining interference phenomena) by making each photon go partly into each of two components. Each photon then interferes only with itself." Indeed, he continues, "Interference between two different photons never occurs." Thus, in the double slit experiment, the wave ψ describing the electron's state divides into two components, ψ_A which passes through slit A and ψ_B which passes through slit B. These two components then interfere between the screen and the detector. The probabilistic description of the electron then becomes: The probability of finding the electron in a region between y and y + dy on the detector when:

1. Only slit A is open, is $P_A(y)\,dy = |\psi_A(y)|^2\,dy$

2. Only slit B is open, is $P_B(y)\,dy = |\psi_B(y)|^2\,dy$

3. Both slits are open, is $P_{A\&B}(y)\,dy = |\psi_A(y) + \psi_B(y)|^2\,dy$ … 12

$$\neq P_A(y)\,dy + P_B(y)\,dy$$

Expanding the expression for $P_{A\&B}(y)$ gives $P_A(y) + P_B(y)\,dy + \psi_A^*.\psi_B + \psi_B^*.\psi_A$

The last two terms in this expression explain mathematically the deviation from the additive pattern; they are called the interference terms. Because this mathematical account of interference finds physical manifestation in terms of the final particle distribution on the detector it is natural to suppose that the ψ function represents something physical and does not merely represent our knowledge, but then the above difficulties attaching to a realistic interpretation of ψ once more appear.

Heisenberg, who accepted Born's interpretation of the square of the absolute value of the wave function in terms of probability, attempted to ascribe a certain reality to the probability wave by distinguishing between aspects of reality, the possible and the actual. He later wrote (1961 p9) that he regarded the probability waves as "a quantitative formulation of the concept of δύναμις (possibility) or, in the later Latin version, potentia, in Aristotle's philosophy. The concept that events are not determined in a peremptory manner, but that the possibility or 'tendency' for an event to take place has a kind of reality - a certain intermediate layer of reality, halfway between the massive reality of matter and the intellectual reality of the idea or the image - this concept plays a decisive role in Aristotle's e philosophy. In modern QT this concept takes on a new form; it is formulated quantitatively as probability and subject to mathematically expressible laws of nature."

In this way Heisenberg ascribes to the concept of probability in QT a certain ontological aspect in addition to its usual epistemological content. This is similar to the conclusion with regard to uncertainty in the UP of MM according to the von Neumann interpretation where the uncertainty associated with complementary variables is accorded ontological force in the sense that it is considered impossible for a quantum object to have simultaneous definite values of complementary variables. While UP may he regarded as the basic principle of QT according to MM, according to WM the superposition principle (SP) may be regarded the basic principle. Before considering the interpretation and implications of the SP we simply state at this point that the UP and SP, although originally appearing independently in MM and WM respectively, can both be derived from the same generalised formalism of QT constructed by Dirac (1930). To this extent the two principles are not entirely independent.

The SP, or principle of superposition of states, has been described by Putnam (1965 p82) as follows: "Let S be a system that has various possible states A, B, C ... according to classical physics. Then in addition to these states, there will also exist in quantum physics certain 'linear combinations' of an arbitrary number of these states, and a system in one of these may, in certain cases, behave in a way that satisfies no classical model whatever." Thus, for example, in the double slit experiment where each possible state is represented by a ψ function, let the state A in which the electron is assumed to pass through slit A be represented by the function ψ_A and the state B in which the electron is assumed to pass through slit B be represented by the function ψ_B. We have seen above in equation (12) how the correct predictions for the probability of arrival of the electron at a certain point on the screen is giveb by taking the state of the system (ie electron when both slits A&B are open) to be represented by the wave function $\psi_A + \psi_B$, ie a linear combination of the two functions representing the two possible states of the system. The surprising thing about this result is that every classically mutually exclusive possibility has an effect on the final outcome. Classically, of course, only one out of the (many) possibilities is assumed to have any reality (although it is not known which) and therefore only one of the possibilities can have any physical effect. In quantum physics, however, it seems that every classical possibility has an effect on the final experimental outcome. It is for this reason that no classical model can be found to account for all the results of, say, the double slit experiment. There is a problem here as to what is meant by a classical possibility when it is admitted that there is no single consistent classical model. This is most easily resolved by saying that 'classical possibility' is to be taken in the sense of classical possibility in the case where a quantum is regarded as a particle, having definite simultaneous values of complementary variables and therefore following a definite trajectory in space and time. Thus, in the double slit experiment, there are two possibilities: the electron may either pass through slit A or slit B. The wave functions describing each of these possibilities are then combined give the wave function for the system as a whole.

It should be remembered that there is a constraint imposed on the linear combination allowed by the SP which appears as a result of the law of total probability which states that the total probability must always equal unity, or in other words the sum of all the probabilities associated with each, possibility must equal one - this ensures that the set of possibilities is complete. Thus the probability P that the electron lands somewhere on the detector plate between y = - ∞ and y = + ∞ is

$$\int_{-\infty}^{+\infty} P(y)\, dy = \int_{-\infty}^{+\infty} |\psi(y)|^2\, dy = 1$$

In the case where slits A and B are symmetrically positioned about the x axis, and therefore where ψ_A and ψ_B appear in equal proportions in the superposition, this normalisation condition imposed on the probabilities implies that, strictly speaking, the superposition of wave functions ψ_A and ψ_B in the case where both slits A & B are open should be equal to $(½)^{½}\psi_A + (½)^{½}\psi_B$. This ensures that

$$\int_{-\infty}^{+\infty} P_{A\&B}(y)\, dy = \int_{-\infty}^{+\infty} \left|(½)^{½}\psi_A + (½)^{½}\psi_B\right|^2 dy = 1$$

It is, however, relatively unimportant to the general argument to keep adding the correct normalisation constants (usually $(½)^{½}$ in this discussion). They will therefore often be omitted in this discussion of superpositions of states. Also, superpositions of states, for example $\psi_A + \psi_B$, will be called superstates for brevity. (This definition of superstates should be clearly distinguished from that of Feyerabend (1962 p215) who defines a superstate as a quantum state which has been further clarified by the addition of new 'hidden variables' (see Ch4, II).)

We have then the situation that in WM the interphenomenon in, for example, the double slit experiment is represented by a superstate $\psi = \psi_A + \psi_B$. If, however, a measurement is made to determine through which slit, A or B, the electron passes, the result of the measurement will be represented by either the state ψ_A or the state ψ_B depending on whether the electron is seen to pass through slit A or slit B, respectively. It is not, however, possible to regard the superstate as representing both mutually exclusive alternatives in a simple disjunctive mixture; this is the ignorance interpretation whose predictions do not agree with the observations. All the elementary states in the superstate are needed in order to account for the phenomena associated with superstates, eg the interference pattern. But if an observation is made to determine in which state of all the possible states in the superstate a quantum object exists (eg whether the electron passes through slit A or B) then the observed result is always one and only one of the states making up the superstate. This would be easy to understand in the ignorance interpretation, but since the superstate may be given ontological as well as epistemological significance due to the necessity of attributing observed interference phenomena to the effective cooperative action of all states in the superstate, the ignorance interpretation seems to be inadequate. We are now in a position to express the most important problem relating to measurement in QT: In a measurement a superstate is 'reduced' or 'col1apses' into an individual state, or, in other words, a probability distribution becomes a single certainty. No mechanistic explanation of this reduction is given in orthodox QT.

Since it is sometimes erroneously supposed by some writers that superstates could never be observed, because they suppose that macroscopic quantities must always have definite values, ie must always be in one and only one definite state, let us consider an extension of the double slit experiment due to Schrödinger (1955) called the 'Schrödinger cat paradox'. Imagine the source in the double slit experiment emits only one single electron which may pass with equal probability through slit A or B. Behind A is a box containing a cat and an electrocuting device attached to the cat which will be switched on automatically if the electron passes through slit A and hence kill the cat. Behind B is an empty box. Both boxes are closed so that we cannot see in. According to WM the wave describing the electron ψ is split into two on arrival at the screen, one part, ψ_A, going through slit A and the other part, ψ_B, going through slit B. Since both slits are equidistant from the source and of equal aperture, the waves ψ_A, and, ψ_B have equal amplitude, that is the probability P_A that the electron passes through slit A is equal to the probability P_B that the electron passes through slit B is equal to a half since the electron is supposed to travel through one or other of the slits and the total probability is equal to unity. So

$$P_A = P_B = ½$$

Since, if the electron passes through A the cat will die and if it passes through B the cat will live, this implies that the probability that the cat is dead, P_d, (after the electron has left the source and passed through the screen) is equal to the probability that the cat is alive, Pa, is equal to a half.

$$P_d = P_a = ½$$

However, remembering the inadequacy of the ignorance interpretation, this does not imply that the cat is either alive or dead but we just don't know which for the state of the system must be described by a superstate $\psi_A + \psi_B$, or in other words $\psi_{dead} + \psi_{alive}$ (the cat now being described quantum mechanically in terms of waves), and this superstate has ontological as well as epistemological significance. Instead, if we presuppose macro-objects are accurately described by QT, we should talk of a 'supercat' which has 50% probability of being found to be dead and 50% probability of being found to be alive when an observation is made but before such an observation cannot be said to be alive or dead, or even one or the other. It is rather in a superposition of both these states. Since many quantum mechanical experiments are of the type in which a microscopic uncertainty is amplified so as to produce something which is macroscopically detectable, these experiments are similar to the cat experiment in that the state of the system prior to measurement is a superstate. If this analysis is correct then far from being clear that superstates can never be observed the question becomes 'why are superstates never observed?' It is in answer to this question that the various interpretations of most readily diverge.

The question is usually seen as this: At what point does the superstate collapse into one or other classical state? Let us consider some various answers to this question. First, that of Schrödinger, according to whom the wave function is to be interpreted as a real physical wave: in this interpretation no account of any change in the wave beside the continuous causal change according to the Schrödinger equation is given. According to this interpretation, therefore, superstates should be observable, in particular a supercat (whatever that would look like) should be observed as the outcome of the cat experiment. This is not the case (an ordinary alive or dead cat being observed) and therefore we need no longer consider this interpretation.

Next we consider the interpretation generally regarded as the orthodox – the CI first espoused by Heisenberg, but closely related to the Born interpretation. Heisenberg seems sometimes to come very close to making the same mistakes as Born did in his original interpretation since he also tries to retain an interpretation of probability in QT as ultimately an expression of our lack of knowledge. As Heisenberg (1930 German edition, sec IV.2) pointed out as early as 1930, the process of measurement must be "divided into two shaply distinguished parts, the first of which is to subject the system to a physically real external interaction which alters the course of events … In consequence, the system is brought into an assemblage (superposition) of states, in general infinite in number … The second part of the measurement selects as the actual state one particular state out of the infinity of states in the assemblage. This second part does not itself influence the course of the process, but merely alters our knowledge of the actual state of affairs."

Thus, according to Heisenberg, we have above only considered part of the measurement process namely that part which produces the superposition. There is a second part to the complete process by which the superstate is reduced to one actual state. By referring both to individual states and to states of knowledge Heisenberg can be seen to be attempting to make an epistemological extension to the ontological interpretation according to which a micro-object in a superstate cannot be regarded as being in any definite classical state.

However, by claiming that the second part of the measurement process, which seems to be the part which according to Heisenberg incorporates the reduction of the wave function, "does not itself influence the course of the process, but merely alters our knowledge of the actual state of affairs," Heisenberg would seem to be committing the same error as Born because only in the ignorance interpretation can reduction of the wave function be regarded as not influencing the course of the process. Perhaps it is for this reason that the above quoted passage does not appear in the 1949 English translation. In clarification of this second part of the measurement process in QT Heisenberg (1959 p54) writes: "The observation itself changes the probability function discontinuously; it selects of all possible events the actual one that has taken place. Since through the observation our knowledge of the system has changed discontinuously, its mathematical representation also has undergone the discontinuous change and we speak of a 'quantum jump'. If the old adage 'Natura non facit saltus' (Nature does not make jumps) is used as a basis for criticism of QT, we can reply that certainly our knowledge can change suddenly and that this fact justifies the use of the term 'quantum jump'."

By the word observation, Heisenberg does not wish to include the mind of the observer or consciousness. As far as Heisenberg is concerned "the transition from the possible to the actual," or in other words the transition from a superstate to a single classical state, "takes place as soon as the interaction of the object with the measuring device, and thereby with the rest of the world, has come into play; it is not connected with the act of registration of the result by the mind of the observer." Thus, for Heisenberg, there is no question of the macroscopic world being at any stage in a superstate.

According to Heisenberg, the macroscopic world is always in a definite classical state in accordance with the suppositions of common sense. The microscopic world is not, however, in general, in a definite classical state but is rather in a superstate which takes the form of 'possibilities' or ‘tendencies’ or 'potentia' in Aristotelian philosophy according to Heisenberg. According to this view, therefore, it must be accepted that elementary processes such as the scattering of an electron in a definite direction could not be assumed to occur since this would require a reduction of the wave function. Only when in interaction with a macroscopic body can such a process be said to occur. Thus, if the whole physical universe were composed entirely of microphysical entities then, according to Heisenberg, the universe would not consist of real events but of evolving potentialities (time dependent ψ functions). The reduction of the wave function in a measurement is thus the transition from the possible to the actual.

A similar view to this has been proposed by Margenau (1954) in his ‘latency interpretation’ of QT. According to this interpretation, what is real is the wave function ψ whereas the classical state observables, such position, momentum or energy may be said to be latent in the sense that their values emerge only in response to a measurement. Thus, latent observables are "not always there" but "they take on values when an act or measurement, a perception, forces them out of indiscriminacy or latency." Since, according to Heisenberg, the measurement of an observable is a "transition from the possible to the actual," Heisenberg’s notion of ‘possible’ may well be likened to Margenau's conception of ‘latency’. Bohr has also adopted a view very close to that of Heisenberg. Hooker (1973 p204), however, writes: Any theory whatever, so far as I can see, could have its problems ‘solved’ by this approach - simply because the concept of an ‘actualization' of a ‘potential’ is so vague and intrinsically not open to direct investigation of its structure. But, whatever their own reasons, it seems that each of these three writers have more or less abandoned this point of view. Bohm has totally abandoned this approach in favour of hidden variables (although according to Hooker he has recently turned to the exploration of a "much more radical conceptual structure for QT”). Heisenberg concentrated on non-linear unified field theories while Margenau has adopted a more positivist approach.

Heisenberg thus attempted to retain the classical position of macro-physical realism whereby all macro-objects are always in a definite classical state, and the view that when an observation of a micro-object is made, the resulting reduction of the wave function corresponds only to a reduction in the sense of the ignorance interpretation, since the measuring instrument is itself a macro-object and is always in a definite state. But his interpretation is not without difficulties. If macrophysical realism is accepted then it is possible to regard the epistemological aspect of reduction in terms of the ignorance interpretation. But there is also an ontological aspect to the reduction process, since the micro-object is assumed by Heisenberg to be in a superstate while the corresponding macro-object (measuring instrument) is assumed to be in a definite classical state. For example, in the cat experiment, the electron is in a superstate $\psi = \psi_A + \psi_B$ while the cat is always in a definite state $\psi = \psi_{dead}$ or $\psi = \psi_{alive}$. But, according to QT, the cat should also be in a superstate, unless it can be shown that an ontological reduction of the superstate has taken place at some stage in the interaction process between electron and cat. This is exactly what Heisenberg does implicitly assume. But he does not explain either why this reduction takes place, or where, in the chain between electron and cat, the reduction occurs. Rather he seems to argue that reduction must occur somewhere between electron and cat because QT requires that the electron is in a superstate albeit potentialised and macrophysical realism requires that the cat is in a definite classical state.

The CI as adopted by most working physicists takes the view that the reduction occurs at the first macroscopic object; however, as pointed out by both Bohr and Heisenberg, there is no clear-cut division between macro and microscopic objects. The division is to a certain extent arbitrary. With regard to the reduction of the wave function the CI is therefore rather unsatisfactory from both a physical and a philosophical point of view. From the physical point of view it is unsatisfactory since superstates and states are ontologically different (the former admitting the possibility of interference), and therefore reduction, if it is to occur prior to the macrophysical level, must correspond to a physical process. Details of this process have not been given although an attempt has been made in this direction by Bedford & Wang (1975) and, more seriously, it seems that the stage at which this process occurs cannot be identified, but must retain an element of arbitrariness. (Remember, for example, how in the analysis of the double slit experiment the screen with the slits was at one time regarded as a measuring instrument and therefore in a definite classical state while at another time it acted as the object under observation and was therefore treated as a superposition of states.) From a philosophical point of view the CI is unsatisfactory because in order to retain the classical ontology of macroscopic objects it entails abandoning the methodological tradition of looking for universalisable theories. If QT is correct and is a universalisable theory, then it must apply to systems of arbitrary size since it applies both to the individual electrons and to nuclei as well as to the interactions between them. In particular, it must apply to macrosystems. Therefore, since there seems to be no mechanism within QT to collapse the superposition, macrosystems should, in certain circumstances (eg the cat in the cat experiment), appear as superstates to conscious perception. Since this is denied by the CI, it must be admitted that QT is not a universal theory, in particular it does not apply to macrosystems. But there are good reasons for supposing on other grounds that QT does in fact apply to macrosystems as well as microsystems. Quantum many-body theory, for example, has yielded a number of new and well confirmed predictions concerning the properties of certain macroscopic systems (eg crystals) which can in no way be explained classically but which require the assumption that the laws of the microworld are also applicable to the macroworld in contradiction to the claims of Heisenberg and the CI.

For Heisenberg it seems that the principal feature of a measuring instrument is the fact that it amplifies or transforms a microphysical (and therefore quantum theoretically describable) event into a macrophysical (classically describable) event. According to the macrophysical realism of Heisenberg the final state of a measuring instrument is definite because it is macroscopic. But Bohr, on the other hand, seems to place much more emphasis on the fact that a certain system is being used as a measuring instrument. For Bohr, it is not the fact that a system is macroscopic that is the significant feature in the problem of the reduction of the wave function in QT, but rather the logical (though not physical) necessity of drawing a sharp distinction between the object and the measuring instrument. The resolution of the problem, according to Bohr, seems to depend on an analysis of the measuring instrument in its function as a measuring instrument (recall the dual treatment of the screen in the double slit experiment).

II

We are now in a position to discuss von Neumann's (1932) theory of measurement. In so far as von Neumann regarded a measurement as logically different from other physical processes he is in agreement with Bohr. In contradiction to Bohr, however, von Neumann treated the measuring instruments as quantum mechanical objects and was able to show that QT is capable of accounting consistently for the operation of measuring instruments. For von Neumann a measurement is essentially "the temporary insertion of a certain energy coupling into the observed system". The wave functions of the observed system or object and the measuring instrument should therefore be combined into a single wave function describing the situation as a whole:

$$\psi_{\text{Total system}} = \psi_{\text{Object}} \otimes \psi_{\text{Instrument}}$$

where ⊗ indicates a tensor product. However, this again reveals that in general the total system is in a superstate, ie $\psi_{\text{Total system}}$ is a superstate.

More recently, Fine (1970 p2783) has given added support to this conclusion in a proof which claims to show that it is impossible to avoid the final outcome that in certain experiments (eg the cat experiment) the resulting state of some macroscopic objects must be a superposition, even if one allows for some degree of doubt as to the initial state of the macroscopic object. Fine concludes that, strictly speaking, no laboratory observations can be cited in support of the QT" because no explanation is given in QT as to how predicted superstates collapse into the individual definite states which are in fact observed without fail in laboratory experiments. Fine (1970 p2787) points out that "in practice one treats the final superposed state of the joint object-apparatus system as having the same observational significance as some corresponding mixed state. The enormous difference between the two, which is the difference between the pointer (say) aiming at some position or other in the mixed state but at no position at all in the superposed state, is treated as though it were no difference." Since this is the only strategy which makes experimental support for QT possible, this is the strategy adopted by all experimental physicists. Many interpretations of measurement in QT can be regarded as alternative formulations of such a strategy.

Von Neumann introduced what he considered to be the essential difference between an ordinary physical interaction and a measurement. A measurement, he reasoned, is not complete until an act of observation has been made by an experimenter in order to determine the outcome of the measurement. It is this act which, according to von Neumann, collapses the superstate. But where and how does this act cause reduction of the wave function? Von Neumann again was confronted with the problem of how a superstate is reduced to an individual state, or in von Neumann's terminology, how the multidimensional superposition of states is 'projected' into one dimension of the underlying Hilbert space (the space spanned by the vectors or wave functions in the superposition). Being influenced by the ideas Szilard (1929) in his study of the relation between the intervention of an intelligent being on a thermodynamic system and the second law of thermodynamics, according to Jammer (1974 p480), von Neumann contended that it is impossible to formulate a complete and consistent theory of quantum mechanical measurement without reference to human consciousness. He thus proposed that the process of reduction or projection can only take place by way of the intervention of a human mind or consciousness. As to how this reduction takes place von Neumann was rather more reticent. He proposed in his axiomatisation of QT that quantum mechanical states can change in two kinds of way:

1. The first way is the one proposed by Schrödinger; that is according to the wave equation 10. This, according to von Neumann (1932), allows "continuous and causal changes in the course of time." These changes von Neumann called "automatic changes"; they are reversible.

2. The second way he called "arbitrary changes by measurement". These changes are "the discontinuous, non-causal and instantaneously acting experiments or measurements". These are irreversible.

$$\sum_i \psi_i \;\rightarrow\; \psi_j$$

But von Neumann did not go further in describing a mechanism or theoretical analysis of processes of the second kind. He writes (1952 p134; 1955 p439): "Now quantum mechanics describes the events which occur in the observed portions of the world, so long as they do not interact with the observing portion with the aid of the process (of the first kind), but as soon as such an interaction occurs, ie a measurement, it requires the application of (a) process (of the second kind)." Thus, according to von Neumann, it is necessary for a complete account of the results of QT to include processes of the second kind. Further, since such processes do not occur in the observed portions of the world they cannot be said to occur anywhere, however deeply, within the observer's body for this too must come under the category of the observed portions of the world governed by "automatic changes". They can only occur in his consciousness. Consciousness, therefore, according to von Neumann, does not play a passive role (as it does, for example in the ignorance interpretation of Born) but an active role in the completion of a measurement in QT. In the terms of Heisenberg, one could say that the whole material world, macroscopic and microscopic, is for von Neumann the 'potentia' which is only 'actualised' through the intervention of mind or consciousness. Von Neumann's theory is thus duelistic since it divides processes into two irreducible categories corresponding to the partition of the world into the observed and the observing portions.

In 1939 London and Bauer decided to write a ‘concise and simple presentation of von Neumann's theory of measurement’. With regard to the reduction of the superposition of states they wrote (1959 p42): "The observer has a completely different point of view: for him it is only the object and the apparatus which pertain to the external world, to that which he calls 'objective'. By contrast, he has with himself relations of a very peculiar character: he has at his disposal a characteristic and quite familiar faculty which we can call the 'faculty of introspection’ for he can immediately give an account of his own (perceptions and hence his own) state. It is in virtue of this ‘immanent cognition’ that he lays claim to the right to create for himself his own objectivity, namely, to cut the claim of statistical coordination expressed by (the superstate $\sum_i \psi^i_{Object} \otimes \psi^i_{Apparatus} \otimes \psi^i_{Observer}$) by stating: 'I am in state $\psi^i_{Observer}$' or more simply: 'I see ... (the observable A has the value) a^i'."

Thus according to London and Bauer the solution to the cat experiment is to be found in the claim that whether the cat is alive or dead is decided only at the moment the observer opens the box and looks at the cat. Prior to that moment the cat cannot be said to be in either state independently of the other. Jammer (1974 p484) has compared the recourse to human consciousness of von Neumann, London and Bauer in their solution to the problem which their physical theory fails to resolve with the invocation of "God's sensorium" by Newton in his search for a fundamental inertial reference system which his equations failed to produce, ie it is a metaphysical solution and hence to be rejected.

A criticism of this theory of measurement may begin from the common sense point of view; surely reference to human consciousness is not entailed by measurement, which might be recorded mechanically. Does von Neumann's account then imply the counter intuitive assertion that even if the record is not 'observed' for years, the reduction does not take place till then? This is exactly what von Neumann does seem to imply because he is taking the measuring instrument, including the recording device, to be susceptible to quantum mechanical treatment, and therefore to be in a superstate until the moment of conscious observation. One might argue, in the case of the cat, that the cat itself is conscious and therefore constitutes the first observer eg Walker (1970). But this still leaves the objectionable assumption that the electrocuting device, which is after all a macroscopic object, is in a superstate until the cat is, or is not, electrocuted. Further, if the cat is killed then it is no longer conscious and cannot constitute an observer in the above sense. Another case is argued by Margenau. He writes (1957 p367): "One might for instance be looking at a recording device and, while daydreaming, fail to take conscious cognizance of the registration. All physical processes, as the term is ordinarily interpreted, are the same as if he were taking conscious notice of the result, but in one case the state function develops continuously, in the other it changes abruptly ..."

This conclusion Margenau regards as absurd. However, if the wave function is to be given anything like a realistic physical interpretation, as it seems it must in order to account for the phenomena of interference, then, since it must change abruptly at some stage according to all the interpretations considered so far, wherever this abrupt change takes place the result is always going to seem absurd given the ordinary interpretation of physical processes. Consciousness, in so far as it is categorically different from physical processes (at least according to the Cartesian tradition) and yet necessary for the completion of any useful and informative measurement, seems indeed to be the least objectionable domain in which such a reduction may take place, for, were the reduction to take place at any stage in physical domain the assumption of the reducibility of physical ontology would be violated. Physical objects prior to the stage at which reduction of the wave function takes place must adhere to a quantum ontology whereas physical objects after the stage of reduction (determined by size or whatever) must adhere to a classical ontology. Ontology thus has the objectionable characteristic of not being universally applicable, which is a serious charge against such interpretations of QT. This objection cannot be maintained if reduction of the wave function is supposed to occur in the domain of consciousness, because in this case a quantum ontology can be universally applied to all physical objects. There however one exception to this, which leads us to a most important objection to the von Neumann theory of measurement.

What happens, asked Wigner (1961 pp284-302; 1971 pp16-17; 1973 pp380-382), if it is someone else besides oneself who makes the observation? According to the above interpretation one is forced to consider him as an apparatus which should be treated quantum mechanically by the ascription of a wave function. Then, in general, after he makes an observation, his state will be described as a linear combination of several observational results, ie a superstate. If he should now tell me the result of his observation, and assuming I believe him, then this is equivalent to an observation made by myself (he is acting as an intermediate mechanical recording device). At this moment the superstate collapses into one definite state according to von Neumann, London and Bauer. Hence one is forced to conclude from this argument that superstates do not collapse at just any consciousness but only at one's own consciousness, at oneself. Consider, for example, the cat experiment again. The experiment is performed and a 'friend' looks into the box containing the cat in order to determine the outcome. According to the von Neumann theory of measurement, I have now to describe the total experimental situation in terms of a product of four wave functions, one for each entity in the combined interaction. So the wave function for the system becomes the superstate

$$\sum_i \psi^i_{Electron} \otimes \psi^i_{Electrics} \otimes \psi^i_{Cat} \otimes \psi^i_{Friend}$$

No definite state can be ascribed to the cat - or even the friend - until he is asked what he saw or until I myself make the observation of the state of the cat. Thus, the wave function cannot said by me to have collapsed when another consciousness makes an observation. Only when I myself make the observation can the superstate be said to collapse as far as I am concerned. This is the only interpretation which is consistent with the reducibility assumption for all physical objects including brains. That is, only if the superpositions are said to collapse at one's own observations is one able to find a universalisable ontology for QT. One ontological consequence of this view should be spelt out. If you look at the cat then for you the cat actually is in a definite ontological state, either alive or dead. (Without this assumption (by me) this interpretation becomes solipsistic.) However, for me the cat, and now also you, is in a superstate, a linear combination of the states ‘cat alive' and 'cat dead’. Thus for us both the cat is in a different ontological state; for me a superposition and for you a definite state either alive or dead. Only when you communicate the result of your observation to me or when we both make the same observation is the cat in the same ontological state for us both. Generalising this result we see that according to this interpretation the worlds for you and I are ontologically different, except for those parts of the world which we share in common as a result of some common experience. In Aristotelian terminology this means that only the world which I myself observe (including those aspects of the world which I can infer without ambiguity from previous observations) is actualised while everything else is potentialised. What is actually the case for one need not be so for another. Let us call this the ‘egocentric interpretation’ of QT.

The egocentric interpretation has one definite attractive feature. According to the classical realist philosophy, physical reality would exist even if no observer existed. This has now come to appear as a common sense view where, at least the macroscopic world is thought to exist independently of our knowledge or observation of that world. Berkeley (1710), however, has argued that "all those bodies which compose the mighty frame of the world, have not any subsistence without a mind - that their being is to be perceived or known." Thus he argues that ontology and epistemology or what exists and what is known are not separate but closely related questions. Unless what exists can be clearly separated from what is known, it would seem to be impossible to maintain a realist philosophy, and it seems from the above discussion of probability and SP in QT that such a separation cannot be achieved. Now, we are accustomed to thinking of the world from a God-like omniscient, omnipresent point of view and, indeed Berkeley has attempted to preserve this view by arguing that the order in nature is created and maintained by God, who secures the reality of all things by His perception. By means of this argument Berkeley is able to retain the common sense position of the perseverance of the world independently of one’s own perception. But according to the egocentric interpretation omniscience is excluded by the theory since ignorance has become a necessary rather than contingent factor. Thus Berkeley’s concept of God, and in particular his related argument for the restoration of an objective reality, become insufficient. One is then left with a reality which is intimately dependent on one's own perceptions. The attractive feature of the egocentric interpretation is, then, that seems to correspond more nearly with the world as actually encountered for we do not in fact look at the world from 'above' as is suggested by realism; we look 'out' as is suggested by the egocentric interpretation. One might object however, that this does not explain how, having once admitted the equally valid experience of other consciousnesses, universal agreement is possible or why there appears to be only one single coherent world shared in common by all consciousnesses. Berkeley's argument to this effect fails

Wigner objects to this interpretation on the following grounds: Imagine, after the cat experiment has been completed by a friend that one then asks him, "What was the result of the exper1ment?" and he replies "The cat is dead." If one were to ask him, "What did you see before I asked you?" he would doubtless answer, “I told you already, I saw a dead cat." Accordingly, Wigner concludes that immediately after the interaction of friend and cat the state of the cat was already resolved and not in a superstate. If one does not make this assumption, writes Wigner (1961 p294), "it implies that my friend was in a state of suspended animation before he answered my question." So, continues Wigner, "it follows that the being with a consciousness must have a different role in quantum mechanics than the inanimate measuring device." It should be clear from the above discussion that if QT is to be accepted in its present form without alteration then this conclusion has the undesirable consequence of abandoning the reducibility assumption; that is the ontology of conscious beings’ brains is different from that of ordinary physical objects. In order to avoid this consequence, Wigner instead proposes what seems to be the only alternative, namely a modification of the SP in

QT. He writes (1971 p17): "There is perhaps no logical or mathematical contradiction here but the conclusion that the state of a 'friend' is a linear combination of several states, indicating different contents of his mind, seems very unnatural. This leads at least me to the opinion that quantum mechanics, in its present form, is not applicable to living systems, whose consciousness is a decisive characteristic."

In particular, Wigner argues, the SP ceases to be linear (and in fact becomes grossly non-linear) when QT is applied to physical processes relating to life and consciousness. Although it seems rather implausible to claim that the fundamental physics of brain processes is different from that of ordinary matter, there is nevertheless at least one good reason for calling into question the universal validity of the SP. This is due to the existence of what are known as superselection rules which forbid the application of the SP in certain physical situations. The fundamental theoretical reason for these rules, which have been added into QT for empirical reasons, is still unclear, but may well be due to some limitation to the validity of the linearity of the SP. It is however very doubtful whether there is any connection between superselection rules and brain processes as such.

Wigner's assertion that "the laws of atomic physics are drastically altered for those atoms that happen to make up the human brain" clearly shows that his interpretation rests on the assumption that the laws applying to human brains cannot be reduced to the laws of QT. Let us now discuss the general question of whether non-reducibility is physically objectionable. Mill (1879 Bk3) has given an example in favour of the 'doctrine of emergence' which is generally formulated as the thesis that it is possible for things or processes at 'higher levels of organisation' to have properties or obey laws which are not predictable from things or processes at 'lower levels of organisation'. In other words, new and unpredictable properties may emerge from a system of highly organised elements even though all the properties of the elements when they exist in isolation from one another are known completely. The example given by Mill is of mechanics and chemistry. The science of mechanics seems to be able to describe the motions and action of the simple material elements, but, writes Mill (1879 p432), "the different actions of a chemical compound will never be found to be the sums of actions of its separate parts." For this reason it is maintained that chemistry cannot be reduced to mechanics, or in other words, when the simple material elements which can be adequately described by mechanics attain a new level of organisation (namely that found in chemistry) then new properties of these highly organised collections of elements (compounds) can emerge which are not predictable from an understanding of the simple elements. Given the theories of mechanics of the19th century then this conclusion would indeed seem to be correct. However, as Nagel (1961 p365) indicates, "What was impossible relative to one theory need not be impossible relative to another physical theory. The reduction of various parts of chemistry to the QT of atomic structure now appears to be making steady if slow headway; and only the stupendous mathematical difficulties involved in making the relevant deductions from the quantum theoretical assumptions seem to stand in the way of carrying the task much further along."

Thus the word 'sum' as used by Mill above should be regarded as meaningful only relative to some assumed theory (in this case the theory of classical mechanics). Relative to another theory (perhaps QT) it may well be that the different actions of a chemical compound can indeed be found to be the sums of the actions of its separate parts. Thus, what were once regarded as emergent chemical properties which could not be explained in terms of classical mechanics might no longer be emergent given a new theory of mechanics. Notice that the parts or elementary constituents in both cases do not remain the same because the different theories will talk in terms of different elementary constituents, thus atoms in classical mechanics are very different from those in quantum mechanics.

Returning to Wigner, we see that a physically acceptable interpretation can be given to his claim that the laws of QT do not apply to matter in the brain. The level of organisation of matter in the brain is well known to be extremely high, perhaps the highest in the universe. This domain is therefore well suited to descriptions in terms of emergent properties and Wigner can then be taken as asserting that it is an emergent property of brain processes that they reduce superstates to individual states. In order to maintain the claim of 'ultimate reductionism’, that is the claim that ultimately the brain can be analysed into parts whose individual properties can predict all brain properties, it would be necessary to devise a new theory of physics more general than QT in terms of which this reduction may be achieved. It is sometimes supposed, with some justification but without proof, that QT is capable of achieving the reduction of the whole of chemistry to physics and modern chemistry of achieving the reduction of biology to chemistry. Were this the case then it should be possible to account for all brain processes in terms of quanta without the need to construct a new physics.

We are left with two rather weak objections to Wigner's suggestion. The first is that the view that brains do not obey the laws of QT, while a logical possibility, requires the invention of a new theory of physics if the reductionist programme is to be extended to the whole of, for example, neurophysiology. Such a theory has not yet been found. We should perhaps note, however, that all those interpretations of QT which rest on the von Neumann distinction between the two ways in which the wave function can change (and this includes in particular the CI) are really subject to a similar criticism because the second kind of change (reduction of the wave function) is an additional assumption which is outside the formalism of QT and these interpretations can therefore be regarded as interpretations of a new theory different from QT. Thus both these interpretations and Wigner's interpretation propose a modification of the original theory. The second objection is simply that it is unpleasant to imagine that the brain can have very large effects on reality as it must if reduction of the wave function is to take place at the brain.

Everett (1973) has agreed with the argument of ‘Wigner's friend’ to the extent that it represents a major difficulty for the interpretation of measurement by London and Bauer, but he does not agree with Wigner's proposed solution in terms of a modification of QT itself. Instead Everett points to his own interpretation which requires no modification of QT and which, I n fact, he regards as the only possible interpretation because he claims that the mathematical formalism defines its own interpretation. According to Everett, the superposition never collapses, and it is for this reason and only this reason that Everett's interpretation may be regarded as a natural interpretation of the formalism of QT, since here is no mechanism for a collapse of the superposition specified by that formalism. The reason why it was found necessary to postulate a collapse of the superposition in the first place was for the simple reason that when we become aware of the final outcome of a measurement we find ourselves to be in possession of one and only one definite result indicating the state of the observed world to be definite rather than a superposition. Had there been no collapse one would expect to observe a superstate, or perhaps, as Putnam (1965 p97) points out, to be oneself to be thrown into a superposition of states: 'my seeing a live cat' and 'my seeing a dead cat', in the case of the cat experiment. A clue to the interpretation of Everett (1957) is found in a comment by De Witt (1971 p226), "… reality is not the reality we customarily think of, but is a reality composed by many worlds." Thus, for example, in the cat experiment, the final outcome, according to Everett, is two distinct worlds, one of which contains a live cat together with me seeing the live cat, and another containing a dead cat together with me (or at least another me!) seeing the dead cat. In general then Everett claimed that a superstate contains a number (possibly infinite) of definite classical states each of which exists, but each in its own world. Looking at the world (or universe) as time progresses individual states interact and become superpositions of states; the world according to Everett, splits into a number of worlds, one corresponding to each of the possible states in the superposition. The world is therefore continually splitting, every possibility being realised in at least one of the resulting worlds.

Because of time symmetry of the Schrödinger equation, it must be possible for superstates to evolve continuously into individual states as well as the other way round. Everett therefore proposed that worlds described by exactly the same wave function recombine to give a single individual world. Thus the world is continually splitting and recombining, every possibility being realised in some world and every change in the state of any system being continuous and causal. In a postscript to Everett's original paper in 1957, Wheeler emphasised that although Everett's interpretation implies a fundamental revision of our traditional conception of reality, the interpretation is nevertheless logically self-consistent. This interpretation is, however, manifestly uneconomical and therefore in conflict with Ockham's principle of economy, for not only does it require one to postulate the existence of every possibility for the unobserved aspects of the world, as is required by the egocentric interpretation, but also the existence of every other possibility besides that which is actually observed consistent with all previous superpositions predicted by the time dependent Schrödinger equation.

One obvious objection to this interpretation which has been anticipated and answered by Everett is this: Why is it that we never experience such a splitting of the world but find ourselves to be in only one of the many worlds in which we are said to exist? This objection is answered by proving that we, to the extent that we can be regarded as automata and therefore on a par with measuring instruments, cannot observe such a splitting because of the laws of WT. What is proved is that if a second apparatus is introduced which both looks at the memory bank of the first apparatus and makes an independent direct check on the value of the observable of the macro-system under observation then these two observations must agree. It is then concluded that therefore there can be no information in the memory bank except that relating to that world in which the observer finds himself. In other words, no experiment in a given branch of the splitting universe can ever reveal the outcome of a measurement obtained in another branch of the universe. Everett is thus able to explain why we see only one universe and not many. But he is unable to explain why we find ourselves in this universe (that is, the one which we are presently experiencing) and not in any other. According to this interpretation there is no a priori difference theoretically between this and any other disconnected but existing world, but in practice of course we do indeed find a difference (one might say an absolute difference) between the world in which we find ourselves to be and all the other possible worlds. This empirical distinction between branches is not, and can never be, explained by the Everett interpretation of QT, since this is the fundamental content of the quantum mechanical notion of probability. If one could explain why we find ourselves to be in this and no other branch then determinism would be restored to physics and QT would have been superseded by a new theory. In order to answer this objection, therefore, Everett would have to produce a new theory of physics and not just a new interpretation of existing theory. Putting this another way, what is lacking in Everett's interpretation of QT is an explanation of why it is impossible to distinguish theoretically between branches when it is manifestly possible in practice since we must find ourselves in one or other world. This objection does not apply to the other interpretations where it is assumed that the world in which we find ourselves is the only existing actual world.

The next criticism of Everett's interpretation is more serious. Imagine a system described by the wave function $\psi(r)$ which is a superposition of functions $\psi_i(r)$, ie $\psi(r) = \sum_i \psi_i(r)$. This implies, according to Everett, that there exists a number of worlds (corresponding to the number of terms in the superposition) each described by one of the wave functions $\psi_i(r)$ in the superposition. If a measurement of the position, r, of the system is made then the outcome will be that we find ourselves to be in one of the worlds corresponding to one of the terms in the superposition, say $\psi_j(r)$. If, however, instead we wish to make a measurement of momentum rather than position of the system represented by $\psi(r)$ then we must transform the wave function from the position representation to the momentum representation by means of the rules in equations (11) to give a new wave function $\psi(p)$ which is again a superposition of functions $\psi(p) = \sum_i \psi_i(p)$. The trouble is that apart from the fact that each world of the many worlds proposed by Everett has now different characteristics to the extent that each is now described by a different wave function, the number of worlds proposed in either case is different. Thus we must accept that, according to this interpretation, the number of worlds which are assumed to exist depends critically on the measurement being performed, or on the observable in terms of which one is talking. It seems hard to believe that a small change in the design of an instrument can change the number of existing worlds.

The interpretation can also be criticised on the grounds that it cannot be disproved by any specific experiment since we cannot observe the split; the interpretation can therefore be regarded as metaphysical. D'Espagnat (1971a pp445-459) has distinguished between two different interpretations of 'splitting' within the Everett interpretation. The first, which he calls ‘solution (a)’, is that which we have taken for granted above where the number of systems etc is not constant but multiplies during a split. According to this solution, "in a measurementlike process that incorporates an observer, we can assume that all the branches have a reality of the type we are familiar with through our personal experience." If, before the process, only one observer exists then after the process there will exist a number of observers, one in each branch. This solves the measurement problem since the 'observed' reduction of the wave function is only apparent, no reduction takes place in fact. However, if the observer is replaced by an instrument then physical multiplication of instruments should take place in measurementlike processes although we cannot observe this multiplicity. Since it is assumed that the interaction between object and instrument is governed by the same laws as those governing any other interaction, similar multiplication should also take place in any interaction process in which the wave function evolves into a superposition. In particular, in the two slit experiment the electron after passage through the screen is described by a superposition $\psi_A + \psi_B$. But if each of these states is believed to exist independently in a separate world then no interference can occur between these two states which is experimentally false. It is therefore necessary to restrict these multiplication processes to measurementlike processes, but it is hard to understand how this can be done since measurementlike processes do not constitute a logically well-defined category, according to the Everett interpretation.

The second interpretation, ‘solution (b)', "consists in assuming that observers, instruments and so forth, instead of being multiplied, as was the case in solution (a), are simply split between the various branches." Thus, if one electron is passed through the screen in the two slit experiment then at any given time there is only one electron in existence but this electron is somehow split among the two slits. Moreover, this assumption is extended to macroscopic systems, instruments and observers. However, in this interpretation, if the impressions of observers are just physical properties of these systems (viz of the observers) then these impressions must be blurred in experiments involving superstates. The reason for this is as follows. Imagine that two systems U and V which are identical to one another in every respect, undergo a similar interaction which leaves their final states described by the superpositions $\sum_{i=1}^{n} \psi_i^U$ and $\sum_{i=1}^{n} \psi_i^V$, which are identical as regards all their physical attributes. According to Everett, there is a one-to-one correspondence between reality and the state functions of U and V, therefore two systems having the same state function must be identical. This applies whether systems U and V are quantum systems or whether they include macroscopic measuring instruments; however, if this is extended to systems U and V which incorporate a conscious observer then a difficulty appears. A conscious observer has the ‘faculty of introspection‘ and is therefore able to say in which of the n states in the superposition he actually finds himself to be. Thus one may say 'system U is in state i = 1' while the other may say 'system V is in state i = 2'. But if it is assumed that the state functions are in a one-to-one correspondence with the whole of reality, including the impressions of the observers, then in our example these impressions cannot differ from one observer to the next, for if they did, the two systems U and V would not be identical to one another in every respect. In other words, if impressions are included as physical attributes in the description of the state of a system then the impressions of observer U must be identical to those of V. Now since there is nothing to single out one particular value of i rather than any other because the individual states ψ_i appear in a symmetrical way in the superstate, the impressions of the observers must be blurred over all values of i. Since blurring of this sort does not occur, this interpretation can only be maintained if it is assumed that consciousness is a property of physical systems which is not described by the wave function and to this extent is different from all other properties. D'Espagnat (1971a p455) has at this stage pointed out an 'unquestionable resemblance' between this approach and epiphenomenalism, the theory that mental events reflect bodily changes but have no causal influence on the body. According to this theory, consciousness is different from all other phenomena since it involves no physical reactions. It would therefore not be surprising according to this theory if the wave function which correctly describes all other physical phenomena, does not describe consciousness.

Solution (b), while logically consistent, yields a conceptual difficulty which is rather hard to accept. Imagine a system S which evolves into a superstate and which is observed by two independent observers O(1) and O(2). The state of the system S + O(1) can be written as $\sum_{i=1}^{n} \psi_i^S \otimes \psi_i^{O(1)}$ while the state of the system S + O(2) can be written as $\sum_{i=1}^{n} \psi_i^S \otimes \psi_i^{O(2)}$. From the above discussion of solution (b) we see that the consciousness of each observer O(1) and O(2) is not split into a superposition of n states but rather is to be found in a state which corresponds to one of the n states in the superposition. Assume then that the consciousness of O(1) adopts a state corresponding to the state i = j and O(2) the state corresponding to i = k, ie $C^{O(1)} \rightarrow C_j^{O(1)}$ and $C^{O(2)} \rightarrow C_k^{O(2)}$. Remembering that the wave function is never reduced and that the state of consciousness of O(1) and O(eo2) may correspond to any value of i independently of the other, we see that O(1) and O(2) can now develop states of consciousness which are not matched to one another. They may correspond to different values of i, ie $j \neq k$. There is no reason why, in this interpretation, j should equal k, and therefore no reason why O(1) and O(2) should have the same, or even similar, experience when both observe the same object. "The idea that we should be able to take tea together," writes d'Espagnat (1971a p459), "while localising differently the same teapot is a logical possibility. It is one, however, that is somewhat hard to accept." No communication between O(1) and O(2) can ever reveal this difference in impressions because any transfer of information from, say, O(2) to O(1) takes the form of a measurement of O(1) on O(2) and since we have shown that O(1) can never observe a split by making any measurement he must get a response which is in agreement with his own perception although this apparent agreement need not necessarily be real.

III

Having considered a number of attempts to resolve the problem of measurement in QT, we now turn to an analysis of Bohr's arguments. In such an analysis one is immediately confronted with an extreme divergence of opinions of various commentators in their exegesis of Bohr. As we have seen, according to the von Neumann, London and Bauer interpretation it is necessary to introduce consciousness explicitly into the discussion of measurement. Does Bohr also require an explicit introduction of consciousness? According to Popper (1967), Hooker (1972) and Hesse (1974) the answer to this is given in the affirmative while according to Feyerabend (1968), Putnam (1965) and d'Espagnat (1971a) the answer is given in the negative. Thus, over this fundamental issue there is still basic disagreement. It is well known that Bohr has repeatedly emphasised the difficulty of distinguishing between subject and object in atomic physics (see 1958 p54; 1934). The problem centres round Bohr's interpretation of 'subject' in physics. According to Feyerabend, 'subject' for Bohr does not mean the consciousness of the observer but rather "the material measuring instrument (including the body of the observer, and his sense organs)" (1969 p92). Thus, Feyerabend argues that the blurring of the subject-object distinction referred to by Bohr does not refer to a disappearance of the boundary between consciousness of the observer and the observed world but its disappearance between atomic phenomena and material measuring instruments. Popper, on the other hand, writes (1967 p7); according to the CI "quantum mechanics does not represent particles, but rather our knowledge, our observations, or our consciousness, of particles." He therefore concludes that the 'subject' in physics according to the CI refers to the observer including "the ghost called consciousness". Popper, himself, does not accept this position however. He believes that the observer in QT plays exactly the same role as in classical physics and he therefore attempts to "exorcise the ghost ... from quantum mechanics."

The attempt to classify Bohr's concept of 'subject' into the category of the mental (consciousness, according to Popper) or the material (measuring instrument including the body of the observer, according to Feyerabend) is, according to Heisenberg (1958 p75) "the root of the difficulty" which "even eminent scientists like Einstein had in understanding and accepting the CI of QT." Heisenberg writes (1958 p73): "The influence of the Cartesian division (into ego and world) on human thought ... can hardly be overestimated, but it is just this division which we have to criticise ... from the development of physics in our time." Again he writes (1958 p75): "The Cartesian partition ... has penetrated deeply into the human mind during the three centuries following Descartes and it will take a long time for it to be replaced by a really different attitude toward the problem of reality." Also, Petersen (1968 Ch.4) writes: "In throwing new light on the subject-object relation, quantum mechanics has added evidence against the Cartesian division between mind and matter." Thus it is, according to Heisenberg, as a result of the rigid and inflexible distinction between subjectivity and objectivity that many attempted interpretations of QT have met with serious difficulty. For example, this rigorous distinction has been maintained by von Neumann, for whom QT is compatible with "the so-called principle of the psycho-physical parallelism - that it must be possible so to describe the extra-physical process of the subjective perception as if it were in reality in the physical world - ie, to assign to its parts equivalent processes in the objective environment, in physical space" (1955 pp418-9). But, according to von Neumann (1955 p420), notwithstanding the fact that the boundary between the two portions "can be pushed arbitrarily far into the interior of the actual observer" rigid division of the world into observed and observing portions is "the content of the principle of the psycho-physical parallelism", and results in the problem of the reduction of the superstate.

In contrast, Bohr's view is characterised by a certain flexibility or lack of precision concerning the distinction between subjectivity and objectivity. As a consequence of this flexibility the characteristic properties of micro-objects can no longer be regarded as being independent of the properties of the observing instrument. Bohr, therefore, insists that all discussion of quantum objects must make appropriate reference to the measuring instruments with which they interact. QT is, therefore, according to Bohr, a theory about quantum phenomena where the word 'phenomenon' is used to refer "exclusively to observations obtained under specified circumstances, including an account of the whole experimental arrangement" (1961 p64). Indeterminacy inherent in quantum processes can then be accounted for by Bohr in the following way by application of the concept of complementarity. Since experimental results and arrangements must be described in "everyday concepts, perhaps refined by the terminology of classical physics" (1961 p26), and since the experimental arrangements involved in the measurement of two complementary variables are mutually exclusive, the transition from a description of a quantum object in terms of one variable to a description in terms of the complementary variable involves a "rupture" in the system as a whole (ie the measuring instrument-object composite) in the transition which cannot be controlled. Thus it appears impossible to unite into a single picture the action of a quantum object. It is therefore also impossible to give a strictly causal account of quantum phenomena, since this would imply the existence of a description of quantum objects which did not depend essentially on the means used to observe them. "The notion of complementarity serves to symbolize the fundamental limitation, met with in atomic physics, of the objective existence of phenomena independent of the means of their observation" (Bohr 1961 p7). Hence Bohr regards complementarity as a generalisation of the classical concept of causality.

One might suppose that the above account of the introduction of complementarity into microphysics is sufficient to explain the problem of measurement which can now be described as the appearance of indeterminacy (or irreducible probability) at the macroscopic level. However, the above account of the appearance of indeterminacy rested on the rupture necessitated by the descriptions of complementary phenomena, whereas the problem of measurement concerns a single phenomenon which theoretically is supposed to be in a superstate. In order to solve this problem, Bohr seems to suggest that the concept of complementarity should be reapplied at a new level, the macroscopic level. The first indication of this generalised use of complementarity is to be found in the analysis of the two slit experiment where, in one set of circumstances, the screen was treated as a measuring instrument and therefore in a classically definite state, while in other circumstances it was treated as a quantum object whose momentum was being measured and which satisfied the uncertainty relations. He writes (1961 p50): "The main point here is the distinction between the objects under investigation and the measuring instruments which serve to define, in classical terms, the conditions under which the phenomena appear. Incidentally, we remark that, for the illustration of the preceding considerations, it is not relevant that experiments involving an accurate control of the momentum or energy transfer from atomic particles to heavy bodies like diaphragms and shutters would be very difficult to perform, if practicable at all. It is only decisive that, in contrast to the proper measuring instruments, these bodies together with the particles would constitute the system to which the quantum-mechanical formalism has to be applied."

Thus, from one point of view a macrophysical object is treated quantum mechanically (in which case it could be described as a superstate) while from another mutually exclusive point of view it is treated as a classical object which is in a definite well-defined state. This would seem to be an application of the concept of complementarity, which was originally formulated in the domain of microphysical objects, to the macroscopic domain, including measuring devices. This generalisation of the principle of complementarity has been expressed more explicitly by Bohm (1951 pp627-8): "The classically definite aspects of large-scale systems cannot be deduced from the quantum mechanical relationships of assumed small-scale elements. Instead, classical definiteness and quantum potentialities complement each other in providing a complete description of the system as a whole." (The senses of complementarity are, however, only loosely connected.) A macroscopic object, therefore, has objective existence and intrinsic properties in the experimental arrangement where it is being used as a measuring instrument and quantum potentialities (or properties relative to the observer) in the circumstance where it plays the part of the object being measured. Bohr has thus avoided committing himself to either realism or positivism (or idealism). The two views, realism and idealism, might be said to complement one another in a complete description of natural phenomena (see pp23-24). As a consequence of this generalised application of complementarity it emerges that a realist macrophysical ontology which has sometimes been attributed to Bohr (eg by Feyerabend), is more likely attributable to Heisenberg and only partially represents Bohr's point of view. Correspondingly, Bohr's occasional characterisation of atoms as real and having individuality (eg 1934 p93) should not be dismissed as is done by Feyerabend.

Since Bohr finds it necessary to extend the application of complementarity to the macroscopic domain in order to give an adequate account of measurement in QT, his extension of the concept to the fields of, for example, biology and psychology should, perhaps, be regarded as more than a hopeful extrapolation, although there is a weakness in this programs since nothing corresponding to the uncertainty relations which establish the predictive power of QT and support an interpretation of QT in terms of complementarity has been found in the domains of biology and psychology - but see Jung (1960 pp229-232 para 439-440). Nevertheless, according to Bohr, his general epistemological view characterised by the concept of complementarity is evidently applicable to all domains of phenomena. This general viewpoint has been described by Shimony (1963 p770) as follows: "Given any domain of phenomena there is a standpoint from which this domain can be observed in mutually exclusive ways." Examples given by Bohr include the description of an organism for which the display of life and the conformity to physical laws are observed in mutually exclusive arrangements. Similarly, determinism and free-will are complementary aspects of the behaviour of an organism. Also, behaviouristic and introspective psychological descriptions he regards as complementary. Even within the domain of introspective psychology itself Bohr finds the necessity for using a complementary mode of explanation. He writes (1961 p101): "... it must be emphasised that the distinction between subject and object, necessary for unambiguous description, is retained in the way that in every communication containing a reference to ourselves we, so-to-speak, introduce a new subject which does not appear as part of the content of the communication." In other words, we are able (and have to be able) to objectify ourselves for the purpose of unambiguous communication in which one makes reference to oneself, even though, from another mutually exclusive point of view, subject and object in this case cannot be separated.

In order finally to attempt a presentation of Bohr-'s solution to the Schrödinger cat paradox, let us analyse Bohr's further references to the relationship between subject and object, for this is vital to his view of the problem of measurement in QT. As an indication of the subtle interplay between observer and observed as conceptualised by the generalised principle of complementarity, Bohr writes (1961 p11): "Without entering into metaphysical speculations, I may perhaps add that an analysis of the very concept of explanation would, naturally, begin and end with a renunciation as to explaining our own conscious activity." In this field one is again confronted with sets of mutually exclusive descriptions appropriate in different circumstances; in order to analyse the concept of explanation one must forego any attempt to explain one's own mental activity. In an earlier publication Bohr makes this point more clearly. "For describing our mental activity," he writes (1934 p96), "we require, on one hand, an objectively given content to be placed in opposition to a perceiving subject, while, on the other hand, as is already implied in such an assertion, no sharp separation between object and subject can be maintained, since the perceiving subject also belongs to our mental content." That is, if our mental activity is to be treated as the object of scrutiny then it must be given objective status by being placed in opposition or contrast with an observing subject. However, since the concept of the observing subject is also part of the mental content, it would, from another point of view, appear impossible to separate clearly subject from object. Consider, for example, a curious thing about breathing: one says in the ordinary way I breathe because one feels that breathing is something one is doing voluntarily just as one might be walking or talking. But when one is not thinking about breathing, breathing goes on just the same. So a curious thing about breath is that it can be looked at both as a voluntary and an involuntary action. One can feel, on the one hand, I am doing it and, on the other hand, it is happening to me. Thus we see from this example that the hard and fast division we make between what we do and what happens to us is arbitrary. Breathing can be considered from one point of view as inseparable from our vo1untary actions and from another (mutually exclusive) point of view as independent of our voluntary actions.

Bohr continues: "From these circumstances follows not only the relative meaning of every concept, or rather of every word, the meaning depending upon our arbitrary choice of viewpoint, but also that we must, in general, be prepared to accept the fact that a complete elucidation of one and the same object may require diverse points of view which defy a unique description." Thus the object which at one time is considered as part of the mental content and therefore inseparable from the perceiving subject can also, on another mutually exclusive occasion, be regarded objectively as possessing properties independent of an observer. Both viewpoints, the first corresponding to idealism and the second to realism, are here seen as complementary, each applying on the appropriate occasion only, but both necessary for a complete description of the phenomena under consideration. The meaning of any concept is thus dependent on the context as a whole in so far as the meaning of the concept depends on the point of view from which it is analysed. Bohr therefore concludes (1934 p96): "Indeed, strictly speaking, the conscious analysis of any concept stands in a relation of exclusion to its immediate application." This follows since the conscious analysis of a concept requires the introduction of a perceiving subject from which it cannot be separated and this situation is complementary to the application of the concept where the concept is given objective content which is in no way relative to the observer. Two such descriptions are therefore mutually exclusive.

Consider, for example, some object O. According to Bohr, no single unique complete description can be given of this object except by means of a complementary relationship between two phenomenal appearances of the object; let us call them O^m when the object is seen as part of the mental content of some observer or perceiving subject S and O^P when the object is regarded as physically distinct from S. The object is then treated in either of two different ways depending on whether the context determines that we are dealing with O^m or O^p, the two characterisations being complementary. O^m makes its appearance through our human "forms of perception" and has to be understood through our usual "categories of understanding"; it is therefore described in classical terms as having definite (objective) characteristics in agreement with our usual well-defined observations. O^p, on the other hand, is described in the terms of QT which involves indefinite superstates and potentialities which only become well-defined through the intervention of S which effectively requires one to deal with the object O^m. O^m and O^p are thus mutually exclusive descriptions of the object O. For this reason, since mental and physical cannot be clearly distinguished but must instead be regarded as complementary, Heisenberg (1958) says that Cartesian dualism which reduces the kinds of existing thing to two basic substances, corresponding to mental and physical in the above sense, "accords very neatly with what the Copenhagen school calls 'complementarity'." Pauli has also suggested a similar conclusion. He writes (1952 p164): "The general problem of the relationship between mind and body, between the inward and the outward, cannot be said to have been solved by the concept of psycho-physical parallelism ... Modern science has perhaps brought us nearer to a more satisfactory understanding of this relationship, by introducing the concept of complementarity into physics itself. It would be the more satisfactory solution if mind and body could be interpreted as complementary aspects of the same reality." This contention of a complementarity between mind and body would indeed seem to be indicated by Bohr's assertion that no clear and ultimate distinction can be maintained between a subject (as a conscious observer) and a perceived object. The alleged complementarity has, however, been taken to its logical conclusion by a number of other writers, for example C A Meier (1935 p362) who writes: "Both sciences have, in the course of many years of independent work, amassed observations and systems of thought to match them. Both sciences have come up against certain barriers which ... display similar basic characteristics. The object to be investigated, and the human investigator with his organs of sense and knowledge and their extensions (measuring instruments and procedures), are indissolubly bound together. That is complementarity in physics as well as in psychology." Between physics and psychology there is in fact "a genuine and authentic relationship of complementarity." Thus, according to Meier, not only is the relationship of complementarity to be found within both the fields of physics and psychology (Jung (1948) considered the relationship between the conscious and unconscious mind to form a complementary pair of opposites in the domain of psychology; this view was also approved by Pauli in a footnote to Jung s essay), but there is also a complementarity between physics and psychology themselves. We have already begun to feel something of this relationship in the attempt by von Neumann to solve the problem of measurement by explicit introduction of "the extra-physical process of the subjective perception" into physics. However, in contrast to Bohr, von Neumann, London and Bauer still maintain an ultimate observer in any act of observation, viewing psychical and physical events as concomitant but disconnected.

Pauli (in the above mentioned note) explains in outline why physics and psychology can no longer be considered as separate in a complete description of natural phenomena: "It is undeniable that the development of 'microphysics' has brought the way in which nature is described in this science very much closer to that of the newer psychology: but whereas the former, on account of the basic 'complementarity' situation, is faced with the impossibility of eliminating the effects of the observer by determinable correctives, and has therefore to abandon in principle any objective understanding of physical phenomena, the latter can supplement the purely subjective psychology of consciousness by postulating the existence of an unconscious that possesses a large measure of objective reality." Thus, while in modern physics it is impossible to achieve an entirely objective description of microscopic entities because the characteristics of the observer cannot be eliminated due to the finite magnitude of the quantum of action, h, at the same time in modern psychology it has become possible to achieve a certain degree of objectivity by means of the postulate of the unconscious or the archetypes which are the "dynamic nuclei of the psyche" (Pauli (1961)) and which are determined as regards their form but not as regards their content. In this way it might be said that the descriptions of classical objective physics and of 'classical' subjective psychology do not contradict one another but complement one another since only together they offer a natural generalisation of the classical mode of description, each involves the other in a complete description of natural phenomena.

In our search for a solution to the problem of measurement in QT, which is exemplified by the Schrödinger cat paradox, let us adopt the particular version of the generalised principle of complementarity which has been explicated above and which is stated explicitly by Bergstein. He suggests (1972 Ch.6): "Consciousness and the external world are complementary concepts." If this were so then it becomes clear why Pauli believed that we should parallel our investigation of outer objects with a psychological investigation of the inner origin of our scientific concepts (see Pauli & Jung (1952 p165)) and also why Heisenberg (1958) believed that when examining nature and the universe, instead of looking for and finding objective qualities, "man encounters himself". If consciousness and the external world are complementary concepts then it follows that no complete description can be given of natural phenomena without due regard for the external world as well as the internal world of subjective perception. This generalised application of complementarity would seem to be in direct conflict with Einstein's view of science for he has said that "the belief in an external world independent of the perceiving subject is the basis of all natural science." However, this conflict is not as immediate as it might first appear since, according to Bohr's analysis, in any given situation there does exist a definite 'cut' or partition line between subject and object or between consciousness and the external world; the cut between subject and object is, however, movable according to the character of the analysis of experience, that is, the position of the cut can vary depending on the particular experimental arrangement under consideration. Concerning the positioning of the subject-object partition, Bohr (1961 pp78-9) says: "Such considerations point to the epistemological implications of the lesson regarding our observational position, which the development of physical science has impressed upon us. In return for the renunciation of accustomed demands on explanation, it offers a logical means of comprehending wider fields of experience, necessitating proper attention to the placing of the object-subject separation. Since, in philosophical literature, reference is sometimes made to different levels of objectivity or subjectivity or even of reality, it may be stressed that the notion of an ultimate subject as well as conceptions like realism and idealism find no place in objective description as we have defined it." Since quantum phenomena can, according to Bohr, only be unambiguously described by means of classical and everyday concepts and since whenever these concepts can be unambiguously applied it is possible to distinguish sharply between the observer and the objects observed (1961 p25), Bohr concludes that the distinction between subject and object is in fact a necessary condition for knowledge and communication. This separation, however, he believes can be performed in arbitrarily many different ways, and never in such a way that an absolute subject or an absolute object exists. Here Bohr's treatment of measurement can be seen to differ from that of von Neumann, London and Bauer all of whom acknowledge an ultimate subject in any act of observation.

In further elaboration of his attitude regarding the "change in the place where the discrimination is made between object and measuring agencies", Bohr writes (1935 p701): "in each experimental arrangement and measuring procedure we have only a free choice of this place within a region where the quantum-mechanical description of the process concerned is effectively equivalent with the classical description." So the arbitrariness as to the positioning of the 'cut' between subject and object is limited to a range in which quantum and classical descriptions are effectively equivalent. That it is necessary to introduce a cut somewhere between the measuring instruments and the objects under investigation forms, according to Bohr (ibid), a "principle distinction between classical and quantum-mechanical description of physical phenomena." This cut indicates a difference in the character of the description which does not feature in classical physics.

In order to clarify Bohr's position, consider again his treatment of the screen in the two slit experiment. In the first instance, when the screen is rigidly attached to its support (Fig.5), the screen can be regarded as a measuring instrument which accurately defines the position of an electron passing through it to be in one of two well-defined regions surrounding each slit. In this case the electron is taken to be the object under investigation while the screen is taken to be the measuring instrument. According to Bohr's analysis, therefore, the screen in this arrangement is to be described in the terms of classical physics and everyday language while the electron is described in the terms of quantum physics. The screen, then, in this case, plays the role of the observing instrument. The partition between observer and observed in this arrangement is then to be drawn between screen and electron. Consider what happens, however, in the second instance when the screen is detached from its support and allowed to move vertically in order to enable a measurement of the transfer of momentum from the electron to the screen and hence a determination of the slit through which the electron passes. In this case the screen plays the role of an object whose momentum is to be measured by means of some other measuring instrument introduced for the purpose. According to Bohr, the screen is no longer to be described in classical terms but in the terms of quantum mechanics, in particular the UP will now apply to the screen. This change in mode of description of the screen is seen as a result of the new positioning of the partition between observer and observed. The screen is now on the observed side of the partition which is in this case drawn between the screen and the new measuring instrument. Thus the position of the partition is dependent on the experimental arrangement, and even a macroscopic object may fall on either side of this partition depending on its role in the experimental arrangement as a whole.

Thus the version of the CI espoused by Putnam (1965 p93) and attributed by him to Bohr wherein "macro-observables retain sharp values at all times" is contrary to the essential content of the generalised principle of complementarity, and if accepted would lead to the problem of the reduction of the wave function. According to Hooker (1972 sec14), however, "for Bohr there exists no reduction of the wave packet - that belongs to the non-Bohrian von Neumann approach." It is true, nevertheless, that the CI as understood by most working physicists is roughly in agreement with Putnam's account. According to this understanding, macro and micro events must be clearly distinguished and macro events must retain their usual classical ontology. Thus in the Schrödinger cat experiment, the first macroscopic event (the triggering of the electrocuting device) must count as the measurement. Then, in agreement with the presumptions of common sense, the physicist assumes that he knows something about the cat without looking, namely that it is either alive or dead. This view therefore accepts that one can't reduce classical to quantum events although no clear explanation is given of why this should be impossible. Also, the line which must be drawn between superpositions of states in QT and classical descriptions is drawn once and for all between macro and micro events (although the distinction between these two is not clear-cut) and no variation in the positioning of the cut as a function of the prevailing experimental arrangement is permitted, in contrast with Bohr's account.

One can at this stage detect an unsatisfactory state of affairs as regards the antithesis between Bohr's point of view and the usual CI concerning the ontological status of macroscopic objects. The former demands a generalised complementarity applicable even at the macroscopic level which implies that no universally applicable ontological status can be assigned to macroscopic objects and which, if successful, goes far in supporting the application of complementarity to microscopic objects (which is maintained by all version of the CI). The latter, however, makes the assumption that all macroscopic objects can be described in the terms of everyday language and therefore always have well-defined properties in agreement with classical physics. So we have two entirely different descriptions of macroscopic objects and it seems that no experiment has been devised which can distinguish between these two views.

Why has no such experiment been devised? To answer this let us look a little closer at what this difference means in the case when there is in fact a difference (for the two views coincide in their description when the macroscopic object is of type O^m). The two views disagree when the object is being observed by a measuring instrument. In this case Bohr wishes to describe the object by means of a wave function ψ which may or may not be a superposition of states - let us assume that it is in a superposition of two macroscopically distinct states ψ_i and ψ_j such that $\psi = \psi_i + \psi_j$. According to the CI the object is always in a well-defined classical state, so, let us say it is in either state C_i or C_j (which correspond to ψ_i and ψ_j respectively). The only way to experimentally distinguish between these two descriptions (ie that in terms of classical physics and that in terms of quantum wave functions) is to observe interference which is possible in the quantum description but not in the classical description since ψ_i and ψ_j can be regarded as simultaneously propagating waves. But in order to obtain interference the two waves must be coherent or in phase, because if they were not any resulting interference would be constructive and destructive in equal proportions leaving no observable effect. We are, therefore, in such an experiment, looking for an effect which is dependent on the phase of the waves in the superposition ψ. Since the quantities which experimenters actually measure in macroscopic systems are gross, in the sense that considerable variation in the microscopic constitution of the system makes no discernible difference in their values, it is difficult to see how the phases of ψ_i and ψ_j can affect the outcome of an actual measurement. It is therefore difficult to see how one can distinguish experimentally between the above two radically different interpretations of measurement in QT.

However, although at the present time no such experiment has been devised, the prospect is not without hope because of the rapid development since about 1950 of a whole new field of physics, solid state physics, in which QT is applied directly to macroscopic objects having deep-going symmetry (eg crystalline solids). In this field, a few effects have been discovered, for example the Mossbauer effect, in which unexpected coherent contributions are obtained from the parts of macroscopic systems; the wave functions of all the atoms in a crystal maintain their coherence in such a way that to an electron scattered off the crystal the whole crystal appears as a single massive quantum. This is not to say that experiments can be performed which can detect interference effects in arbitrarily complex macroscopic objects; it does, however, indicate two cautionary conclusions. The first is that merely because, on certain occasions, the application of QT to macroscopic objects leads to the peculiar conclusion that the object is in a superposition of clearly distinct states (eg cat dead + cat alive) one should not be too hasty in drawing the conclusion that macroscopic objects are not described by QT. The successful application of QT to regularly shaped macroscopic solids indicates that QT is at least applicable to some macroscopic objects (those which are most readily amenable to mathematical analysis due to their symmetry). The second caution is against underestimating the ingenuity of experimenters in devising experiments concerning phase relations of macroscopic systems, especially since unexpected coherence has already been discovered.

Returning to our analysis of the relationship between subject and object we can now see that the fundamental distinction introduced by Bohr between instruments and quantum systems is not at all based on the macroscopic nature of the latter, but entirely on the use that we make of them for defining the conditions under which we observe. According to Bohr (1961), "proper measuring instruments ... serve to define, in classical terms, the conditions under which the phenomena appear." Bohm (1951 p627) was therefore in error (or at least in conflict with Bohr's point of view) when he wrote that "large-scale and small-scale properties are both needed to describe complementary aspects of a more fundamental indivisible unit, namely, the system as a whole", since it is not a question of size but rather of the use to which it is put which determines the type of properties, classical or quantum, which an object is described as having. Later, Bohm rewords this complementary relationship giving a version which does seem to be in agreement with Bohr. He writes (1951 pp627-8), "classical definiteness and quantum potentialities complement each other in providing a complete description of the system as a whole." The earlier disparity between Bohr and Bohm does seem however to indicate a fundamental misunderstanding on Bohm's part of this generalised complementarity introduced by Bohr. Bohm's position of macro-objectivism (ie all macroscopic objects are assumed to have well-defined properties) leads to the problems associated with the usual CI of measurement.

Let us consider why it is necessary to postulate a complementarity between classical and quantum descriptions. According to Bohm (1951 p624), classical concepts are characterised by three assumptions concerning the properties of matter:

1. The world can be analysed into distinct elements.

2. The state of each element can be described in terms of dynamical variables that are specifiable with arbitrarily high precision.

5. The interrelationship between parts of a system can be described with the aid of exact causal laws that define the changes of the above dynamical variables with time in terms of their initial values. The behaviour of the system as a whole can be regarded as the result of the interaction of all its parts.

We have seen, however, that in the quantum domain the properties of matter are to be associated with incompletely defined potentialities which can only be more definitely realised in interaction with a classically describable measuring instrument. The wholeness of a quantum phenomenon, moreover, implies that the properties of quantum systems can only be defined in the experimental context in which they appear when considered as a whole (postulate (b)). Thus we have contradicted classical assumptions 1 and 2. That 3 is also not satisfied is not surprising since exact causal laws would be meaningless where there existed no precisely defined variables to which these laws could apply. Nevertheless, despite this clear distinction between quantum and classical concepts, it is necessary, according to Bohr, to appeal to the classical level in order to obtain unambiguous communicable knowledge concerning quantum phenomena, since it is in terms of the concepts of this level that the results of all experiments on quantum systems must be described. Therefore, we conclude that "QT presupposes the classical level and the general correctness of classical concepts in describing this level; it does not deduce classical concepts as limiting cases of quantum concepts" (Bohm 1951 p625). Both the quantum and the classical levels are necessary for the complete elucidation of phenomena but one is not deducible from the other, they are instead to be regarded as complementary in accord with Bohr's "renunciation of accustomed demands on explanation" (1961 p78).

How does this complementary mode of explanation account for the paradox of Schrödinger's cat? The answer to this question centres round Bohr's attitude to the problem of the reduction of the wave packet. According to Bohr, wave-packet ‘reduction is not a physical process which would take place when an instrument comes into play or when a subject has a perception. Rather, it has already taken place at the time when we consider any definite experimental arrangement. In other words, it is ultimately determined by the actions we make in order to acquire knowledge as well as by the consciousness that we then have of purposefully performing such actions. Thus the experimental arrangement determines which aspects lie on the subject side of the subject-object partition, and thus which properties can be ‘described in the terms of the classical level. Therefore, writes Shimony (1971 p475), Bohr was able "to consider the demand for an explanation of the reduction of wave packets as illegitimate, for such an explanation would have to regard the observer and his instruments, which lie on the subject side of the subject-object separation, as objects for theoretical investigation." So, in the cat experiment, if the cat can be considered to lie on the subject side of the partition due to the particular experimental arrangement then it will be describable in classical terms (that is, it will be either dead or alive). No reduction of the wave packet is involved in this situation because of the complementarity (rather object than continuity) between classical and quantum levels, or between subject and object. Of course, if, in a different experimental arrangement, the cat is regarded as being on the object side of the partition then it has to be described in quantum terms as a superposition of cat alive and dead.

The most serious difficulty in Bohr's account is again (as in the case of wave-particle duality) one of finding a single ontological framework in which the subject-object separation can be seen as a natural event rather than as merely a mode of organising the content of experience. Bohr would appear to believe that the renunciation of an ontological framework is imposed by "the old truth that we are both onlookers and actors in the great drama of existence" (1934 p119; also 1948). According to Petersen (1963), Bohr, in an oral comment, said: "It is wrong to think that the task of physics is to find out how nature *is*. Physics concerns what we can say about nature." From this comment it would seem that Bohr did not consider physics to be concerned with ontology but instead wishes to shift the emphasis to epistemology as the most important field. Petersen who was Bohr's assistant for many years has, however, characterised epistemology as "the ontology of mind" (1968 Ch.I). The above complementarity between subject and object might then be regarded as a complementarity between the ontology of mind and the ontology of the world.

Chapter Three: Completeness and Reality

I

One of the fundamental problems immediately raised by QT concerns the question whether the theory is complete, in the sense of furnishing a complete description of the state of an individual object. Ihis question arises directly from the fact that the quantum theoretical state description is, in general, a probabilistic one. Since, classically, probability-free descriptions are assumed to be in principle an attainable ideal, the appearance of probability in QT is indicative of incompleteness in the quantum theoretical description. The first to react to the fundamental indeterminacy of QT in this way was Einstein, who, at the Solvay Conference of 1927, argued that a complete specification of the initial conditions of a process should enable one in principle to determine the course of the process, and therefore the statistical character of quantum mechanical predictions implies that the quantum description is incomplete (see Jammer 1966 pp357-9). Arguments of this type have been repeated recently by Belinfante (1975 p9) who writes: "As it is quantitatively meaningless to talk about probabilities for a single system, theories of this kind must be deterministic ... When people want a theory of elementary systems, and yet they claim that they want to maintain QT, they obviously are contradicting themselves."

In such arguments there is a clear prejudice in favour of the retention of strict causality in any complete theory. No compelling reason is given, however, why one should accept that nature is fundamentally deterministic. Einstein did in fact reveal his strong bias toward strict adherence to the principle of causality three years earlier in letters to Born and Ehrenfest. On April 29, 1924, Einstein wrote to Born: "Bohr's opinion of radiation interests me very much. But I don't want to let myself be driven to a renunciation of strict causality before there has been a much stronger resistance against it than up to now. I cannot bear the thought that an electron exposed to "a ray should by its own free decision choose the moment and the direction in which it wants to jump away. If so, I'd rather be a cobbler, or even an employee in a gambling-house than a physicist. It is true, my attempts to give the quanta palpable shape have failed again and again, but I'm not going to give up hope for a long time yet." Similarly, two days later on May 1, 1924, Einstein wrote to Ehrenfest giving as one of the reasons for rejecting Bohr's suggestions that "a final abandonment of strict causality is very hard for me to tolerate." Again, on December 4, 1926, Einstein wrote to Born saying that he was convinced that "God does not play dice." In other words, Einstein was saying that he did not believe that probability could play a fundamental role in physics. Such a belief is tantamount to an uncritical acceptance of causality as an indisputable regulative principle in science. Einstein's point of departure can thus be seen to be in conflict with Bohr's conclusion regarding causality in QT since Bohr (1931; Ch.2) claims that "the nature of QT forces us to regard the space-time coordination and the claim of causality, the union of which characterises the classical theories, as complementary but exclusive features of the description.." Thus, for Bohr (1948; 1949), "complementarity presents itself as a rational generalisation of the ideal of causality." Since this implies that a causal description cannot always be given of quantum processes, Einstein's point of view must remain clearly distinct from that of Bohr.

Perhaps for this reason Heisenberg (1955; see also 1959 p115) has characterised Einstein as belonging to a group which "expresses rather its general dissatisfaction with the QT, without making definite counter-proposals, either physical or philosophical in nature." However, as we shall see, this characterisation is unjust since Einstein, although never relinquishing his confidence in causality, has made numerous specific arguments both against the orthodox CI and in favour of his own statistical interpretation (SI). These arguments have had the effect of forcing a change in the CI (from a disturbance to a relational conception of quantum attributes) while the SI would appear to be gaining support even to the extent of sometimes being confused with the CI (see Belinfante 1975 p7 note 14). Nevertheless, Einstein's idealistic conception of physical determinism has been described by Rosenfeld (1963) as bordering on "mystification", and his view of QT is throughout correspondingly coloured; for example he was convinced that the two aspects of light involved in wave-particle duality can be causally related to one another.

Einstein was very much concerned with the question of complete causality and also sensed the threat to the principle of causality from the direction of QT a number of years before the appearance of the final statistical equations in 1925-6, as can be seen from a letter to Born dated 27 January, 1920, wherein Einstein remarked: “The question of causality worries me also a lot. Will the quantum absorption and emission of light ever be grasped in the sense of complete causality, or will there remain a statistical residue? I have to confess, that I lack the courage of a conviction. However, I should be very, very loath to abandon complete causality ..." The strength of Einstein's conviction is evident from his essentially unsuccessful search from 1911 until the end of his life in 1955 for a unified causal theory of all physical phenomena. In a letter to Schrödinger in 1950 Einstein continued to maintain that "the fundamentally statistical character of the theory (QT) is simply a consequence of the incompleteness of the description ..." - see Przibram (1967).

It should be noted that Einstein's original reaction to the MM (together with the UP) of Heisenberg and its interpretation by Born and Bohr was initially one of disbelief. In a letter to Schrödinger on 26 April, 1926, he wrote: "I am convinced that you have made a decisive advance with your formulation of the quantum condition, just as I am equally convinced that the Heisenberg-Born route is off the track." Presumably he was not aware at this time of Schrödinger's demonstration of the equivalence between the two formalisms since Schrödinger’s paper was not published until 4 May, 1926. But right up until the 1950 Solvay Congress Einstein continued to remain antagonistic towards the claim of logical consistency of the formalism of QT, constantly attempting to show its inconsistency by means of though experiments designed to obtain measurements of complementary quantities to an accuracy greater than that allowed by the UP. A thought-experiment of this kind was described by Einstein at the Solvay Congress (the ‘photon box experiment’) but, as has been clearly documented by Bohr (1949), after an almost sleepless night, on the next morning Bohr rebutted Einstein's challenge of inconsistency of the time-energy uncertainty relation by means of Einstein's own general theory of relativity. Further to this, Jammer (1974 p155) reports that: 'Einstein by introducing five-vectors into a four-dimensional space-time cherished the hope that a unified field theory would dispense with the Heisenberg indeterminacies, for they could then be regarded as merely projections onto a world of four-vectors, and their statistical implications could be regarded as the result of the suppression of the fifth component, which is necessary for a complete strictly deterministic description of five-dimensional physical processes."

The abortive nature of this attempted explanation together with the demonstrable failure of his thought-experiment to show the inadequacy of the Heisenberg- Bohr point of view led Einstein finally to admit the consistency of the uncertainty relations. Later he wrote (1949 p666): "the correctness of ... (Heisenberg's indeterminacy relation) is, from my own point of view, rightfully regarded as finally demonstrated." Instead of this approach, after 1950 Einstein changed his tactics directing his criticism to the incompleteness rather than the inconsistency of Bohr's approach. Einstein therefore continues

(1949 p666): "I am, in fact, firmly convinced that the essentially statistical character of contemporary QT is solely to be ascribed to the fact that this (theory) operates with an incomplete description of physical processes." That Einstein was willing, at least after much debate, to change his mind on matters of fundamental physics is to his credit but it does indicate something of the profundity of QT as well as Einstein's limitations, at least in his original conception of the theory. (QT seems also to have had a hand in influencing his philosophical standpoint away from positivism towards realism.)

Einstein's main concern in physics would seem to have always been directed towards his theory of relativity, possibly because, according to this theory, physical reality could be considered to be a four-dimensional space-time continuum in which all events are strictly determined in agreement with Einstein's beliefs concerning determinism. Indeed, according to Jammer (1974 p115), the first words spoken by Einstein at the Fifth Solvay Congress in 1927 were: "I have to apologise for not having gone deeply into quantum mechanics," which indicates again that he was primarily concerned with relativity. However, it would not be unreasonable, as a final introductory comment to Einstein's interpretation of QT, to characterise him as the 'grandfather' of QT for the following reason: Many of the decisive advances in QT were made as a reaction to a comment or paper by Einstein. For example, one can trace a connection between the development of Schrödinger's WM and the MM of Heisenberg-Born which are usually said to have originated in totally independent ways to the influence of Einstein. In 1905 Einstein published his celebrated paper on the photoelectric effect wherein he postulated that light itself is constituted of a beam of corpuscles or photons of energy hν and velocity c (= 3.10^{10}cms/sec) thus extending Planck's postulate which is limited to the introduction of a discontinuity in the absorption or emission mechanism. Certain phenomena, previously inexplicable in terms of the wave theory of light can then be explained on the basis of a corpuscular theory (eg the photoelectric effect and the Compton effect). Light thus can be seen to present itself in two forms: wave and corpuscle, each of these aspects appearing in more or less clear-cut fashion depending on the phenomenon under consideration. Thus Einstein's postulate stands at the basis of Bohr's wave-particle duality. Following in analogy with Einstein's lead, de Broglie then postulated a similar duality for matter as Einstein had proposed for radiation (this postulate was again quickly confirmed experimentally). Following from the generalised Einstein-de Broglie duality, Schrödinger was able to construct an equation which incorporated this duality. As regards the interpretation of Schrödinger's wave function (the solution to the above equation) in terms of probabilities, which is always credited to Born, it has been admitted by Born (1955) that the idea was fundamentally Einstein's (see also Jammer 19714 p41).

Concerning MM on the other hand, Heisenberg (1971 pp62-69) has described how, in conversation with Einstein, he was greatly impressed by a remark made by Einstein who said: "It is the theory which decides what we can observe." This statement was, said Heisenberg in a lecture delivered in Vienna on December 9, 1930, the point of departure for his reasoning which led him to the indeterminacy relations (see Jammer 1974 p57 and p76). Thus it is to a considerable extent as a result of the influence of Einstein that others under this influence have been able to develop the foundations of QT. However, despite Einstein's immense contribution to fundamentals of QT, he is nevertheless a self-named "renegade" since heterodoxically he "does not wish physics to be based on probabilities" (Born-Einstein Letters 1971 p163).

Let us begin our account of the SI by returning to the Schrödinger cat paradox. On 9 August, 1939, Einstein wrote to Schrödinger: "I am as convinced as ever that the wave representation of matter is an incomplete representation of the state of affairs ... The prettiest way to show this is by your example with the cat." Einstein is here agreeing with Schrödinger (1935a) who claimed that it would be "naive" to consider the wave function $\psi = 2^{½}(\psi_{cat\ dead} + \psi_{cat\ alive})$ as depicting reality since the cat must be in a definite state, either alive or dead, at all times. Thus they would seem to be claiming that any description which does not determine precisely in which state the cat will be found when observed is necessarily incomplete. This seems to be the view taken by a number of writers, eg Putnam (1965 p97); Popper (1967 pp9-10); Park (1968 p227); Ballentine (1972 p1768); Belinfante (1975 p xiii); etc, but the above argument is insufficient to substantiate the claim of incompleteness since it relies wholly on the unreserved acceptance of strict causality in principle. That the principle of causality is not necessarily universally applicable has been argued in the preceding two chapters which contained various probabilistic (ie not strictly causal) interpretations of the superstate ψ. It is therefore not adequate as an argument to claim that QT is incomplete simply because it is unable to determine (with probability equal to unity) the outcome of all events; either one must construct a new theory which is able precisely to determine all outcomes (in which case one can then hold onto the principle of causality and regard QT as incomplete), or else one must give a plausible reinterpretation of the wave function which enables one to interpret probability or uncertainty in QT as merely indicative of lack of knowledge. Einstein has pursued the latter weaker programme in an attempt to restore causality, the classical concept of physical reality and consequently (and perhaps most importantly for Einstein) what Einstein considers to be a "useful basis for the whole of physics" (1949 p666; see also 1936 p374 and p378; 1949 p683).

Einstein first expressed his own view of the interpretation of the wave function at the Fifth Solvay Congress in 1927 when he distinguished between two possible interpretations of the quantum state function:

I. According to the first viewpoint QT is regarded as a theory which completely described individual processes. This was the view of de Broglie, Born and Heisenberg in 1927 who said: "We maintain that quantum mechanics is a complete theory; its basic physical and mathematical hypotheses are not further susceptible of modifications." This was also the view of Bohr and of all varieties of the CI (see Ballentine 1970 p560 note 2). Einstein seems at this stage to have regarded the problem of the reduction of the wave packet as the greatest difficulty in this view since the localisation of a particle initially described by an unlocalised wave packet appears to involve the assumption of a peculiar 'action-at-a-distance' in violation of local causality, the assumption that causes and their effects transform the one into the other in a restricted region of space-time.

II. According to the second viewpoint the wave function is considered to represent not one individual particle but rather an ensemble (in the Gibbsian sense) of similarly prepared particles. Thus, for example, consider a particle which is ascribed a wave function: $\psi(x) = \delta(x - x')$. According to the first interpretation (I) this wave function represents a particle which has a definite position $x = x'$ but whose momentum p_x is therefore (as a result of the UP) completely unknown (this is the weakest version of CI). According to the second interpretation (II), however, $\psi(x)$ does not describe an individual particle. Instead $\psi(x)$ is the description of an essentially infinite number of particles each of which is regarded as having been subjected to the same initial conditions as the actual particle in the given experimental arrangement whose position is known to be x'. Each particle in this ensemble is then assumed to have a definite momentum but the range of momenta of the particles is $-\infty \leq p_x \leq +\infty$ as a consequence of the SI of the UP. When a momentum measurement is made on the actual particle under examination (as compared to the conceptual (infinite) set) then a definite value for the momentum will be obtained indicating in which sub-ensemble the real particle under examination belongs. In this interpretation there would be no difficulty in assuming that each particle has at all times a definite position and momentum since the UP then refers to the uncertainty as to which sub-ensemble of the infinite ensemble the real object or system belongs.

There is no doubt as to which of these two interpretations Einstein favoured. In (1949 p671) he wrote: "One arrives at very implausible theoretical conceptions, if one attempts to maintain the thesis that the statistical QT is in principle capable of producing a complete description of an individual physical system. On the other hand, those difficulties of theoretical interpretation disappear, if one views the quantum-mechanical description as the description of ensembles of systems." Thus Einstein is convinced that if one adopts interpretation II then all the difficulties in the interpretation of quanta disappear. This same conclusion he again emphasises (ibid): "The attempt to conceive the quantum-theoretical description as the complete description of the individual systems leads to unnatural theoretical interpretations, which become immediately unnecessary if one accepts the interpretation that the description refers to ensembles of systems and not to individual systems." Again we find that Einstein believes that all the unnatural elements in the interpretation of QT become immediately unnecessary if one accepts interpretation II. This view, then, according to Einstein (1949 p682), "blasts the framework of the ‘orthodox QT’." One important feature of the difference between interpretation II, the ensemble or statistical interpretation, and interpretation I, the CI in its many variations, is the difference between the two in their interpretation of probability. Heisenberg has compared probability in the CI to the Aristotelian notion of ‘potentiality’ while in the SI Einstein regards the probabilities in QT as relative frequencies of the results of an ensemble of identical experiments, a view which is probably more acceptable to the majority of physicists due, at least in part, to its greater familiarity.

It should now be clear how interpretation II solves the problem of measurement. Consider, for example, Schrödinger's cat: according to II, the wave function $\psi = \psi_{\text{cat alive}} + \psi_{\text{cat dead}}$ does not describe a single cat but rather an infinite ensemble of cats under similar circumstances. Thus the probabilities derived from the above wave function are not inherent irreducible descriptions of individual cats (hence involving, for example, the introduction of the Aristotelian concept of 'potentia'). Instead, these probabilities can be interpreted as relative frequencies in the same way as one interprets probabilities in games of dice.

Before going on to consider a detailed argument given by Einstein in favour of his own interpretation of QT, let us clear up a couple of difficulties. The first difficulty concerns an erroneous interpretation which could be given to Einstein's SI as a result of his terminology. He has, on certain occasions (see eg Einstein & Infeld 1938 pp299-502), replaced the word ‘ensemble’ by the word 'aggregate' or 'crowd' or 'congregation' or (see Feyerabend 1962 pp212-3) ‘collective’. Thus he says: "Quantum physics deals only with aggregations, and its laws are for crowds and not for individuals." This might be misleading in that it encourages the view that Einstein implies that what was once thought to be a single individual particle is now, in his view, to be considered as a composite made up of a large number of component particles. This is not at all what he intends. A second difficulty concerns the validity of the argument leading to the conclusion that QT is incomplete. If Einstein is arguing that QT is incomplete because it is a statistical theory then this argument is not necessarily valid since it depends on the assumption that it is meaningless to talk about probabilities for a single system. Sometimes Einstein appears to be arguing in this way, for example (1936 p374) he writes: "The incompleteness of the representation is the outcome of the statistical nature (incompleteness) of the laws." This, however, is not an argument but only the statement of an opinion. The justification of this opinion is given in the SI. It is because the state function according to this interpretation does not describe the single system that Einstein is able to conclude that QT is incomplete. The incompleteness or completeness of QT therefore rests not on whether the description given by QT is probabilistic (of this there is no question) but rather on the interpretation adopted for the state function, in particular whether this function refers to a single system or to an ensemble of systems. A final difficulty concerns a second meaning for completeness used by Einstein. Completeness in this sense means finality and logical consistency of the mathematical relations of QT. In this sense QT, according to Einstein, is complete. So he writes of QT (1949 pp666-7): "This theory is until now the only one which unites the corpuscular and undulatory dual character of matter in a logically satisfactory fashion; and the (testable) relations, which are contained in it, are, within the natural limits fixed by the indeterminacy-relation, complete." Hence, Einstein continues, the formal mathematical relations of QT "will probably have to be contained, in the form of logical inferences, in every useful future theory." The two sense of completeness must be clearly distinguished.

We now turn to an explicit argument by Einstein, Podolsky and Rosen (EPR) (1935) which is intended to prove that the quantum mechanical wave function does not provide a complete description of physical reality. It may be noted that, according to Rosenfeld (1967), the basic idea behind this argument was almost entirely due to Einstein who had expressed the essentials of the argument in a discussion in 1933. In support of Rosenfeld's claim, we might also refer to a letter from Einstein to Epstein, dated November 10, 1945, in which Einstein wrote concerning the EPR argument: "I myself arrived at these ideas starting from a simple thought-experiment." Rather than being based on the statement of any specific principle of causality, the argument is based on two criteria; one of physical reality (which is indirectly related to the requirement of strict causality), and the other of completeness. From the outset the authors make the explicit assumption that, for the purposes of the argument, the predictions of QT are in agreement with the results of experiment; that is, the theory of quantum mechanics is assumed to be correct. Hooker (1970; 1971; 1972 sec6) has pointed out that a number of critics have dismissed the EPR argument on the grounds that QT contains no contradictions, but this is not the purpose of EPR, which is a "physical and metaphysical critique not a formal critique." We shall therefore ignore the criticisms of these authors (eg Jauch, Kripps and Sharp).

EPR adopt, as a sufficient rather than a necessary, condition of reality:

(i) If, without in any way disturbing a system, we can predict with certainty (ie with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.

This criterion they claim to be "reasonable" and also "in agreement with classical as well as quantum-mechanical ideas of reality". With regard to the requirement for a complete theory, EPR propose the following necessary condition for completeness:

(ii) Every element of the physical reality must have a counterpart in the physical theory.

In other words, combining (i) and (ii), we find that as far as EPR are concerned, for every physical quantity which can be predicted with certainty, any complete theory must contain a counterpart for that quantity. This is indeed a reasonable hypothesis. However, it should be clearly understood that the transpose of this statement, which amounts to a statement of the principle of causality, does not in any way follow from the above two criteria, viz the counterpart for every physical quantity in a complete theory can be predicted with certainty. Seen in this light the above conditions may be regarded as a weakened version of Einstein's old demand for the preservation of causality in QT.

It is well known from the basic equations of QT that if a system can be characterised by means of a wave function ϕ(p) which is an eigenfunction of an operator P which corresponds to some observable quantity, say momentum, ie

$$P\,\phi(p) = p\,\phi(p) \qquad \ldots\ 13$$

then the value of the momentum of the system can be predicted with certainty, and is equal to p, the eigenvalue of the eigenequation 13. In this situation, given the condition of reality (i), it follows that the momentum of the system characterised by function ϕ(p) is real. However, it also follows from the foundations of QT that there exists another operator, X representing another physical observable quantity, position, which does not commute with the first operator, P, ie PX ≠ XP. It follows that ϕ is not an eigenfunction of the operator X, since if it were then X and P would obviously commute. In order to obtain predictions of the position of the system, the function must be Fourier transformed into a set of functions which are eigenfunctions of X, ie

$$\emptyset(p) = h^{-\frac{3}{2}} \int \Psi(x)\, e^{-i(p.x)/\hbar}\, dx \qquad \ldots\ 14$$

This yields the prediction that the system is to be found with equal probability over a whole range of positions. Clearly, in this case, the position of the system cannot be predicted with certainty and can only be determined by a direct measurement which will influence the system, changing its wave function due to the action of the process known as reduction of the wave packet. Thus it seems that no element of physical reality can be associated with the position of the system in this circumstance. EPR therefore write: "The usual conclusion from this in quantum mechanics is that when the momentum of a particle is known, its coordinate has no physical reality." Or more generally it can be said that two quantities cannot have simultaneous reality if they are represented by operators which do not commute. If, however, QT was admitted to be incomplete, then every element of physical reality need not necessarily have a counterpart in the theory (from (ii)) in which case the theory might not predict with certainty the value of the corresponding physical quantity. In this case, then, two quantities represented by non-commuting operators could still have simultaneous reality. EPR therefore conclude that either (1) the quantum-mechanical description of reality given by the wave function is not complete or; (2) when the operators corresponding to two physical quantities do not commute the two quantities cannot have simultaneous reality. Since the CI rejects (1), it seems forced to accept (2). For various refinements to this argument see eg Bohm (1951 Ch22 sec15), Ballentine (1970 p362 note 4), Hooker (1972 sec4). This concludes the first stage of the FER argument.

The second stage of the argument begins by assuming that the quantum-mechanical wave function does give a complete description of the physical reality. It is then shown that this assumption, together with the above criterion of reality, leads to a contradiction. In order to derive this contradiction, EPR consider two systems, I and II, which interact for a time from t = 0 to t = T, after which time it is assumed that there is no longer any physical interaction between I and II. Let us state this assumption of the mechanical isolation of the two systems after time T explicitly as a ‘locality assumption’ or ‘principle of separability of mechanically isolated systems' (see d'Espagnat 1971a pp113-4):

(iii) If "at the time of measurement ... two systems no longer interact, no real change can take place in the second system in consequence of anything that may be done to the first system" (EPR 1935 p779).

Clearly, this is just a statement of what is meant by the absence of an interaction between I and II if 'consequence' is taken to mean the result of a physical cause. Suppose that the states of the two individual systems prior to t = 0 are known, then the state of the combined system I + II can be calculated by means of the Schrödinger equation for any time t > 0 in particular t > T. Let us write the wave function for the combined system ψ^{I+II}. Consider a physical quantity A relating to system I which has eigenvalues a_l, a_2, a_3 ... and eigenfunctions $R_1^I(x)$, $R_2^I(x)$, $R_3^I(x)$ ... where x stands for the variables used to describe the system I. The wave function for the combined system can then be written as

$$\psi^{I+II} = \sum_{n=1}^{\infty} U_n^{II}(y)\, R_n^I(x) \qquad \ldots 15$$

where $U_n^{II}(y)$ are to be taken as the coefficients of the expansion of ψ^{I+II} into a series of orthogonal functions $R_n^I(x)$ and y stand for the variables used to describe the second system II. If, now, a measurement is made to determine the value of the quantity A for system I then it will be found that A has some value, say a_k. After the measurement, therefore, the first system will be left in state $R_k^I(x)$ while the second system will be left in the state given by $U_k^{II}(y)$. Thus the infinite series in equation 15 will be reduced to the single term

$$\psi^{I+II} = U_k^{II}(y)\, R_k^I(x) \qquad \ldots 16$$

as a result of the process of reduction of the wave packet. If instead of quantity A we had decided to measure quantity B having eigenvalues b_1, b_2, b_3 … and eigenfunctions $S_1^I(x)$, $S_2^I(x)$, $S_3^I(x)$ ... then 15 would be replaced by the expansion

$$\psi^{I+II} = \sum_{m=1}^{\infty} V_m^{II}(y)\, S_m^I(x) \qquad \ldots 17$$

where V_m^{II}'s are the new coefficients. If a measurement of quantity B yields the result b_j say, then in this case the first system will be left in state $S_j^I(x)$ and the second in state $V_j^{II}(y)$. That is, the wave function for the combined system will then be

$$\psi^{I+II} = V_j^{II}(y)\, S_j^I(x) \qquad \ldots 18$$

Comparing 16 and 18 we see that as a result of performing two different measurements on system I, the second system may be left in states with two different wave functions, $U_k^{II}(y)$ and $V_j^{II}(y)$. EPR now argue that by way of the locality assumption (iii) since system II has not in any way been disturbed measurements on system I, it therefore is possible to assign two different wave functions to the same reality, namely system II.

Now, if it can be shown that the two different wave function U_k and V_j can be eigenfunctions of two non-commuting operators then we shall have proved that two complementary variables can have simultaneous reality since either one or the other can be predicted with certainty without in any way disturbing the system II. This result would be in contradiction to the conclusion of stage I. In order to demonstrate that the two different wave functions U_k^{II} and V_j^{II} can be eigenfunctions of two non-commuting operators, EPR consider the example of two interacting particles taking position and momentum as the two complementary physical quantities. Instead of this example we shall consider the simpler example discussed by Bohm (1951 Ch22 sec16). Imagine a particle having total spin zero which at a certain time t_0 splits into two particles which then move in opposite directions. If the spin of one particle, I, is measured in any of the orthogonal directions x, y or z then, as a consequence of the conservation of total angular momentum, the spin of the second particle, II, in this direction can be predicted with certainty without disturbing II - it will be equal and opposite to the spin of I. This prediction by QT of a correlation between the pair of systems I and II has apparently been confirmed experimentally by Clauser and Holt in the case of the polarisation of photons (see Ch 4, I).

A similar situation in classical physics would present no problems since it could be assumed that at all times subsequent to t_0 system II actually has a definite spin in all directions x, y and z. The correlation between measurements of the spin of system I and those of system II is then regarded as a consequence of the initial interaction at t_0. In QT, however, the x, y and z spin operators do not commute, implying that all the spin components cannot have simultaneous reality according to the analysis in stage one using the criterion of reality (i) given above. This raises the question as to how particle II 'knows' (see Belinfante (1973 p12)) which spin component has been measured on the distant non-interacting system I (by assumption (iii)), as well as the result of this measurement in order that II can adopt an equal and opposite spin such that if its spin is measured one will with certainty obtain that result. If, prior to a measurement on I, II cannot be said to have definite values of all three spin components then one would tend to hypothesise some sort of action at a distance between I and II whereby, if a measurement is made of I's x-spin, a signal is sent from I to II causing II to have, at that moment, a corresponding definite value of x-spin. In the quantum-mechanical treatment of this experiment, however, the interaction term has been taken to be zero. Therefore any further interaction corresponding to the above suggested signal is outside the realm of QT and is therefore non-physical as far as any known physical force is concerned. The type of causality concerned in this interaction is called 'non-local' since, if it exists, according to QT it acts over indefinitely large distances without the transmission of energy of, any sort, as far as one can see at the present stage in the development of physics. While this sort of interaction would conflict with the locality assumption (iii), it is not logically impossible. It would, however, have the consequence that the measurement which one chooses to perform on one system in a particular region of space can have a causal influence on another remote system and this influence can persist over indefinite times (if no measurement is made on this second system) which implies that the system has some sort of a 'memory' of events which occur in another physically isolated system although this need not have observably awkward effects, since the memory is destroyed at every measurement.

Returning to the EPR argument we now see that the above example of spin correlation serves to show that it is possible to find an example in which the two different wave functions U_k and V_j are eigenfunctions of non-commuting operators, since, in the case where R_n^I corresponds to the x-spin wave function of particle I, U_n^{II} will be the x-spin wave function of particle II. If, alternatively, S_m^I is the y-spin eigenfunction of particle I then V_m^{II} will be the y-spin eigenfunction of particle II (by factorisation of the total wave function). Now, the x and y spins are represented by non-commuting operators and therefore U_k and V_j are eigenfunctions of non-commuting operators as required. By measuring either the x or y spin of particle I we are therefore in a position to predict with certainty, and without in any way disturbing particle II (assuming (iii)), either the x or y spin of II. As a result of (i) and the conclusion that U_k and V_j can be ascribed to the same reality, it follows that there exists an element of physical reality corresponding to both the x and y spins of II. Thus, assuming that the wave function does give a complete description of physical reality, we conclude that two physical quantities which are complementary can have simultaneous reality. It has been demonstrated in stage one of the argument that either (1) the quantum-mechanical description of reality given by the wave function is not complete or (2) when the operators corresponding to two physical quantities do not commute the two quantities cannot have simultaneous reality. Since we have shown in stage two that assuming (1) is false leads to the conclusion that (2) is false and since from stage one we have concluded that either (1) or (2) is true, we are forced to conclude, according to the EPR argumentation, that (1) is true, that is, the quantum-mechanical description of physical reality given by the wave function is not complete. QED.

Criticism of this argument has been made by Furry (1936) on the following grounds. Furry considers two possible interpretations of the correlation between systems I and II:

<u>Assumption</u> A. According to this interpretation when the systems I and II become macroscopically separated then the quantum theory of many-bodies breaks down. In particular, when the sub-systems I and II of the combined system I + II become widely separated then transitions occur according to the laws of probability such that the systems I and II then acquire well-defined states (ie reduction of the wave packet occurs spontaneously prior to measurement when the systems become widely separated). Correlation can then be viewed classically as a correlation between states which are already there in a well-defined form in each sub-system. That is, instead of remaining in a pure state described by the wave function ψ^{I+II} the state of the system I + II is described by a mixture.

<u>Assumption</u> B. According to this second interpretation, the combined system I + II is described as a pure state (rather than a mixture) by the wave function ψ^{I+II} in agreement with the formalism of ordinary QT.

Now, according to Furry, it was assumption A rather than assumption B which was adopted by EPR since he (Furry) interpreted criterion (i) as implying that when the value of a physical quantity can be predicted with certainty then this value belongs to the system whether or not a measurement has been, or can be, carried out. Also, it has been reported by Feyerabend (1962 p212) that "it was indeed Einstein's guess that the current formulation of the many-body problem in quantum mechanics might break down when particles are far enough apart." Following up this line of thought Furry analysed the experimental differences between the consequences of assumptions A and B and found that calculations using B lead to the appearance of an additional interference which is not present in the conclusions based on A. Furry thus showed by means of the EPR thought-experiment that A leads to a contradiction with the UP. He therefore concluded that "the assumption, a system when free from mechanical interference necessarily has independent real properties, is contradicted by quantum mechanics." As a result of Furry's analysis, experiments have been suggested (see Bohm and Aharonov (1957)) and have been performed by Wu and Shaknov (1950), which might be able (but see Ch4, I) to decide between the predictions based on assumption A and those based on assumption B. According to Bohm and Aharonov, the Wu and Shaknov experiment has disproved assumption A, the results being consistent with the predictions of QT and therefore assumption B.

We find, therefore, that not only is assumption A in contradiction with QT but it may also be ruled out by experiment. If, therefore, EPR did make use of this assumption then the argument would appear to be invalid. The question is, then, whether EPR have made use of assumption A. The answer to this would seem to be negative, for, although Einstein may well have believed the assumption to be correct, the EPR argument rests only on a correlation between the two sub-systems I and II and not on assumption A. Supporting this conclusion, Feyerabend (1962 p212) says that, after all, the EPR argument is not attempting to give a well-defined description of the states of the separated systems I and II in terms of quantum-mechanical wave functions, it is only attempting to show that a description in terms of the wave function of QT cannot be regarded as complete, given that QT is correct.

Consequently, Furry's analysis does not constitute a refutation of the EPR argument as he had intended. Since EPR do not claim that the Schrödinger equation of motion yields incorrect predictions of the final quantum state of system I + II (ie a pure state rather than a mixture) but rather that no quantum state, pure or mixed, can provide a complete description of an individual system, Furry's refutation is based on a misconception and is therefore "irrelevant" (Ballentine 1970 p370 note 13) to the EPR argument.

There is, however, as a result of the invalidity of Furry's assumption A, a clear difficulty for the ensemble SI which Einstein has advocated for QT. According to this interpretation, the EPR experiment is about an ensemble $E^{I,II}$ of pairs of systems I and II that interacted in the past. Consider, for example, systems whose spins are correlated as described above. A measurement of the x-spin of I will then enable one to predict with certainty the x-spin of II even if I and II are widely separated. This correlation can be explained if we assume that the pair on which the x-spin has been measured is a member of a sub-ensemble $E^x \subset E^{I,II}$ and in this sub-ensemble the x-spins of I and II are correlated in the appropriate way, ie measurements of the x-spins of I and II yield with certainty equal and opposite results. However, if we had chosen to measure the y-spin of the particular particle in question in I (and, in principle, this decision can be made - ie whether to measure the x or y spin component - even after the two systems I and II have ceased to interact) then a correlation will be found between the y-spins of I and II implying that this sub-ensemble $E^y \subset E^{I,II}$. The sub-ensembles E^x and E^y are different because pairs in E^x are assumed to have definite (correlated) x-spin components while pairs in E^y are assumed to have definite (correlated) y-spin components. But since assumption A is false, no system can have definite values of both x and y spin components and therefore no system can belong to both the E^x and the E^y sub-ensembles. Now, since no signal from I telling which measurement (x or y) had been made could reach the distant II in time before the measurement there in order to tell it to which ensemble it belonged, the observed correlations which are predicted by QT are seen to be paradoxical even in a classical ensemble interpretation. It therefore seems impossible to describe the predictions of QT by the methods of classical physics, in particular, we cannot understand the information which QT provides about the systems II in the ensembles E^x and E^y by regarding the individual systems II as carriers of this information. Concerning this strange result, that correlation between distant systems is determined only after a measurement has been made on one of the systems which cannot be explained in classical terms even by the SI, Schrödinger writes (1935b p556): "It is rather discomforting that the theory should allow a system to be steered or piloted into one or the other type of state at the experimenter's mercy in spite of his having no access to it." As far as Schrödinger was concerned, this result indicated not merely that QT was incomplete which was the conclusion of Einstein et al, but rather the manifestation of a serious flaw in the very foundations of the theory.

Pin-pointing the origin of this alleged flaw, Schrödinger (1935b sec 1) writes: "When two systems, of which we know the states by their respective representatives, enter into temporary physical interaction due to known forces between them, and when after a time of mutual influence the systems separate again, then they can no longer be described in the same way as before, viz by endowing each of them with a representative of its own. I would not call that one but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought." This aspect of QT has also been invoked by Sharp (1961 p229) in his attempt to refute the EPR argument when he writes: "strictly speaking, only the entire system has a state function; separate 'parts' of the system will not be representable by pure states." The point being made by Schrödinger is that after the two systems I and II have interacted they cannot be described each by a separate wave function ψ^I and ψ^{II} which give exhaustive descriptions of the two systems independently of one another. Instead, the combined system I + II must be described by a single wave function ψ^{I+II} even if I and II no longer interact. Schrödinger therefore continues: "The best possible knowledge of a whole does not necessarily include the best possible knowledge of all its parts, even though they may be entirely separated and therefore virtually capable of being 'best possibly known', ie of possessing, each of them, a representative of its own. The lack of knowledge is by no means due to the interaction being insufficiently known - at least not in the way that it could possibly be known more completely - it is due to the interaction itself." However, if now a measurement is made to determine the state of particle I then it becomes possible to infer the state of particle II without interfering with it. At this stage, then, one is able to 'disentangle' the wave function ψ^{I+II} into separate wave functions ψ^{I} and ψ^{II} which give the 'best possible knowledge' of each individual system. This disentanglement forms the basis of the EPR experiment and is regarded by Schrödinger as being of "sinister importance" for QT, indicating to Schrödinger deficiency rather than incompleteness in the theory.

In a recent paper, Baracca, Bohm, Hiley and Stuart (1975 p454) have further analysed this quantum mechanical feature of disentanglement. They argue that a two-particle system which is initially described by the entangled superstate ψ^{I+II}, after a measurement on one of the particles, becomes a product of two well-defined states ψ^{I} and ψ^{II} which they call 'local' states (since these states approach delta-functions in the appropriate phase space). Thus a measurement on one particle produces spontaneous localisation of the other particle by a process which seems to imply instantaneous action at a distance. Considering the implications of the usual QT they write: "Without interaction in any of the usual senses the second ... (particle II) seems to 'know' instantly in which direction its spin has to be well defined, and in which directions its spin has to be not well defined (after a measurement on I). Thus it can be said that the measurement must, in some sense, actually produce suitably localized states that are properly correlated from an initial state that is not localised in this way at all." So the authors conclude that "we have to understand nonlocality as intrinsic to quantum mechanics". Actually the paper is intended as support for a hidden variables theory which eliminates the need for explicit reference to measurement in this situation by proposing a process of spontaneous localisation prior to measurement in a way similar to that suggested by Furry (A) but avoiding the ensuing difficulties. However, non-locality is still involved in a fundamental way.

Attempting to give some credibility to the concept of non-local causality, which seems to be required by QT in order to account for the EPR experimental results, these authors introduce the concept of 'formal causality' as compared to the usual concept of dynamical causality. Formal causality they claim to be a development of Aristotle's original ideas concerning formal cause. According to this point of view, the quantum mechanical wave function gives a description of a form and not a dynamical sequence. As in the case of music there is no dynamical causal relationship determining the order of successive themes, an entire theme being a single whole form, so there is no dynamical cause of successive quantum states. Rather the whole order and form of the development is the cause. The spontaneous process of localisation of states is then regarded as a change of form which requires no dynamical causal explanation. One might, however, criticise this approach for the same reason as the introduction of the concept of potentiality into QT, since it would appear to explain away (rather than clarify) the meaning and role of probability or indeterminism in QT.

Although in the EPR paper extended discussion was not given to the point of view of the assumption (iii) of separability. Recognising that, according to Bohr, even if the partial systems (I and II) are spatially separated from each other at the particular time under consideration, the fact that the partial systems I and II form a total system which is described by the single wave function ψ^{I+II} implies that there is no reason as far as QT is concerned why any mutually independent existence (state of reality) should be ascribed to the partial systems I and II viewed separately. Therefore, to assert that the real situation of II could not be (directly) influenced by any measurement on I is unfounded from the point of view of QT and is also unacceptable since the EPR argument shows that such an influence exists according to QT. Following this argument, Einstein (1949 p682) arrives at a modification of the EPR conclusion by proposing a possible alternative conclusion. This has been formulated as a theorem by Ballentine (1970 p363):

Theorem 1.

The following two statements are incompatible:

(1) The state function provides a complete and exhaustive description of an individual system;

(2) The real physical states of spatially separated (non-interacting) objects are independent of each other.

The EPR conclusion is thus weakened, for one can still claim that QT provides a complete description of physical reality if one is willing to forego statement (2). This seems, according to Ballentine (ibid), to have been the position taken by Bohr while Einstein preferred to maintain (2) at the expense of (1). Thus, although Einstein is still able to maintain the locality assumption (iii), he now admits that it is logically possible to renounce this assumption and thence uphold the view that QT is complete.

We thus see that the conclusion that QT is an incomplete theory is not an inescapable conclusion as was first suggested in the EPR paper. Given, however, that QT might be incomplete, it is clear that Einstein et al believe that it should be possible to develop a new theory which is complete and which is thus able to give a complete description of physical reality. In the conclusion to their paper EPR therefore write (1935 p780): "While we have thus shown that the wave function does not provide a complete description of the physical reality, we left open the question of whether or not such a description exists. We believe, however, that such a theory is possible."

Let us now take three suppositions made by Einstein concerning the foundations of this proposed theory. First of all, he considers such a theory to be local for he writes (1949 p85): "But on one supposition we should, in my opinion, absolutely hold fast: the real factual situation of system II is independent of what is done with the system I, which is spatially separated from the former." The second assumption concerns the deterministic nature of the theory. It will be remembered that Einstein (1940 p666) said that he was "firmly convinced that the essentially statistical character of contemporary QT is solely to be ascribed to the fact that this (theory) operates with an incomplete description of physical systems." This indicates that as far as Einstein is concerned a theory which provides a complete description of physical systems will not be essentially statistical but will be deterministic. This theory will therefore not merely yield probabilistic predictions but will give contingent general predictions of the individual results of measurements ie the predictions are general in that they can be given exactly for any particular type of measurement and contingent in that they can be given for various possible settings of a particular apparatus and not only for the setting which is actually chosen by the experimenter. Finally, the predictions of the new theory must be compatible with the statistical predictions of QT. That this appears to be accepted by Einstein can be seen from his comment (1949 p667): "The formal relations which are given in ... QT - ie, its entire mathematical formalism - will probably have to be contained, in the form of logical inferences, in every useful future theory." Thus, Einstein, giving a classical interpretation to the concept of probability, holds that the probabilistic predictions are correct as far as they go, but that a future theory will predict exact results rather than probabilities. While all three assumptions are plausible, since these assumptions would be natural from a classical point of view, we shall now demonstrate as was first done by Bell (1964) that these three assumptions are mutually incompatible, that is, their combination leads to a contradiction.

Consider the example above where a particle initially of total spin zero splits into two spin ½ particles I and II (see Stapp 1971 p1306). The two particles are then analysed by means of two Stern-Gerlach devices A_1 and A_2 whose axes point in the directions of the vector $\underline{a}$ and $\underline{b}$, the cosine of the angle between them being equal to the scalar product $\underline{a}.\underline{b}$. Taking j as a label to denote the individual experiment then we shall define $n_{1j}(\underline{a},\underline{b})$ to be plus one or minus one depending on whether the new theory predicts that the particle from the j^{th} pair which passes through A_1 is deflected up or down given that the axes of A_1 and A_2 point in the directions $\underline{a}$ and $\underline{b}$ respectively. Similarly, numbers $n_{2j}(\underline{a},\underline{b})$ are defined for the particle passing through A_2. If, in this experiment, we allow A_1 and A_2 to be capable of being rotated into any of the directions $\underline{a}^1$ or $\underline{a}^2$ for A_1 and $\underline{b}^1$ or $\underline{b}^2$ for A_2 then the condition that the theory is able to give contingent general predictions of individual results means that for each pair j the numbers $n_{ij}(\underline{a},\underline{b})$, i= 1,2 are defined for all four combinations of the variables $\underline{a}$ and $\underline{b}$. Of course, in any given experiment only two of these eight numbers can actually be confirmed, the others corresponding to results which would have been obtained had the experiments with the alternative experimental arrangements been performed.

The assumption of local causality can now be incorporated since it implies that if A_1 and A_2 are widely separated then the deflection of the particle in the device A_1 does not depend on the orientation of the second device A_2 and vice versa because the settings of $\underline{a}$ and $\underline{b}$ can be made just before the particle arrives and therefore if $n_{1j}(\underline{a},\underline{b})$ was dependent on $\underline{b}$ or $n_{2j}(\underline{a},\underline{b})$ on $\underline{a}$ this would require the introduction of an instantaneous effect from a distant cause in contradiction to the requirement of local causality. Consequently:

$$n_{1j}(\underline{a}^1,\underline{b}^1) = n_{1j}(\underline{a}^1,\underline{b}^2) \equiv n_{1j}^1 \quad \ldots 19a$$

$$n_{1j}(\underline{a}^2,\underline{b}^1) = n_{1j}(\underline{a}^2,\underline{b}^2) \equiv n_{1j}^2 \quad \ldots 19b$$

$$n_{2j}(\underline{a}^1,\underline{b}^1) = n_{2j}(\underline{a}^2,\underline{b}^1) \equiv n_{2j}^1 \quad \ldots 19c$$

$$n_{2j}(\underline{a}^1,\underline{b}^2) = n_{2j}(\underline{a}^2,\underline{b}^2) \equiv n_{2j}^2 \quad \ldots 19d$$

According to QT the following relationship holds for the product of the spins measured by A_1 and A_2 with increasing accuracy as N increases:

$$(1/N) \Sigma_{j=1..N}\, n_{1j}(\underline{a}^1,\underline{b}^1)\, n_{1j}(\underline{a}^1,\underline{b}^1) = -\,\underline{a}.\underline{b} \quad \ldots 20$$

Now, if the predictions of the new complete theory are to agree with those of QT then equation 20 must be satisfied for all combinations of $\underline{a}$ and $\underline{b}$. Let us therefore choose the directions $\underline{a}^1$, $\underline{a}^2$, $\underline{b}^1$ and $\underline{b}^2$ such that:

$$\underline{a}^1.\underline{b}^1 = 1 \quad \ldots 21a$$

$$\underline{a}^1.\underline{b}^2 = 0 \quad \ldots 21b$$

$$\underline{a}^2.\underline{b}^1 = -1/2^{½} \quad \ldots 21c$$

$$\underline{a}^1.\underline{b}^1 = 1/2^{½} \qquad \ldots\ 21d$$

Following Stapp (1971 p1307), we shall now show that these assertions lead to a contradiction. Inserting 19 and 21 into 20, one obtains:

$$(1/N)\ \Sigma\ n_{1j}^{1}\ n_{2j}^{1} = -\underline{1} \qquad \ldots\ 22a$$

$$(1/N)\ \Sigma\ n_{1j}^{1}\ n_{2j}^{2} = 0 \qquad \ldots\ 22b$$

$$(1/N)\ \Sigma\ n_{1j}^{2}\ n_{2j}^{1} = 1/2^{½} \qquad \ldots\ 22c$$

$$(1/N)\ \Sigma\ n_{1j}^{2}\ n_{2j}^{2} = -1/2^{½} \qquad \ldots\ 22d$$

Remembering that $n_{ij}^{k} = \pm 1$, from 22a one concludes that

$$n_{1j}^{1} = -n_{2j}^{1} \qquad \ldots\ 23$$

Combining this with 22b gives

$$(1/N)\ \Sigma\ n_{2j}^{1}\ n_{2j}^{2} = 0 \qquad \ldots\ 24$$

Subtracting 22d from 22c we get

$$(1/N)\ \Sigma\ n_{1j}^{2}\ (n_{2j}^{1} - n_{2j}^{2}) = 2^{½} \qquad \ldots\ 25$$

Now since $n_{2j}^{2} = \pm 1$ then $n_{2j}^{2}\ n_{2j}^{2} = 1$ and we can rewrite 25 as

$$(1/N)\ \Sigma\ n_{1j}^{2}\ n_{2j}^{2}\ (n_{2j}^{2} n_{2j}^{2} - 1) = 2^{½} \qquad \ldots\ 26$$

Using the fact that the absolute value of a sum is less than or equal to the sum of the absolute values and the $|\ n_{1j}^{2}\ n_{2j}^{2}\ | = 1$ since $n_{ij}^{k} = \pm 1$, 26 becomes

$$2^{½} \leq (1/N)\ \Sigma\ |\ n_{1j}^{2}\ n_{2j}^{1} - 1|$$

$$\leq (1/N)\ \Sigma\ (1 - n_{1j}^{2}\ n_{2j}^{1})$$

$$\leq\ 1 - (1/N)\ \Sigma\ n_{1j}^{2}\ n_{2j}^{1}$$

$$\leq\ 1, \text{ from equation 24.}$$

Since $2^{½} = 1.414...$ the conclusion $2^{½} \leq 1$ is false. Thus the three assumptions of locality, causality and correctness of quantum mechanical predictions (even to within 5% since small variations of 5% in equations 22 cannot remove the contradiction) are mutually incompatible since their simultaneous affirmation leads to a contradiction. This result may be expressed in the form of a theorem which is known as Bell's theorem:

<u>Theorem</u> 2.

No theory can (a) give contingent general predictions of the individual results of measurements, (b) be compatible with the statistical predictions of QT (even to within say 5%) and (c) satisfy local causality.

Thus any new theory which attempts to reproduce the results of QT by means of the introduction of some sort of deterministic sub-structure is forced to abandon the principle of separability to which Einstein felt "we should... absolutely hold fast." Therefore Bell (1966 p452) writes that by means of this sort of theory (which was advocated by Einstein himself - see his 1935 p780; 1936 p377; 1949 p672) "the EPR paradox is resolved in the way which Einstein would have liked least." Further, combining; theorems 1 & 2, we see that, assuming the predictions of QT to be approximately correct (ie to within 5%), in order to be able to conclude from theorem 1 that QT is incomplete one must assert the principle of separability. However, from theorem 2 the assertion of this principle implies the abandonment of the principle of strict determinism. Therefore, in order to be able to argue that QT is incomplete, as Einstein desires, one must abandon strict determinism, which Einstein is "very, very loath" to do. We can now conclude that, given what Einstein regards as natural, it is not the case that "unnatural theoretical interpretations ... become immediately unnecessary" or that "those difficulties of theoretical interpretation disappear" even if one accepts that "the (quantum-mechanical) description refers to ensembles of systems and not to individual systems." Einstein's metaphysical assumptions are simply not compatible with the results of QT, even in his own SI.

II

At the end of Einstein's (l949) comments concerning the interpretation of the EPR experiment we find a short passage which at first sight appears quite unconnected with the problems of QT under discussion. He writes (1949 p685): "I close these expositions ... concerning the interpretation of QT with the reproduction of a brief conversation which I had with an important theoretical physicist. He: ‘I am inclined to believe in telepathy.’ I: ‘This has probably more to do with physics than with psychology.’ He: ‘Yes.’ - " No further clue is given to indicate the relevance of this passage and we can therefore only guess. The word 'telepathy' was coined in 1882 by Myers who defined it (1907 p xvii) as "the communication of impressions of any kind from one mind to another, independently of the recognised channels of sense." As an example of this type of phenomenon, let us consider the case of Commander Aylesbury quoted by Gurney, Myers and Podmore (1886 Vol.II p227): "The writer, when thirteen years of age, was capsized in a boat ‘W when landing on the island of Bally, east of Java, and was nearly drowned. On coming to the surface, after being repeatedly submerged, the boy called his mother. This amused the boat's crew, who spoke of it afterwards, and jeered him a good deal about it. Months after, on arrival in England, the boy went to his home, and while telling his mother of his narrow escape, he said, ‘While I was under water, I saw you all sitting in this room; you were working something white. I saw you all - Mother, Emily, Eliza, and Ellen.’ His mother at once said, 'Why yes, and I heard you cry for me, and I sent Emily to look out of the window, for I remarked that something had happened to that poor boy.’ The time, owing to the difference of East longitude, corresponded with the time when the voice was heard."

Further details of corroboration are given by Aylesbury and one of his sisters. This case may be taken as typical of those cited as evidence for telepathic communication since it contains the essential ingredients of lack of physical interaction yet of simultaneous experiences of two (typically although not necessarily distant) parties which are meaningfully correlated. Statistically significant semi-repeatable laboratory experiments have been pioneered by J B Rhine and S G Soal using 'Zener card' guessing techniques. If the results of these experiments are accepted then possibly one of the most interesting and significant discoveries associated with this work is that the distance between the agent and the percipient does not appear to affect the telepathic powers (as well as the other powers associated with the so called psi-faculty). This fact alone would appear to rule out all known physical modes of interaction, for example radio waves (see Sinclair (1930) - foreword by Einstein) which have in fact been ruled out on other experimental grounds, since all these modes involve attenuation with distance (but see Wassermann 1956 p53).

Why then, should it be Einstein's opinion that telepathy is probably explicable in terms of physics rather than psychology when all forms of physical interaction presently known can be dismissed? If no physical interaction can be postulated between Aylesbury and his mother at the time of the occurrence of the telepathic phenomenon (and they were, after all, thousands of miles apart with no obvious means of communication) then it appears that we are dealing with "the simultaneous occurrence of two meaningfully but not causally connected events." This is how Jung (1960 p441) working in collaboration with Pauli (who might easily have been the "important theoretical physicist" alluded to by Einstein) defined his concept of "synchronicity" which he regarded as an "acausal connecting principle", and which he uses as a basis for the explanation of telepathic phenomena. This concept is not necessarily in conflict with relativity since it is information rather than energy which is connected acausally. Kammerer (1919) had also made use of a similarly defined acausal principle which he called "seriality". It is, perhaps, significant to find that Einstein was favourably disposed to Kammerer's work which, according to Przibram (1926), he called "original and by no means absurd."

The close connection between synchronistic events, or meaningful coincidences, and the correlated events predicted by QT in the EPR experiment should now become clear. In the EPR experiment we are confronted with two events, for example the x-spins of I and II, which are not related by any (known) physical cause and yet which display a meaningful correlation (ie the x-spin of I is equal and opposite to the x-spin of II). Since no causal explanation can be given of this correlation, the result may be regarded as a coincidence of the type exemplifying the principle of synchronicity. Similarly, synchronicity can be employed as an explanatory principle in accounting for the acausal connection common to telepathic phenomena, for example the connection between Aylesbury's cry and his mother's 'simultaneous' hearing. The following objection might now be made concerning this postulated identity between the type of correlation found in the EPR situation and that in telepathic cases: The EPR correlation comes as a result of an initial interaction between I and II after which and prior to measurement on either I or II the systems I and II have to be described in terms of a single wave function ψ^{I+II}. It is this composite wave function which accounts for the correlations found in QT. Since there would appear to be no initial interaction between Aylesbury (A) and his mother (M) giving rise to a composite wave function ψ^{A+M} it would seem that it is impossible to account for the observed correlation between A and M in the same way as that found in QT. It should be remembered, however, that the correlation between I and II will hold for an unlimited period of time after the initial interaction as long as I and II do not become 'spontaneously localized' by means of some external interaction on either one or the other. Therefore, as long as there has been an interaction in the past, however long ago, and as long as those interacting elements remain 'unmeasured' or 'unobserved' (or, following the von Neumann interpretation to its logical conclusion, unconscious) then the elements will still be described by a composite wave function - which enables non-local 'communication' to occur as soon as either of the elements in the composite are observed.

The purpose of this account, however, is not to develop a specific theory of telepathy in terms of QT. (Specific theories of this type have, in fact, been proposed recently eg by O Costa De Beauregard in a lecture entitled 'The EPR Paradox' delivered on 17 April 1976 at MIT and by W von Lucadou & K Kornwachs in a lecture entitled 'Can QT explain Paranormal Phenomena?' delivered at the 19th Annual Convention of the Parapsychological Association, Utrecht, on 18 August 1976.) It is rather to show that it is possible by means of an acausal principle such as that of synchronicity to demonstrate some similarity between the novelties required in the resolution of the difficulties to be found in the interpretation of QT and those required to account for the H observed coincidences met with in the examination of psi-phenomena, in particular telepathy. If this has been done then it is hoped that the relevance of Einstein's brief conversation will have been to some extent clarified. (For some further comments on the relationship between QT and parapsychology, see Note).

III

Let us now turn to see what bearing the EPR argument has on the points of view of some various representatives of the Copenhagen School. Consider first Heisenberg's (early) version of the CI, and in particular his theory of measurement. According to Heisenberg, if a measurement is performed to determine the value of some physical quantity A then this measurement will cause a disturbance of the system under consideration which is such that the value of some other complementary quantity B becomes entirely unknown as a result of this disturbance. Since this disturbance is in principle unavoidable and uncontrollable, measurements in QT are "theoretically opaque" in the sense that the simultaneous assignation of precise values to the complementary quantities A and B would imply an "operational contradiction" (see Bub 1973 p6). This claim is supported by the arguments from, for example, Heisenberg's γ-microscope, which is designed to show that any procedure for measuring position involves an interference with the system in such a way that the value of the complementary momentum quantity is changed in an indeterminate way. This theory of measurement is the one to be found most commonly in textbooks on physics. For example, Dirac (1930 p3) writes: "there is a limit to the fineness of our powers of observation and the smallness of the accompanying disturbance - a limit which is inherent in the nature of things, and can never be surpassed by improved technique ..." And in a more recent book Matthews (1968 p17) writes: "The quantum physicist thus appears, if not as a bull in a china shop, at least as a man with his eyes shut, liable to knock down anything he touches, trying to obtain a clear picture of the delicate objects which surround him." Although of appealing simplicity, we shall now show (following Bub 1973) that this disturbance theory of measurement is refuted by the EPR argument.

According to Heisenberg's interpretation, QT is both complete and irreducibly statistical. The reason for the statistical nature of the theory, ie the absence of dispersion-free statistical operators, is taken to be a result of the theoretical opacity of measurement disturbances which have the effect of making it physically impossible to determine the simultaneous values of complementary quantities A and B since a measurement of A leads to an unavoidable, uncontrollable and unpredictable disturbance of the value of the quantity B. The probability assignments represented by the pure statistical operators therefore supply the maximum possible simultaneous knowledge of the micro-level consistent with the opacity of measurement disturbances. Now, if QT is assumed to be complete then it follows that no two complementary quantities of a given system can be assigned simultaneous precise values as a result of Heisenberg's measurement theory. However, given the EPR criterion of reality which is consistent with the ontological status of micro-objects on Heisenberg's version of the CI, EPR have shown that in certain cases the conjunction of the completeness assumption with their criterion of reality leads to the conclusion that certain pairs of complementary quantities can be ascribed simultaneous precise values, since it is shown that no disturbance is necessarily involved in the prediction of the precise values of either of these quantities. Thus Bub (1973 p15) concludes that "the completeness assumption of Heisenberg's interpretation is inconsistent with a criterion of reality that is consistent with Heisenberg‘s interpretation. In other words, Heisenberg's interpretation turns quantum mechanics into an inconsistent theory." Taking again the example of spin correlations between systems I and II and remembering that the x and y spins of both systems are complementary quantities, we see that the fact that both the x and y spins of the system I are not empirically precisely determinable does not imply, according to EPR, that the value of the x spin of II cannot be determined without altering the value of the y spin of II or vice versa. Since this implication is necessarily involved in Heisenberg's disturbance theory, we conclude that the EPR argument can be regarded as a refutation of Heisenberg's version of the CI.

At length we come to the interpretation which, of those that consider QT to be complete, according to Einstein (1949 p681), "seems … to come nearest to doing justice to the problem." This is the interpretation due to Bohr. Less than two months after the appearance of the EPR paper, Bohr wrote a paper of the same title in which he argued that "the argumentation of the mentioned authors does not justify their conclusion that quantum-mechanical description is essentially incomplete" (1935 p700). Concerning Bohr's view, Frank has said (1936; English translation 1949) "I believe that, as a starting point for a correct formulation of the complementarity idea, one must retain as exactly as possible the formulation set forth by Bohr in his reply in 1935 to Einstein's objection against the present QT." We therefore follow closely Bohr's line of reasoning.

Bohr's analysis centres round what he considers to be an "essential ambiguity" in the EPR criterion of reality (i). However "reasonable" (EPR (1935 p777) one might feel this criterion to be, it is interesting to find (Jammer 1974 pp244-7) that Friedberg (unpublished 1969) has shown it is possible to reformulate (i) in such a way that the reformulated criterion (i') is incompatible with the statistical predictions of QT. The reformulated criterion which is only slightly stronger than that of EPR since it implies that, in certain situations such as the EPR experimental situation, theoretical correlation is sufficient to establish the existence of elements of physical reality (ie no experiment need be performed to give the definite predictions), can be written:

(i') If the result of a measurement on a system I which (ie the measurement) in no way disturbs a system II certainly agrees with the result of a measurement on II which in no way disturbs I, then the result is part of reality - even if neither measurement is performed.

Jammer (1974 p247) therefore concludes that since this criterion of reality can be shown to be in conflict with the results of QT "it is known that quantum mechanics itself excludes certain almost self-evident criteria of reality." Thus 'reasonableness' cannot be regarded as sufficient grounds for the adoption of a criterion of reality in QT. This caution has indeed been indicated both by EPR (1935 p777) and by Bohr (1935 p696), the latter of whom writes: "The extent to which an unambiguous meaning can be attributed to such an expression as ‘physical reality’ cannot of course be deduced from a priori philosophical conceptions, but - as the authors of the article cited themselves emphasise - must be founded on a direct appeal to experiments and measurements." The "essential ambiguity" which Bohr detects in the EPR criterion of reality (i) is to be found in the meaning of the expression "without in any way disturbing a system" which is an essential part of (i). Although it is this expression which is fatal for the Heisenberg view just considered, Bohr's criticisms with regard to the meaning of this expression should in no way be considered as an attempt to recover the Heisenberg disturbance point of view for he states explicitly that in the EPR experimental situation there is "no question of a mechanical disturbance of the system under investigation during the last critical stage of the measuring procedure" (1935 p700) the like of which is demanded by the Heisenberg CI.

The ambiguity is not to do with a mechanical disturbance of the system but rather concerns the question of "an influence on the very conditions which define the possible types of predictions regarding the future behaviour of the system" (1935 p700). In other words, while the system is not disturbed mechanically, there is an influence concerning predictions relating to the system. Without consideration of this influence, (i) becomes ambiguous when applied to quantum phenomena since, according to Bohr (ibid), the above-mentioned conditions "constitute an inherent element of the description of any phenomenon to which the term 'physical reality' can be properly attached." These 'conditions' will now be investigated with the help of a specific experimental arrangement proposed by Bohr. As regards these conditions, it is essential to recognise that for Bohr they include not only the correlation effects between two previously interacting systems I and II but also the particular experiment which one chooses to make on one of the systems since, on his view, no phenomenon exists until a measurement has been made. On the other hand, of these two conditions, EPR use the first and ignore the second (see Scheibe 1973 p194).

As far as Bohr is concerned there is nothing exceptional about the physical situation described by EPR for he showed that their deductions may be considered as "an immediate consequence of the transformation theorems of quantum mechanics" (p696 note 1). Consider a system consisting of two partial systems I and II which are described by the two pairs of canonically conjugate variables (q_1, p_1), (q_2,p_2) pertaining to I and II respectively and satisfying the usual commutation rules:

$$[q_i,p_j] = ih/2\pi\ \delta_{ij}$$

where $\delta_{ij} = 1$ if i = j and zero otherwise. According to the transformation theorems of QT, (q_1,p_1), (q_2,p_2) can be replaced by two new pairs of conjugate variables (Q_1,P_1), (Q_2,P_2) which are related to the first variables by a simple orthogonal transformation, corresponding to a rotation of angle α in the planes (q_1,q_2), (p_1,p_2) giving:

$$Q_1 = q_1 \cos\alpha + q_2 \sin\alpha \qquad P_1 = p_1 \cos\alpha + p_2 \sin\alpha \qquad \ldots 27$$

$$Q_2 = -q_1 \sin\alpha + q_2 \cos\alpha \qquad P_2 = -p_1 \sin\alpha + p_2 \cos\alpha$$

where the new pairs of variables satisfy:

$$[Q_i,P_j] = ih/2\pi\ \delta_{ij}$$

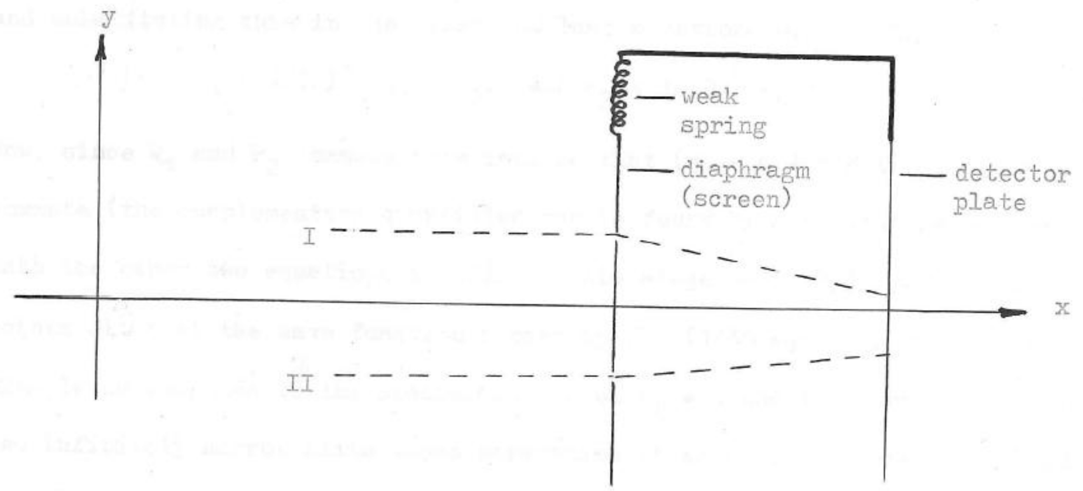
In order to demonstrate the relevance of the relations 27 to the EPR argument, Bohr considered "a simple experimental arrangement, comprising a rigid diaphragm with two parallel slits, which are very narrow compared with their separation, and through each of which one particle with given initial momentum passes independently of the other" (1935 p699). Now, it is clear that given this experimental arrangement, the difference in the initial positions of the two particles emerging from the diaphragm is equal to the difference between the two slits which is constant and can be accurately measured. Further, if the momentum of the diaphragm is accurately measured before and after the passage of the two particles in the direction perpendicular to the slits (ie y-direction in Fig.5), the sum of the momenta of the two particles in this direction can be obtained. That is, the two quantities (q_1 –

q_2) and ($p_1 + p_2$) can be accurately and simultaneously determined in this experimental arrangement. That these two quantities do in fact commute can be seen by taking $\alpha = -\pi/4$ and substituting this in the first and last equations of 27. This gives:

$$Q_1 = 1/(2)^{½} (q_1 - q_2) \quad \text{and} \quad P_2 = 1/(2)^{½} (p_1 + p_2)$$

Now, since Q_1 and P_2 commute this implies that ($q_1 - q_2$) and ($p_1 + p_2$) also commute (the complementary quantities can be found by substituting $\alpha = -\pi/4$ into the other two equations in 27). At this stage Bohr (1935 p699 note 3) points out that the wave function chosen by EPR (1935 equation 9) in their example corresponds to the special choice of $P_2 = 0$ and the limiting case of two infinitely narrow slits whose separation is equal to D where $Q = -D/(2)^{½}$. Thus Bohr has associated the rather abstract mathematical formulation of the EPR argument with a concrete experimental arrangement.

Figure 5



Bohr's Thought-Experiment

It should now be clear that given this experimental arrangement and knowing the values of ($q_1 - q_2$) and ($p_1 + p_2$), any subsequent measurement of either the position or the momentum of any one of the particles will enable one to determine, with any desired accuracy, the position or momentum of the other particle. Thus, in agreement with EPR, we see that we have a completely free choice at this stage whether we want to determine the position or momentum of the second particle by a process which does not directly interfere with this particle. However, as far as Bohr is concerned, this example "does not actually involve any greater intricacies" than the choice between experimental procedures suited for the prediction of the position or the momentum of a single particle which has passed through a slit in a diaphragm. The reason for this is that the "freedom of choice" offered in the above example is "just concerned with a discrimination between different experimental procedures which allow of the unambiguous use of complementary classical concepts" (1935 p699). In other words, the freedom of choice to measure either q_1 or p and hence compute q_2 or p_2 respectively involves a discrimination between different and mutually exclusive experimental procedures. While EPR are seemingly unconcerned by this aspect of the situation in their view of physical reality, Bohr, on the contrary, wishes to "emphasise that in the phenomena concerned we are not dealing with an incomplete description characterised by the arbitrary picking out of different elements of physical reality at the cost of sacrificing other such elements, but with a rational discrimination between essentially different experimental arrangements and procedures ..." (ibid). And again he writes: "Indeed we have in each experimental arrangement suited for the study of proper quantum phenomena not merely to do with an ignorance of the value of certain physical quantities, but with the impossibility of defining these quantities in an unambiguous way" (ibid). Let us now see how these conclusions, with which we are familiar from the discussion of Bohr's concept of complementarity in Chapter One, can be applied to the two particle situation of EPR.

Imagine, for example, that one chooses to determine the position q_1 of I. In order to perform such a measurement one requires an instrument which is rigidly fixed to the support which defines the space frame of reference (see Fig.5). Then, writes Bohr (1935 p700): "Under the experimental conditions described such a measurement will therefore also provide us with the knowledge of the location otherwise completely unknown, of the diaphragm with respect to this space frame when the particles passed through the slits. Indeed, only in this way we obtain a basis for conclusions about the initial position of the other particle relative to the rest of the apparatus." However, since in this measurement an essentially uncontrollable momentum is allowed to pass from the particle whose position is being measured into the support, it is therefore impossible thereafter to apply the law of conservation of momentum to the system consisting of the diaphragm and the two particles. We have therefore "lost our only basis for an unambiguous application of the idea of momentum in predictions regarding the behaviour of the second particle." Thus, given the experimental arrangement which is required in order to determine the position of I and hence the position of II, the possibility of predicting the momentum of II is automatically lost due to the mutual exclusion (seen here as an operational contradiction) of the two experimental procedures. If, on the other hand, we had decided to make a measurement of the momentum of I and hence predict the momentum of II, it is clear that since such a measurement of momentum requires a device which is not rigidly attached to the support (see Fig 5), we lose through the uncontrollable displacement inevitable in such a measurement any possibility of being able to determine the position of the diaphragm relative to the rest of the apparatus. We therefore have "no basis whatever for predictions regarding the location of the other particle." Thus we see that any two arrangements which admit accurate measurements of q_1 and p_1 will be mutually exclusive and therefore predictions as regards q_2 or p_2 respectively will pertain to phenomena which are basically of complementary character. It is, then, just this circumstance which, according to Bohr, implies that in QT we are not dealing with a situation which requires an arbitrary renunciation of a more detailed analysis of quantum phenomena but with a recognition that such an analysis is in principle excluded.

The ambiguity detected by Bohr in the phrase ‘without in any way disturbing a system’ should now be plain. While he admits there is no mechanical interference of system II whose state is being predicted from a measurement on system I, there is nevertheless an influence resulting from the unanalysable unity of a quantum phenomenon which involves both the object and a given experimental arrangement. Since the experimental arrangement is involved in an inextricable way, it must be included in the description of any phenomenon to which the term ‘physical reality’ can be properly attached. At this stage there is a clear distinction between the views of Bohr and EPR. For Bohr a system and measuring instrument are indivisibly (ie unanalysably) linked so that it is impossible in principle to separate the two (ie physically impossible and conceptually incoherent). The two form a single whole situation no part of which can be abstracted out. Therefore, the object cannot be ascribed an independent reality. (This is also in contrast to Heisenberg's disturbance point of view.)

Einstein, however, who at one time has said that "such an interpretation is certainly by no means absurd from a purely logical standpoint" (1949 p671), has on another occasion declared that this does not provide a "reasonable definition of reality" (1935 p780). Indeed, as the trend of the EPR argument shows, Einstein struggled to preserve a view of reality in which objects distinct from the observing instruments are ascribed an independent reality. Concerning the supposition that instrument and object form an unanalysable whole, Pauli has written in a letter of 1954 to Born that "Einstein wants to know nothing of this" (see Born 1971 p218). The feature of wholeness fundamental to Bohr's interpretation does however provide a means for avoiding the EPR conclusion since predicting the value of a physical quantity is only meaningful in a definite experimental context. Now, since complementary physical quantities require mutually exclusive experimental arrangements (including those to be predicted from measurements on distant systems), according to Bohr, such pairs of quantities cannot be simultaneously defined and it is in this sense that there is an ambiguity in the expression ‘without in any way disturbing a system’. Bohr has thus given an epistemological or conceptual reply to Einstein who expected an ontological reply. Bohr's claim that QT is complete is not an a priori dogmatic claim as has sometimes been suggested. According to Bohr, QT is complete because there is no physical experiment which cannot be fitted into its framework. Nothing has been said about the final correctness or adequacy of the theory, it is the interpretation and treatment of the theory as a certain kind of theory that is complete (see Hooker 1972 sec 12). Thus for Bohr QT indicated "the necessity of a final renunciation of the classical ideal of causality and a radical revision of our attitude towards the problem of physical reality" (1935 p697; see also 1935 p702; 1937; 1949).

At the beginning of his reply to EPR, Bohr writes (1935 p696): "I shall ... be glad to use this opportunity to explain in somewhat greater detail a general viewpoint, conveniently termed 'complementarity', ... from which quantum mechanics within its scope would appear as a completely rational description of physical phenomena." This statement suggests that no fundamental change has taken place in his point of view as a result of the EPR argument. However, the discussion and controversy initiated by the EPR paper has in fact caused an important change in the CI, replacing the disturbance point of view by a relational conception of quantum states. Thus Feyerabend (1962 p254 note 6) states: "For it was not before 1935 that the idea of the relational character of the quantum mechanical states was added to the CI. Despite later assertions to the contrary this meant a tremendous change of point of view." As an indication of this change we find Bohr (1949) writing: "... and, in commenting on Einstein's views as regards the incompleteness of the quantum-mechanical mode of description, I entered more directly on questions of terminology. In this connection I warned especially against phrases, often found in the physical literature, such as 'disturbing of phenomena by observation' or 'creating physical attributes to atomic objects by measurements'. Such phrases, which may serve to remind of the apparent paradoxes in QT, are at the same time apt to cause confusion, since words like 'phenomena' and 'observations', just as 'attributes' and 'measurements' , are used in a way hardly compatible with common language and practical definition. As a more appropriate way of expression I advocated the application of the word phenomenon exclusively to refer to the observations obtained under specified circumstances, including an account of the whole experimental arrangement."

As we have seen it is Bohr's appeal to the feature of wholeness involved in a quantum phenomenon which enables him to avoid the EPR conclusion. But, at the same time, by admitting that no mechanical interaction takes place between the second system II and the measuring apparatus involved in the measurement of I, it seems that one can no longer assume that the quantum-mechanical state description of II when all interference has been eliminated is the description of a property of II since the description of II is not independent of the situation at I. Rather, it is not possible, according to Bohr, to conceive of the quantum-mechanical state of an isolated microscopic system which one is observing (whether this observation is made directly or indirectly) but only of such a system plus the entire experimental apparatus used in the process. The state description should therefore not be regarded as characterising a property of the system under investigation but rather a relation between the system and the measuring device. Thus Feyerabend (1962 p217) writes: "EPR seems to assume that what we determine when all interference has been eliminated is a property of the system investigated. As opposed to this Bohr maintains that all state descriptions of quantum mechanical systems are relations between the systems and measuring devices in action and are therefore dependent upon the existence of other systems suitable for carrying out the measurement." This, then, is a basic postulate of Bohr (cf postulate (b')):

<u>Postulate</u> (e). A quantum-mechanical state is a relation between (usually microscopic) systems and (macroscopic) devices.

It can now be seen how this postulate explains, at least conceptually, the problem raised by EPR. Although a property cannot be changed except by interference, a relation can be changed without such interference. Feyerabend (ibid; also 1968 p102) gives the example of the state "being longer than b" of a rubber band which may change when we compress it (by physical interference) but may also change without any interference when we change b. Jammer (1974 p198) gives a further example of the state "being hotter than a body G". Thus the state of II may change without any external physical interaction but by a change in the device used in the measurement of I if 'state' is regarded as a relational conception. This follows from Bohr's statements to the effect that the conditions of observation influence in a fundamental way the physical reality itself, that is, anything that we can call physical reality. But the arguments of Feyerabend and Jammer do not quite work because, in the EPR experiment, measurements at II yield different properties which is not the case for 'hotter than' etc. If, however, one continues to attempt to maintain that quantum objects have an independent reality, it must be admitted that "in the correlation-at-a-distance phenomena, the fact that an instrument at one given place should instantaneously modify the list of the observables that have sharp values (ie, that can be conventionally attributed to the quantum system alone) at some other distant place, remains disturbing" (d'Espagnat 1971a p405).

If an ontological conclusion can be drawn from Bohr it must be this: the quantum description of physical reality is not a description of isolated quantum objects but rather a description of objects including an account of the entire experimental arrangement; objects then cannot be ascribed independent reality, only the object-apparatus totality is real. This point of view is reflected in Bohr's use of the word 'phenomenon'. It is sometimes conjectured that there might be some fundamental reality which underlies Bohr's reality but difficulties arise in QT when one tries to construct a model of reality common to complementary situations. It is Bohr's refusal to bridge the gap between complementary situations which generates the consistency in his point of view.

IV

Following Hooker (1972 sec.2), we now list some of the various assumptions of the classical conception of reality and indicate how they are challenged by QT.

C1. Physical reality is divisible into conceptually distinguishable elements and all elements have equal ontological status.

This corresponds to the firmly entrenched classical assumption that no object is any more real than any other which allows one to interpolate into the interphenomena and describe it in the same terms as the phenomena. Bohr's point of view is, however, in conflict with this assumption since he claims that measuring instruments are to be described in classical terms while quantum objects must be described in terms of QT. As a result the ontological status of measuring instruments is at all times unchanging and well-defined while that of a quantum object is in one experimental arrangement defined in one way and in another defined in quite a different way. Thus the two cannot be said to have equal ontological status.

C2. All complex objects consist in definite structures of the fundamental elements which are their constituents.

This assumption forms the basis of ontological reductionism which was formulated by Bacon (1620 sec.124) as "without dissecting and anatomising the world most diligently" we cannot "found a real model of the world in the understanding, such as it is found to be, not such as man's reason has distorted." This is echoed in Descartes' second 'Rule for Investigation' and Galileo's 'Metodo Resolutivo'. Although the development of atomic physics owed much to the application of Bacon's 'principle of dissection', nevertheless Bohr's interpretation of the wave function in QT implies a rejection of C2 due to his insistence on the unanalysable wholeness of a quantum phenomenon. Consider, for example, the complex object I + II in the EPR experiment. If it is assumed that this complex object consists in a definite structure of the fundamental elements I and II then one arrives at the conclusion, following EPR, that an object can have definite simultaneous values of complementary variables which is in conflict with the orthodox interpretation of the UP. In order to avoid this conflict, Bohr introduces his relational and holistic conception of the state of a physical system according to which the structure of the complex I + II is dependent on the particular experimental context in which the complex is situated. This involves the assumption that not only is the complex not a unique definite structure of its constituent elements but that the state of a complex object (or any other object) is not even defined at all except in relation to some experimental arrangement.

C3. The basic elements of physical reality, and structures of them, are precisely and exhaustively physically characterisable within the classical conceptual scheme of physical attributes and the objects of attribution are at definite spatio-temporal locations.

However, according to the complementary mode of description, there is a limitation on the simultaneous application of certain classical concepts in a single situation, in particular one is forced to renounce the combination, characteristic of classical physics, of the space-time coordination of events with the general conservation theorems of dynamics. This leads to a denial of C3 with the adoption of complementarity as a generalisation of the classical conceptual scheme of explanation.

C4. Physical theories correspond to, or 'mirror', the world, when they are adequate. In a completely adequate physical theory, every relevant element of reality and every relevant physical attribute of these elements has a corresponding counterpart in the theory.

This assumption is derived from the EPR definition of the term 'complete'. They write (1935 p777): "Whatever the meaning assigned to the term 'complete', the following requirement for a complete theory seems to be a necessary one: every element of the physical reality must have a counterpart in the physical theory." Thus they suppose that in any adequate theory there is a one-to-one correspondence between assertions of the theory and the world. Bohr, however, who rejects the EPR criterion of reality and who believes QT to be complete, must again reject assumption C4 since in QT there is not a one-to-one correspondence between assertions of the theory and the world.

C5. A complete description of a physical system S during a time interval T is one for which every attribute of S is precisely determined for every temporal instant $t \in T$.

Reichenbach (1944) has shown that every attempt to describe the interphenomena of quantum processes seems to lead to causal anomalies, implying that a description of the type characterised by assumption C5 cannot be given in QT.

C6. The temporal sequence of states of any system S is such that every instantaneous state of S is causally or functionally generable from the immediately temporally preceding state of S and its physical environment.

According to the CI, however, QT is complete and fundamentally indeterministic which implies that C6 is in general false.

C7. Statistical theories represent the average behaviour of physical magnitudes for a large number of distinct physical systems, identical otherwise, but whose precise particular magnitudes for the quantities in question are distributed randomly. Each of the elements of such a statistical ensemble is, however, definitely characterisable in all relevant respects. Thus, statistical theories represent less than complete knowledge of the state and behaviour of the ensemble.

Since this conclusion is denied by the proponents of the CI who regard QT as both statistical and complete, C7 is rejected in this interpretation of QT.

C8. Knowledge of the states of physical systems is gained by making measurements on the systems. A measurement is a straightforward physical interaction between a measuring instrument and measured system, the outcome of which is directly related to the feature of the system under investigation in a known way.

Bohr's relational conception of states conflicts with the first part of this assumption since, according to this conception a measurement does not give knowledge of the state of a physical system but rather of the unanalysable system-instrument whole. That a measurement is a straightforward physical interaction between instrument and system is also denied by those of the Copenhagen school for whom a measurement is supposed to cause a reduction of the wave packet, by whatever means. In this case a measurement is clearly not a straightforward interaction if reduction of this sort only takes place during a measurement.

C9. Physical systems exist and evolve independently of the presence of observers, qua observers.

This assumption is repudiated by, for example, London and Bauer for whom it is consciousness, and therefore the presence of observers qua observers, which is supposed to be the cause of the reduction of the wave packet.

Summarising the views of the various authors, Hooker (1972 sec 2) claims that Heisenberg can hardly avoid rejecting C2,3,5,6,7,8; von Neumann, London and Bauer, and Wigner, reject at least C6,7,8,9; Bohr rejects all except C9 which he retains since his is not a subjectivist position (but see discussion in Ch.2). Finally there are a number of physicists and philosophers who wish to restore much of the classical conception of reality. These include EPR, Popper, de Broglie, Bopp, Schrödinger, etc. for whom QT is either incomplete or incorrect.

To the above nine postulates, Hesse has suggested the addition of a tenth.

C10. If two identical isolated systems have been prepared in exactly the same state then they will produce exactly the same effects.

This brings out another aspect of the difference between the views of Einstein and Bohr. If one assumes the correctness of C10 and also that systems which have been identically prepared are identical then one is led to the conclusion that QT is, at least incomplete since it predicts that identically prepared systems can have different effects. Einstein concluded from this that QT is incomplete and therefore although two systems might appear to have been identically prepared, making the two systems seem identical, other factors must be considered if the two systems are to be truly identically prepared for identical systems must produce the same effects. Two systems which do not are therefore, according to Einstein, not identical. The SI is then supposed to allow for the possibility of the addition of further variables which would then explain why seemingly identical systems produce different effects. The CI, on the other hand, maintains the completeness of QT and reject C10.

One further postulate might be added to this list, the postulate of locality.

C11. The state of an isolated system I is in no way dependent upon what happens to a distant system II, whether or not there has been prior interaction.

Again this is rejected by Bohr in his relational conception of state while Einstein feels he should hold this assumption "absolutely fast". According to Bell's theorem (2) however, if one maintains that the predictions of QT are correct then one cannot maintain simultaneously C6 and C11. Einstein would therefore seem to be faced at this stage with a disagreeable decision while, according to Stapp (1971 p1307), proponents of the CI are able to circumvent this dilemma by the use of two complementary descriptions - one of which omits the requirement of lawfulness, and the other of which omits the requirement of separability.

Chapter Four: Hidden Variables

I

According to the CI, "quantum-mechanical description of physical phenomena would seem to fulfil, within its scope, all rational demands of completeness" (Bohr 1955 p696). EPR, on the other hand, conclude their paper with the words: "While we have thus shown that the wave function does not provide a complete description of the physical reality, we left open the question of whether or not such a description exists". "We believe, however, that such a theory is possible" (1935 p730). Thus we see that there is a clear distinction between the views of Bohr and Einstein. Bohr considers QT to be a complete theory while Einstein, following up the SI, considers the theory to be incomplete and believes, therefore, that one must search for a new theory which is complete. Such a theory, although perhaps still far out of reach of precise formulation, if indeed such a theory exists at all, has come to be known as a 'hidden variables theory'. The variables of this proposed theory are said to be 'hidden' because "at this stage we can only conjecture their existence and certainly cannot control them" (Bell 1971 p172) and because "we cannot at present measure them" (Bohm 1952b p185) - see also Bell (1966 p447); d'Espagnat (l971a p141); Belinfante (1975 p8). These variables may or may not be eventually observable; it does however seem probable that if these variables are to be directly observable then this will involve new forms of measuring techniques and instrumentation since present measuring instruments fail entirely to indicate their presence.

Let us begin by examining Einstein's views concerning plausible hidden variable theories (hvts). We can immediately reject one commonly held misconception concerning Einstein. Bell (1964 p195) writes: "The paradox of EPR was advanced as an argument that quantum mechanics could not be a complete theory but should be supplemented by additional variables" where these "additional variables" are the hidden variables mentioned above (see also Shimony 1973 p593). However, Einstein did not think that a satisfactory hvt could be developed merely by supplementing QT with additional variables, that is by replacing the quantum-mechanical description of state in terms of the wave function ψ by a new description $\{\psi,\lambda 1,\lambda 2,\ldots\}$ where $\lambda 1$, $\lambda 2$, etc are the additional hidden variables which, together with the original wave function, give a complete state description. In a letter to A Kupperman in 1953 Einstein wrote: "I think that one can in no way reach a description of the individual system by a mere completion of the current statistical QT." Instead, Einstein seems to have believed that what was required was a radical departure from the conventional quantum theoretical approach, "somewhat analogous to the way in which his theory of general relativity supplanted Newton's theory of gravitation" (Jammer 1974 p254). Further, concerning the future conceptual basis of physics, Einstein (1949 p666) declared: "I reject the basic idea of contemporary statistical QT, insofar as I do not believe that this fundamental concept will provide a useful basis for the whole of physics." Thus, for Einstein, QT should not be regarded as the starting point for the development of an adequate hvt, on the contrary he writes (1936 p374): "I believe that the theory (QT) is apt to beguile us into error in our search for a uniform basis for physics, because, in my belief, it is an incomplete representation of real things ..." Consequently, while EPR attempted to prove that QT gives an incomplete description of physical reality, Einstein does not believe that a complete theory (hvt) can be devised by merely completing the current QT. Rather, it is because QT in his view is incomplete that one must, according to Einstein, look elsewhere for a complete description of physical reality.

In order to clarify his view of the direction which physics should take and of the task confronting the physicist in developing a complete hvt, Einstein has suggested the following analogy. He writes (1949 p672). "For if the statistical QT does not pretend to describe the individual system (and its development in time) completely, it appears unavoidable to look elsewhere for a complete description of the individual system; in doing so it would be clear from the very beginning that the elements of such a description are not contained within the conceptual scheme of the statistical QT. With this one would admit that, in principle, this scheme could not serve as the basis of theoretical physics. Assuming the success of efforts to accomplish a complete description, the statistical QT would, within the framework of future physics, take an approximately analogous position to the statistical mechanics within the framework of classical mechanics. I am rather firmly convinced that the development of theoretical physics will be of this type; but the path will be lengthy and difficult."

Thus, just as the statistical nature of classical statistical mechanics (CSM) can be seen at a deeper level to result from the deterministic theory of classical mechanics (CM), in the same way, Einstein conjectures, the statistical QT will be seen to result from a deterministic theory at the sub-quantum mechanical level. This sub-quantum hvt will then appear in a similar logical relation to QT as CM is to CSM. However, while CM historically preceded the development of CSM, the situation with regard to QT is the reverse, the development of QT has preceded the sub-quantum hvt. This makes the path of the future development of theoretical physics "lengthy and difficult" since in the same way as CM cannot be derived from CSM so an hvt cannot be derived from QT. An hvt logically, although not historically, precedes QT. Einstein (1936 p378) therefore writes: "I do not believe that quantum mechanics will be the starting point in the search for ... [any future theoretical] basis, just as, vice versa, one could not go from thermodynamics (resp statistical mechanics) to the foundations of mechanics." After all, the fundamental concepts of thermodynamics, viz temperature, pressure, etc, cannot be taken as the starting point for the development of CM which makes use of entirely different concepts. Therefore, according to Einstein, the task facing physics is analogous to that of the construction of a theory of the CM of individual systems from the statistical mechanics of their ensembles. This task is much more complicated than its reverse.

This analogy has been somewhat clarified by Furth (see Bohm 1957a p33) who has demonstrated that a relation can be derived from classical Brownian motion which closely resembles that of Heisenberg's UP. According to the theory of Brownian motion (first proposed by Einstein in 1905) the distance Δx covered by a ‘Brownian particle in its random motions is connected to the time interval Δt during which the particle's motions are being observed by the relation:

$$\langle(\Delta x)^2\rangle = a\,(\Delta t)$$

The root mean square of the distance Δx is then

$$\delta x = \sqrt{\langle(\Delta x)^2\rangle} = a^{1/2}\,(\Delta t)^{1/2}$$

Defining the mean velocity during the time interval Δt as $\langle v\rangle = \delta x/\Delta t$, the root mean square of the random fluctuation in $\langle v\rangle$ is

$$\delta v = \sqrt{(\langle(\Delta x)^2\rangle/(\Delta t)^2)} = a^{1/2}\,(\Delta t)^{-1/2}$$

Finally, writing $p = mv$, where m is the mass of the particle, we obtain

$$\delta x\,\delta p = ma \qquad \text{... 23}$$

which is independent of Δt. Clearly equation 28 is in close analogy with the UP where the constant ma is replaced by the constant $\hbar = h/2\pi$. It is therefore not unreasonable to begin a consideration of hvts by taking Brownian motion as being analogous to QT. Note, however, that two essential features of QT, viz quantization and wave-particle duality, are not present in Brownian motion (although Bohm (1957a) has attempted to develop these features in Brownian motion). Therefore, at this stage at least, the analogy should not be regarded as in any way complete. It is however perhaps of some help in understanding the nature of the proposed sub-quantum mechanical level. Thus Bohm (l957b p80) expresses the programme of the development of an hvt as follows: "If we study the level of Brownian motion itself, we can expect to treat, in general, only the statistical regularities, but for a study of the precise details of the motion this level will not be complete. Similarly, one might suppose that in its present state of development, the QT is also not complete enough to treat all the precise details of the motions of individual electrons, light-quanta, etc. To treat such details, we should have to go to some as yet unknown deeper level, which has the same relationship to the atomic level as the atomic level has to that of Brownian motion." It is contended, then, that at a level deeper than the quantum mechanical level there are new and (for the present at least) unobserved or 'hidden' variables which are numerous and complex enough in their movements that they impart to the quantum mechanical variables a certain degree of random motion such that the UP ($\delta x\ \delta p \geq \hbar$) is ‘satisfied. The problem of an hvt is therefore to define more precisely the nature of this inner structure.

It is interesting and perhaps revealing at this stage to point out the continuity of Einstein's scientific methodology as regards Brownian motion vis-a-vis QT. At the time when Einstein proposed his solution to the problem of Brownian motion in terms of the kinetic energy of invisible constantly colliding molecules, one has to remember that atoms and molecules were still far from being considered as real - there were still physicists, including Mach and Ostwald, who did not believe in them. The success of Einstein's theory, however, seems to have had the effect of virtually proving to the physicist the existence of molecules. “The old fighter against atomistics, Wilhelm Ostwald," Sommerfeld has stated, "told me once that he had been converted to atomistics by the complete explanation of the Brownian motion" (see Clark 1973 p73). Thus Einstein succeeded in demonstrating convincingly the existence of a sub-microscopic (ie invisible) level which could explain the statistical observed effects of 'Brownian particles’. It is exactly this programme, which has already been successful for Einstein in his explanation of Brownian motion that he now wishes to extend to the atomistic level. The statistical descriptions of QT will then, Einstein believes, like the statistical descriptions of Brownian motion before it, be explicable in the deterministic terms of a new deeper level which had previously been assumed not to exist, or at least not to have any physical repercussions.

Einstein even raised his approach to the status of a methodological principle when, in a letter to Popper in 1935, he wrote: "A (method of) description which ... is statistical in principle, can only be a passing phase, in my opinion" (Popper 1959 p459). Thus, according to Einstein, for every fundamental statistical theory there must exist a deeper level of explanation in terms of which determinism is restored. In contrast to this, however Rosenfeld (1957 p42), for example, points out that there is "no logical justification for requiring a ‘deterministic substratum’ to a given set of statistical laws. There may or may not be such a substratum: this is a question to be decided by experience, not by metaphysics." Einstein cannot claim to have shown that such a deterministic substratum is logically necessary, he only states that it is his opinion that it exists.

Having to some extent clarified Einstein's view as regards a theory which would, as far as he is concerned, give a complete description of physical reality (see EPR 1935 p780), let us consider the specific requirements which, according to Einstein, such a theory must have if it is to be regarded as suitable.

1. The first requirement would seem to be that the theory is deterministic. Thus Bohm (1952a p166) writes: "Einstein ... has always believed that, even at the quantum level, there must exist precisely definable elements or dynamical variables determining (as in classical physics) the actual behaviour of each individual system, and not merely its probable behaviour."

2. As we have seen in the preceding chapter, a second requirement which Einstein demanded of an adequate hvt is that the theory be local, ie it does not admit instantaneous action-at-a-distance.

It is clear, however, that, as a consequence of Bell's theorem, the attempt to devise a theory which maintains causality and locality is doomed unless one accepts that the predictions cannot all be correct. Therefore, if Einstein maintains the correctness of the quantum mechanical predictions (as he did for the purposes of the EPR argument) then his conception of an hvt must be false and will lead to failure. It is claimed, however, by Shimony (1971 p193 note 8), following the opinions of Rosen and Podolsky, that this was not, in fact, a primary belief of Einstein. Hence Shimony writes (1971 p182), "Einstein had no confidence that the statistical predictions of quantum mechanics were all correct, and therefore he felt no obligation to seek an hvt in exact statistical agreement with quantum mechanics." Bell's problem concerning local hvts is thus avoided.

Two questions naturally arise at this point; (i) Can an hvt which satisfies conditions 1 and 2 be constructed that agrees with those predictions of QT which have been experimentally confirmed? and (ii) Is there any experimental evidence to indicate that QT is predictively false? Clearly, only when both of these questions have been answered in the affirmative would an hvt of the type advocated by Einstein be applicable. This type of hvt, some of whose predictions are in conflict with those of QT even in circumstances where elementary QT had been assumed to be valid, has been called by Belinfante (1973) an hvt of the second kind.

Before we consider the experimental evidence relating to hvts of the second kind it is interesting to note an important "difference in attitude and outlook" (as Bohr described it) between Bohr and Einstein which became clear to them both on the occasion of their first meeting in 1920. At this time Einstein expressed the view, which we find utilised above in his search for an hvt, that physicists ought to begin with certain approved fundamental propositions. However, to this Bohr strongly objected: "No, never! I would regard it as the greatest treachery on my part if, in embarking on a new domain of knowledge, I accepted any foregone conclusions" (see Clark 1973 p247). Thus it seems that from Bohr's point of view an approach towards a new deeper physical theory wherein certain preconceived demands are made prior to the construction of the theory is unacceptable. In fact Einstein's is not only the approach adopted by all those who attempt to construct non-self-contradictory hvts, but it is also their motivation for developing such a theory. Disenchanted, not by the marvellous predictive accuracy or physical usefulness of QT but by certain metaphysical presumptions in the interpretation of the theory which are considered ‘distasteful’ or 'objectionable', the proponents of hvs are motivated to develop a theory which eliminates those objectionable presumptions and includes only those which are considered 'satisfactory' or 'desirable'. This for Bohr is 'treachery'. At least it is clear that if this approach were continually adopted then one could never learn anything fundamentally surprising. Rosenfeld (1957 p45) summarises the Copenhagen view of the future development of physics with the words: "The true road to the future progress in this domain is clearly marked: the new conceptions which we need will be obtained not by a return to a mode of description already found too narrow, but by a rational extension of QT." The purpose of an hvt, however, is to attempt a return to certain classical preconceptions, such as determinism and locality, by means of a new theory of physics. In this chapter we aim to see to what extent this would be possible.

In discussing the experimental evidence relating to hvts of the second kind, it is first of all necessary to determine which of the quantum mechanical predictions would fail to be confirmed according to such theories. A quantitative answer to this question has been given by Bell (1966; 1971) who considers the possibility of a local hvt (ie an hvt of the second kind) which could account for the correlation effects predicted by the EPR analysis. Consider the experimental situation (first suggested by Wheeler (1946) as a suitable means for experimentally testing quantum mechanical correlation predictions) where an electron and positron annihilate giving two high energy photons which are emitted with equal momentum in opposite directions along the z axis. According to QT, the state of the combined system may be written as the superposed wave function

$$\Psi^{1.2} = 2^{-½} (\psi_R{}^1\psi_R{}^2 + \psi_L{}^1\psi_L{}^2)$$

where $\psi_R{}^1$ is a wave function for photon 1 which is right circularly polarised and $\psi_L{}^2$ is a wave function for photon 2 which is left circularly polarised, etc. Alternatively, the state of the combined system may be written as

$$\Psi^{1.2} = 2^{-½} (\psi_x{}^1\psi_y{}^2 - \psi_y{}^1\psi_x{}^2)$$

where $\psi_x{}^1$ is a wave function for photon 1 which is linearly polarised along the x direction and $\psi_y{}^2$ is a wave function for photon 2 which is linearly polarised along the y direction, etc. Thus, if one measures the circular polarisation of photon 1 and finds it to be right circularly polarised (ie $\psi_R{}^1$), then, according to QT, photon 2 has wave function $\psi_R{}^2$ and is therefore also right circularly polarised. Or, if the linear polarisation of 1 is measured giving 1 to be in state $\psi_x{}^1$ then photon 2 is, according to QT, in state $\psi_y{}^2$. One can therefore predict either the circular or the linear polarisation of 2 by means of an appropriate measurement on 1. However, according to QT, photon 2 cannot have simultaneous definite values of both these kinds of polarisation. The situation here regarding photons is therefore completely analogous to that considered by EPR with regard to particle positions and momenta and that by Bohm with regard to particle spins (this was first pointed out by Bohm and Aharonov 1957; 1960). Now, Wheeler (1946) suggested, it is possible to test the perpendicular correlation of polarisation by means of measurements of the coincidence-counting rates of photons scattered by carbon scatterers placed at positions z' and -z' along the z axle, the source of annihilation photon pairs being at the origin O.

If such an experiment were to verify the existence of correlation between the polarisation of photons 1 and 2, then how could this correlation be accounted for in terms of a local hvt? That is, if photon 1 is found to have linear polarisation in the x direction then how can photon 2 'know' that it should have linear polarisation in the y direction? For, according to QT, the polarisation of a photon remains indefinite until the photon meets a polarisation filter that distinguishes between two orthogonal polarisations, and the filters determining the polarisations of photons 1 and 2 are, in general, widely separated thus allowing no time for a signal to pass from one apparatus to the other. According to a local hvt, the polarisation of a photon must be determined locally at the apparatus where the measurement takes place and without reference to the outcome of any measurement made at a distant apparatus on another photon. Thus the photon (and perhaps also the apparatus) must be endowed with a set of hvs which are at the time of emission of the photon determined and act as a "tabulation of instructions" (Belinfante 1973 p13) for the particles future behaviour; in particular they enable it to decide locally whether or not to pass through a certain polarisation filter and thus account for the correlation between separated photons in terms of correlation between sets of hvs. However, since according to QT the result of a measurement on photon 1 is correlated to and not independent of the result of a measurement on photon 2 at some distant apparatus, while an hvt (of the second kind) predicts the result of a measurement on photon 1 without reference to what happens to 2, we should not be surprised that the predictions of such an hvt violate the predictions of QT. This leads to the possibility of a decisive test of the whole class of local hvts (including local hvts in which 'external' hvs in the apparatus may be assumed to play a role). Notice that in this account of hvts of the second kind we are not concerned with the construction of elaborate theories which explain all the results found in arbitrary experiments, and certainly do not try to prove that such hvts could explain all the results of QT, for we have seen that Bell (1964) has proved that we cannot do this. Instead we shall postulate some properties which such hvts have to have if they are to explain the observed correlation between the spatially separate parts of composite systems. We find that this suffices for proving contradictions with QT of a nature that is experimentally verifiable. These differences between the predictions of QT and local hvts turn out to be small and it is possible that they could have escaped detection in the past. New experiments have therefore recently been set up which are intended to decide between the predictions of QT and those of a local hvt.

We begin with a brief account of some of the earlier experiments which have been performed and whose results have a bearing on the question of the existence of local hvs. These experiments, following the suggestion by Wheeler, examine the correlation effects of photons produced in electron-positron annihilation processes using scattering techniques for polarisation determinations and have been performed by, amongst others, Bleuler and Bradt (1948) , Wu and Shaknov (1950), Langhoff (1960) and Kasday, Ullman and Wu (1970). Not all these experiments were equally accurate but, within possible errors, agreement with QT was always found. Unfortunately these experiments cannot be taken as conclusive evidence against local hvts without the inclusion of extra additional assumptions which might be false. The most important such assumption is that "the results obtained in a Compton scattering experiment are correctly related by QT" (see Kasday 1971 p200). Bell has shown (see Belinfante 1973 Appendix G of Part III) that, if one does not make this assumption, it is possible to construct an ad hoc hvt of the second kind which yields for the above experiments the quantum theoretical predictions. Now since all the above experiments use scattering techniques to determine photon polarisation (it is not possible to use filter techniques for hard, ie high energy, photons), the assumed correctness of the quantum theoretical treatment of Compton scattering is explicitly involved in the calculation of experimental results. Due to the existence of Bell's ad hoc hvt it is therefore not possible to conclusively rule out all local hvts as a result of these experiments.

Kasday (1971 p207) draws the following conclusions from his experimental results:

1. We find no evidence for a breakdown in the quantum predictions for Compton scattering of annihilation photons.

2. If we make the assumption that it is possible in principle to construct an ideal linear polarisation analyser and that QT correctly relates results obtained with ideal and Compton analysers, then it follows that a local hvt cannot describe the annihilation photons.

3. Assuming that the QT of Compton scattering is correct, Furry's assumption, that when the parts of a composite system become widely separated the state of the system, originally described in terms of a pure wave function, instead becomes a mixed state, is excluded by the experimental results, given that QT correctly describes Compton scattering.

Thus, given the assumption relating to Compton scattering and the assumption that it is possible in principle to construct an ideal linear polarisation analyser, then the above experiments on hard photons are able to exclude conclusively all local hvts as well as Furry's solution of the nonlocality paradox which appeared as an offshoot of the EPR argument on completeness. Having recognised the insufficiency of these experiments due to the use of scattering techniques in polarisation measurements, it would clearly be of great interest to conduct similar photon correlation experiments using alternative polarisation measuring instruments. This involves the detection of soft correlated photons by means of filters. Experiments using soft photons are technically more difficult to perform and involve theoretical results which have only recently been worked out by Bell (1964) and by Clauser, Horne, Shimony and Holt (1969). Nevertheless, such experiments have been under way since 1969 by Clauser (1969) and Freedman at Berkeley and Holt at Harvard. No absolutely conclusive results have yet been obtained but the indications seem to be that the early results favour QT and refute local hvs of every type. Before going on to the theory and experiments let us reflect on the significance of the above conclusion as regards the interpretation of QT.

If the hypothesis of local hvs is experimentally shown to be false then it is clear that the correlation effects of superstates cannot be accounted for in terms of the classical scheme of explanation wherein an influence by one system on another spatially separated system must be caused by the propagation of a physical entity through the intervening space from the first system to the second. The theory of relativity accepts this hypothesis and further clarifies it by showing that such an influence can propagate at most at the speed of light. Now, although the non-local influence involved in QT does not, according to QT, carry energy and is therefore, in this sense, not in conflict with relativity, nevertheless if one assumes, as Einstein seems to do (see Bohm 1971b p430), that the quantum mechanical correlations of superstates could only reasonably be explained in terms of signals carrying energy and propagating at a speed not greater than that of light then it seems that there is a conflict between the approaches of relativity and QT as regards their mode of explanation of information correlation. For Einstein it is QT which is regarded as suspect rather than relativity but the following are important reasons for questioning the ultimate acceptability of relativity rather than QT:

(i) It is well known that when one attempts to combine logically the theories of special relativity and QT then one is led to what we could call mathematical nonsense, constant observables become infinite. This is usually taken as some sort of argument against QT, but we see from the above conclusions that it might perhaps rather be taken as evidence against relativity. This claim is further substantiated by the fact that by means of a rather simple technique (ie renormalisation) it is possible to obtain experimental conclusions from the infinities of relativistic QT which are confirmed by experiment to an extremely high degree of accuracy. These results could not conceivably be obtained from relativity alone but instead are characteristically quantum mechanical. Also, as we shall see in Chapter Five, the results of RQT have many related philosophical problems which only come to light in this new quantum mechanical framework.

(ii) The cosmological evidence seems to indicate a deep inadequacy in the theory of general relativity. This is indicated by the 'big bang' theory of the origin of the universe due to Gamow and the proposed 'presence' of singularities or 'black holes' in the fabric of space-time. Thus Ellis (1971) writes: "The existence of a singularity in our past implies a breakdown of ordinary physical laws." These physical laws are not those of QT but rather those of general relativity. This gives further reason for questioning the ultimate correctness of relativity rather than QT as the probable source of the difficulties encountered in RQT.

If then Einstein's solution to the problems of interpretation of QT in terms of some theory which preserves local causality is erroneous, it seems that we must return to QT per se.

So, what bearing does the proved non-existence of local causality have on the problems discussed in Chapter 1-3? Consider first the double slit experiment. It will be remembered that in his reply to EPR, Bohr (1935) demonstrated how the double slit arrangement could be used for an exposition of the EPR correlation effects. Since his explanation of the correlation effects of the superstate is in no way dependent on the postulated existence of local causality his arguments are not in the least influenced by its breakdown and therefore his complementarity interpretation of the double slit experiment remains unscathed. The Schrödinger cat experiment can also be seen to be based on the concept of a superstate and can therefore also be worded in terms of an EPR correlation effect (magnified). Imagine two cats, one in box A and one in box B facing slits A and B respectively. The superposed wave function for the Schrödinger situation can then be written:

$$\Psi = 2^{-\frac{1}{2}} (\psi_{alive}^{A}\psi_{dead}^{B} + \psi_{alive}^{B}\psi_{dead}^{A})$$

where ψ_{alive}^{A} represents the final state where cat A is alive, etc (cf Ch.3 equation 15). This shows a correlation which is magnified into macroscopic magnitudes between cats A and B. The correlation in this case would seem to be rather trivial, for all it says is that if cat A is found to be alive then cat B will be found to be dead. This, we might argue, is simply because the electron passes through either hole A or hole B but not through both. However, according to Bohr's interpretation of the double slit experiment, it is not possible to talk like this. Instead we must accept that no local causal explanation can be given of the simple correlation between the health of cats A and B. Thus we are led back to philosophically radical non-classical interpretations of QT such as that of Bohr.

Let us now return to the experiments of Clauser and Holt. In his experiment Clauser measured the coincidence rates of pairs of soft photons emitted in opposite directions from excited calcium atoms while Holt did the same for excited mercury atoms. Polarisation filters were used in both experiments to determine the polarisations of the photons, thus assumptions concerning the validity of the application of QT to Compton scattering were eliminated. The theory behind these experiments can be derived as follows. Let A be a function describing photon 1 which is equal to +1 if the photon is transmitted by filter 1, and equal to -1 if the photon is stopped by filter 1. Similarly a function B is defined for photon 2. According to the locality hypothesis of a local hvt, A will be a function of the orientation direction of filter 1, which we write as $\underline{a}$, as well as a function of some hv or set of ' hvs which we write as λ. A, by hypothesis, will be independent of the orientation $\underline{b}$ of filter 2 (and similarly for B). Thus,

$$A = A(\underline{a}, \lambda) = \pm 1 \text{ and } B = B(\underline{b},\lambda) = \pm 1 \qquad \dots \ 29$$

If ρ(λ) is the probability distribution of λ then the expectation value of the product of the two quantities A and B is

$$P(\underline{a},\underline{b}) = \int A(\underline{a},\lambda)\, B(\underline{b},\lambda)\, \rho(\lambda)\, d\lambda \qquad \ldots\ 30$$

Since for $\underline{a} = \underline{b}$ the expectation value is -1, this implies that $B(\underline{b},\lambda) = -A(\underline{b},\lambda)$ for all λ. Therefore, from 30

$$P(\underline{a},\underline{b}) = -\int A(\underline{a},\lambda)\, A(\underline{b},\lambda)\, \rho(\lambda)\, d\lambda$$

Taking now a third direction $\underline{c}$, we have

$$P(\underline{a},\underline{b}) - P(\underline{a},\underline{c}) = -\int A(\underline{a},\lambda)\, A(\underline{b},\lambda) - A(\underline{a},\lambda)\, A(\underline{c},\lambda)\, \rho(\lambda)\, d\lambda$$

$$= \int A(\underline{a},\lambda)\, A(\underline{b},\lambda)\, [\, A(\underline{b},\lambda)\, A(\underline{c},\lambda) - 1\,]\, \rho(\lambda)\, d\lambda$$

So finally we get

$$|\, P(\underline{a},\underline{b}) - P(\underline{a},\underline{c})\, | \le \int [\, 1 - A(\underline{b},\lambda)\, A(\underline{c},\lambda)\,]\, \rho(\lambda)\, d\lambda = 1 + P(\underline{b},\underline{c})$$

ie

$$|\, P(\underline{a},\underline{b}) - P(\underline{a},\underline{c})\, | \le 1 + P(\underline{b},\underline{c}) \qquad \ldots\ 31$$

Equation 31 was first derived by Bell (1964) and is called Bell's inequality.

This inequality has been derived on the assumption of perfect correlation, ie $P(\underline{b},\underline{b}) = -1$. This assumption is however experimentally unrealistic and therefore Clauser, Horne, Shimony and Holt have considered the case of imperfect correlation. For perfect correlation, if $\underline{b}' = -\underline{b}$ then $P(\underline{b}',\underline{b}) = 1$. For imperfect correlation they assumed that for every $\underline{b}$ there exists a $\underline{b}'$ such that $P(\underline{b}',\underline{b}) = 1 - \delta$, where $0 \le \delta \le 1$ and for good correlation δ is close to zero. Given this generalisation Clauser et al (1969) then derived the generalised inequality

$$|\, P(\underline{a},\underline{b}) - P(\underline{a},\underline{c})\, | + P(\underline{b}',\underline{b}) + P(\underline{b}',\underline{c}) \le 2 \qquad \ldots\ 32$$

Further generalisation of this relation has been given by Bell (1971 p179) who has shown that relation 32 continues to hold even when the instruments themselves are assumed to contain hvs which locally influence the results. Note also that with $\underline{b} = \underline{c}$ in 32 and assuming $P(\underline{c},\underline{c}) = -1$, relation 32 yields the original relation 31.

The question now arises, how can the inequality 32 which is predicted to hold experimentally for all local hvts be tested? Let a_l and a_2 be two arbitrarily chosen orientations of filter 1 (and similarly b_1 and b_2). Relation 32 can then be written

$$|\, P(a_1,b_1) - P(a_1,b_2)\, | \le 2 - P(a_2,b_1) - P(a_2,b_2) \qquad \ldots\ 33$$

If $w[A(a)_{\pm},B(b)_{\pm}]$ is the probability that $A(a) = \pm 1$ and $B(b) = \pm 1$, then

$$P(a,b) = w[A(a)_{+},B(b)_{+}] - w[A(a)_{+},B(b)_{-} - w[A(a)_{-},B(b)_{+}] + w[A(a)_{-},B(b)_{-} \quad \ldots\ 34$$

and for perfect efficiency of the coincidence counting,

$$R(a,b)/R_0 = w[A(a)_+,B(b)_+]$$

$$R_1(a)/R_0 = w[A(a)_+,B(b)_+] + w[A(a)_+,B(b)_-]$$

$$R_2(b)/R_0 = w[A(a)_+,B(b)_+] + w[A(a)_-,B(b)_+]$$

where R(a,b) is the coincidence counting rate for photons 1 and 2 passing through filters 1 and 2 which are oriented in directions a and b respectively, $R_1(a)$ is the coincidence rate with polarisation filter 2 removed and filter 1 with its polarisation direction at angle a, $R_2(b)$ is the coincidence rate without filter 1 and with filter 2 at angle b, and R_0 is the coincidence counting rate of the two detectors where both filters 1 and 2 have been removed. Substituting ‘ in 34 then gives

$$P(a,b) = [\ 4R(a,b)-2R_1(a)-2R_2(b)\ ]/R_0 + 1$$

Inserting 35 into 35 one finds

$$|\ 4R(a,b)-2R_1(a)-2R_2(b)\ | \le -4R(a_2,b_1)-4R(a_2,b_2)+4R_1(a_2)+2R_2(b_1)+2R_2(b_2)$$

Removing the redundant modulus sign, bringing everything to the left hand side of the relation and dividing by 4 then yields the relation obtained by Clauser et al (1969).

$$\Delta \equiv R(a_1,b_1) - R(a_l,b_2) + R(a_2,b_1) + R(a_2,b_2) - R_1(a_2) - R_2(b_1) \le 0$$

ie $\Delta \le 0$

Thus, according to any hvt of the second kind, for perfectly efficient experiments the measurable quantity Δ is zero or negative. It can be shown, however, that, according to QT, it is possible with an appropriate choice of angles to make Δ positive (see Belinfante 1975 pp291-3). In the experiments of Clauser and Holt, the angles $a_1 \dots b_2$ are chosen so as to maximise the quantum theoretical value of Δ, then a mere verification of the sign of Δ as calculated from the measured coincidence rates can verify whether the hvt is false or whether QT is false.

Of course, in any realistic experimental situation the instruments will not be perfectly efficient, in particular the efficiency of the detectors of photons 1 and 2 and the efficiency of the coincidence counting mechanism will not be equal to unity. Belinfante (1973 pp299-300) has shown, however, that even for the case of imperfect efficiency the inequality $\Delta \le 0$ remains valid to a good approximation, therefore the experiment should still be able to give conclusive results. According to the most recently published results of Freedman and Clauser (1972) the results of their experiment are now conclusively in favour of QT, even when incorporating the assumptions concerning imperfect efficiency. According to Belinfante (1973 p290), Clauser is now devising a new experiment which will eliminate any need for assumptions about the efficiency of the counters. The assumptions concerned here are that the efficiency of the detectors are approximately independent of the presence and orientation of a filter, ie on whether the incident light is unpolarised or polarised in any specific direction. Since the detectors used (photomultipliers) respond to the energy rather than the polarisation of a photon and since it is assumed that a photon leaving a polarisation filter has practically the same energy as the photon incident upon the polarisation filter, there is good reason to believe that the above assumption is correct. Therefore, if we can go by the results of Freedman and Clauser as is suggested by Shimony (1971 p193 - note added in proofs), all local hvts are experimentally refuted. We find, however, in a footnote (25[b]) in Belinfante (1975 p290) that the results obtained by April 1973 by Holt, who started similar experiments for his Ph.D. work at Harvard University, "apparently contradict QT. Others plan to continue measurements with his apparatus." Also in a footnote (182), Jammer (1974 p339) writes: "It should be noted, however, that the results obtained ... by Holt ... do not fully agree with the predictions of quantum mechanics. Whether this is due solely to experimental errors or has a deeper cause remains to be tested with improved instrumentation." Belinfante (1975 p313) concludes: "Unless Ho1t's experimental results would be confirmed and the results of Freedman and Clauser would be shown to be faulty, we would have to drop theories of the second kind for their disagreement with experimental data." This seems to be as far as one can go at the present time concerning the experimental evidence as regards local hvs. But, remember, there is here no question of suppression of a new theory by the orthodox. There is no local hvt. Rather it seems such a theory would be false.

If one is still determined to maintain local causality then there is yet one more argument which could be made. According to Bell's arguments, the condition of locality implies that A ≠ A(<u>b</u>) and B ≠ B(<u>a</u>) unless relativity theory is violated since no signal can, in general, travel over the arbitrarily large distance between filter a and filter b at less than the speed of light in time to tell how the two results should be correlated. It might be argued, however, that at the moment when the two filters are set up and oriented there is communication between the hvs of one filter and those of the other at speeds not greater than that of light which enables each filter to 'find out' the orientation of the other and thence produce the correct polarisation correlation for incident photons without violating relativity. In order to eliminate this possibility Clauser has suggested the use of Kerr cells for changing the polarisation orientation of the filter axis while the photons are in flight, which would ensure that A = A(<u>b</u>) and B = B(<u>a</u>) only if relativity is violated (see Shimony 1971 p191; I973 p598). This would seem to be experimentally possible but has never been done.

Finally, on the question of local hvts, it has sometimes been suggested that Einstein's views on this matter can be proved to be impossible without reference to experiment. Thus Gardner (1972a p18) writes: "There can be no doubt that Einstein's programme - recovering the quantum statistics from a distribution in an underlying phase space - is impossible, at least on Kochen and Speaker's interpretation of the programme." This interpretation of Einstein's programme is, however, different from that considered above, since Kochen and Speaker (1967), who, in fact, do not mention Einstein in their paper, only deal with hvts which reproduce the quantum mechanical predictions. Since it seems, according to Rosen and Podolsky (see Shimony 1971 p182 & p193 ref.8), that Einstein had no confidence that all the statistical predictions of QT were correct, we have not made Gardner's assumption concerning Einstein's hv programme. On this view, therefore, the above proof of impossibility does not apply to Einstein because we have taken Einstein's proposals to be in favour of an hvt of the second kind.

II

We now turn to a different kind of hvt. The characteristic of this kind of theory (which Belinfante (1973) calls an hvt of the first kind) is that they reproduce all the usual predictions of QT. In all the theories of this kind which have been proposed (eg Bohm 1952a,b; Wiener & Siegel 1953,1955; Bohm & Bub 1966) this requirement is satisfied in the following way. The statistical predictions of QT are assumed to arise as a result of QT not taking into account the existence and influence of one or an unspecified number of 'hidden' factors at a sub-quantum mechanical level which precisely determine the outcomes of physical interactions, including measurements, in analogy with molecules which were proposed in order to explain deterministically the apparently random motions of Brownian particles suspended in a liquid. It is then supposed that if these factors are taken into account by appropriate specifications of hvs then it would be possible to explain in deterministic language the outcomes of measurements on quantum systems which at present can only be described statistically by QT. If this were all that was to be said concerning these hvs then it can easily be shown that the predictions of such hvts would be able to contradict the predictions of QT since one could, by means of appropriate consecutive measurements on a system, restrict the range of values of these hvs in such a way that non-commuting observables relating to the system could be simultaneously and precisely determined, in disagreement with the UP (see Wigner 1970 p1009 note 2). In order to avoid this conflict it is supposed that some (unspecified) mechanism exists which randomises the hvs, producing in a very short time an equilibrium distribution which is self-perpetuating. It is then conjectured that in the case of an equilibrium distribution of hvs, the hvt will make exactly the same predictions as ordinary QT. This is again in close analogy with the case of Brownian motion wherein the ‘hidden’ causes producing the statistically described movements of Brownian particles are believed to be the randomly (Gaussian) distributed velocities and consequent forces of the microscopic molecules of the liquid. If the velocities of these molecules should significantly change their distribution from a random distribution then the original statistical predictions of the motions of the Brownian particles would no longer be found to be correct (cf hvs and the statistical predictions of QT).

There is, however, one way in which the above analogy breaks down in theories of the first kind. In CSM the parameters which are found to be useful in describing, for example, a gas of molecules (eg temperature, pressure, etc.) are of no use in the descriptions by CM of the individual motions of each particle of the gas, for these parameters represent only an average and incomplete description of the microscopic nature of the gas. Although, according to hvts of the first kind, the wave function of a quantum system again represents only an incomplete description of the state of the system, it is found in these hvts that the wave function plays a necessary and very important part in the complete description of the system. Thus, the analogy proposed by Einstein (1949 p672) breaks down in hvts of the first kind.

The "very essence of an hvt", according to Bohm & Bub (1966 p474), is that it restores determinism to physics although it is not clear how such a theory is possible. In an hvt of the first kind this is attempted by regarding quantum mechanical states as "ensembles of states further specified by additional variables, such that given values of these variables together with the state vector determine precisely the results of individual measurements" (Bell 1966 p448). That is, while the state of a system described merely by a wave function ψ is regarded as being incompletely described, it is postulated (contrary to the views of Einstein Ch4, I) that a complete description can be given by means of the wave function ψ together with certain other hvs, which we shall signify by the letter λ. Thus the state of a system is completely specified by a specification of ψ and λ. Further, the characteristic feature of an hvt of the first kind is that it reproduces the predictions of QT when one has an equilibrium distribution of hvs. Representing this equilibrium distribution by $\bar{\lambda}$ we have that the probability predictions of the hvt, $P(\psi, \bar{\lambda})$ are equal to those of QT ie $|\psi|^2$ when $\lambda \rightarrow \bar{\lambda}$. Hence

$$P(\psi, \bar{\lambda}) = |\psi|^2 \qquad \dots 37$$

Also it is postulated that as time increases any non-equilibrium distribution hvs in an undisturbed system will rapidly tend to equilibrium and this equilibrium is self-perpetuating:

$$P(\psi,\lambda) \rightarrow{}^{t}\!\rightarrow P(\psi,\bar{\lambda}) \rightarrow{}^{t}\!\rightarrow P(\psi,\bar{\lambda}) \qquad \ldots 38$$

Thus in a theory of the first kind, as opposed to a theory of the second kind, deviations from QT only occur in non -equilibrium distributions of hvs.

Since these theories reproduce the predictions of QT, it is clear that, as a result of Bell's theorem, they cannot hope to restore local causality to physics. This, although in itself a disturbing conclusion for proponents of hvts, raises another question as regards these theories. We have seen in the introduction how Feyerabend has criticised the CI for the vagueness which has allowed complementarity to be "reinterpreted in the light of objections and still give the impression that these objections had really missed the point." While this criticism is to some extent justified, in particular as regards the CI of the EPR paradox, it would seem that it is far more relevant as a criticism of hv interpretations of quantum mechanical predictions. This is so since in the early stages of hvts the proponents of these theories have attempted to restore a number of classical conceptions, such as local causality, and when it has turned out that these conceptions cannot be restored to physics they have not simply abandoned the interpretation but have turned to other features which might conceivably be restored and have continued to call their theories by the same name, hvts. Thus there are a number of hvts which attempt to do different things, including some which have been shown to be impossible in the sense that they are self-contradictory. When it becomes clear that one aim is hopeless, they turn to another and make that the fundamental purpose of their theory. Since there is no end to this type of approach it is clear that hvts of all conceivable sorts (even if we restrict ourselves to hvts of the first kind) can never be totally disproved. Some definite conclusions can, however, be obtained concerning what any hvt definitely cannot achieve and what any hvt might possibly achieve (at least in as much as these things are not at present known to be self-contradictory). Further than this it will not be possible to go unless an hvt is developed which actually makes definite predictions which go beyond those of QT and are experimentally verified. No such theory has yet been proposed, although if it should ever be developed there is no doubt that "science will have ascended to a new plane of power and fertility" (Hanson 1962 p93) and that this hvt will leave all previous interpretations of QT behind. Hvts will have proved themselves to be the superior approach to QT. However, no such theory is anywhere in sight, therefore Hanson (1962 p92) writes, "Speculation is never an alternative to a working theory, however imperfect that theory may be. If Bohm, Vigier and Feyerabend as much as hint that they have an available alternative theory they do themselves a disservice. Until their speculations churn out numbers it is highly reasonable to doubt that they ever will ... it is also reasonable for practising physicists to act on the assumption that Bohm, Vigier and Feyerabend will not succeed with the development of their speculations, and unreasonable for anybody to interpret this lack of confidence as a symptom of dogmatic orthodoxy." We begin here with an examination of the nature of these speculations and with the motivations and aims of various attempts to construct acceptable hvts.

The following is a list of some of the various demands which are made, either together or individually, of ‘acceptable’ hvts. Each demand can be seen to be an attempt to restore some basic characteristic of classical physics which is thought to be desirable in any reasonable physical theory. It should be understood from the beginning that these hvts are not intended simply as an interpretation of the basic principles of QT such as the SP but rather as an expression that something is wrong with the basis of QT and as an attempt to go beyond this basis to a new and fundamentally different basis.

I. There is one demand which seems to be essential to all hvts that "for a completely specified state of the system, in which the values of the hvs are all determined, the result of a measurement of any observable can be predicted with certainty" (Bohm & Bub 1966b p474). This motivation to find a deterministic hvt can, historically, first be seen from Einstein's reaction at the 1927 Solvay Conference to the indeterminacy in QT when he argued that a complete specification of the initial conditions of a process should suffice in principle to determine the course of the process, and therefore the statistical character of the predictions of QT implies that the quantum description is incomplete (see Jammer 1966 pp357-9). In terms of an hvt where the state of a system is described by means of a wave function, ψ, together with hvs λ, this demand for determinism can be viewed in the following way. According to QT, if a measurement of an observable A is made on a system described by the wave function ψ which is not an eigenfunction of the operator A representing the observable A, then the outcome of the measurement will be some eigenvalue ψ_i and the system will be left in the state where, in Dirac notation,

$$|\psi\rangle = \sum_i c_i \ |\psi_i\rangle$$

where the c_i are, in general, complex numbers, ie where the set of all ψ_i form a complete orthogonal set of vectors spanning the space containing the vector ψ and which are eigenvectors of the operator A. The expectation value, $\langle \psi|A|\psi\rangle$, then predicts the outcome of the measurement to be a_i with probability $|c_i|^2$, a_j with probability $|c_j|^2$, etc. Now, according to requirement I, an hvt should, by supplementing the state $|\psi\rangle$ with hvs λ giving $|\psi,\lambda\rangle$ yield a unique eigenvalue prediction. That is,

$$\langle\psi,\lambda|\ A\ |\psi,\lambda\rangle = a_n, \text{ say, where } n = n(\psi,\lambda) \qquad \ldots 39$$

In other words, determinism can only be restored to physics if one assumes that the reduction of the wave packet taking place during a measurement and written $\psi \rightarrow \psi_n$, where ψ_n is the particular final state of the system, can be described, causally by an equation where n is shown to be a function of the hvs λ and also, possibly, a function of ψ, ie $n = n(\psi,\lambda)$. Unless such an equation can be found no causal explanation of the reduction of the wave packet can be given and hence no deterministic physics is possible. If, however, such an equation should be found then this would have many implications; practically all the discussion in Chapter Two would become irrelevant. In particular, it would allow one to reason 'if different members of an ensemble described by the wave function ψ are found upon measurement to have different values a_i of some observable A, then these individual systems must have been in different initial microstates'. It will be remembered from Chapter Three, IV that this basic criterion of classical reality is rejected in the CI of QT. We shall see later, however, in the context of the work of Gleason (1957) and of Kochen and Specker (1967), that it is not possible to restore determinism by means of an equation as simple as $n = n(\psi,\lambda)$. This has the effect of replacing the demand for predictive or ontological determinism of macroscopic systems by, at best, one of ontological "cryptodeterminism" for microscopic systems undisturbed by interactions with macroscopic systems. That is, the best that can be done by an hvt of the first kind is to justify the belief that 'everything happening in nature is predetermined by previous happenings in the physical wor1d', but the knowledge that the hvs predetermine the results of a measurement would be of no practical use in that they would not enable us to predict with certainty the results of measurements. Belinfante (1973 p18) calls this type of determinism 'cryptodeterminism'. It is now argued (see eg Siegel 1966), not that hvts are invented to restore determinism, but rather that they are invented for making the laws of nature cryptodeterministic instead of being governed by unexplainable randomness. It is doubtful whether a cryptodeterministic explanation of quantum mechanical randomness would have satisfied Einstein. Having in mind the randomness associated with the time of emission of a particle in the process of radioactive disintegration (see 1936 p374; also 1949 pp667-8), he asks (1936 p377): "Is there really any physicist who believes that we shall never get any inside view of these important alterations in the single systems, in their structure and their causal connections?" It is most likely that Einstein would not have been totally satisfied with a theory unless it could actually predict the observed time instant of the disintegration of a radioactive atom, for example. This, as we shall see, no hvt of the first kind can possibly do. Thus Bub (1968 p186) explicitly now claims that the hv approach is not an attempt to restore classical determinism.

II. In our discussion of the CI in Chapters One to Three we have seen how, according to Bohr, a description of the measuring apparatus must be given in ‘the usual terminology of CP and everyday language’ while the description of microscopic objects is given in terms of QT. This produces a peculiar dichotomy of physics into quantum and classical descriptions which has given rise to speculations about the existence and positioning of a 'cut' or 'boundary' between quantum and classical domains. Now Bell (1971 p172) writes: "For me the possibility of determinism is less compelling than the possibility of having one world instead of two." And so he says (ibid): "It is this possibility, of a homogeneous account of the world, which is for me the chief motivation of the study of the so-called ‘hidden variable’ possibility." This request for a homogeneous account of the world can be divided into two distinct aspects. Firstly, it can be seen as a desire to restore to physics a single ontology for micro- and macro-domains as is to be found in classical physics. Thus the world would be homogeneous in that there would be ontological reducibility of macroscopic to microscopic entities. Secondly, the request might be taken as meaning that an acceptable hvt should postulate laws which are applicable to macro- and micro-domains, that is there should be reducibility of laws. According to the CI, QT does not satisfy either of these requirements while CP satisfies both. Thus, if the CI of QT is regarded as progress in the development of physics, the request for a unitary account of the world (as well as the request for determinism) can be viewed as reactionary. Ontological reductionism is not, however, regarded by all proponents of hvs as being a necessary or even desirable aspect of an hvt, thus Bub (1968 p186) claims that "the deep intention behind the development of hvts is the realisation of a 'natural philosophy’ which incorporates a concept of 'wholeness' as a new ontological thesis." The concept of unanalysable wholeness of a quantum phenomenon forms one of the foundational pillars of Bohr's philosophy and appears as revolutionary in that it conflicts with Bacon's 'principle of dissection‘ which had been regarded as indispensible in the investigation of microscopic phenomena. Thus, as Bub (1968) argues, hvts, when conceived as incorporating this concept, contain a significant portion of the revolutionary doctrine of the CI thereby avoiding the charge of being reactionary. Others, however, for example Fine (1972 p4), wish to deny that the innovations of QT support a conceptual superstructure that is in any way revolutionary. Instead Fine (1972 p27) supports what he calls a ‘revisionist thesis’ which he hopes is not reactionary wherein an attempt is made to salvage ontological reducibility by means of a reinterpretation of the two slit experiment. It turns out, as a result of the work of Gleason and Kochen and Specker, that the best that an hvt can possibly do for the restoration of reductionism is to postulate a single ontology for macro- and micro-domains, but reducibility of macro- to micro-laws is not possible. This situation is really unsatisfactory since without reducibility of laws the ontological state of a combined macro-micro-system may ‘jump’ discontinuously on observation from one ontological state to another macroscopically distinct state in an unpredictable way. Such an ontology would hardly seem much more intelligible than that of the CI in terms of superstates.

III. A third motivation for the search for an hvt arises as the result of a suggested solution to the problem of measurement (see Shimony 1973 pp593-4). Here it is proposed that hvs might enable one to return to the realistic, definite state ontology of CP. If this were possible then one could say that the result of a measurement involving a macro-micro-interaction is always definite and objective prior to the registration of the result on the consciousness of the observer, thus restoring realism to physics. Such a definite state ontology involves states which are known as 'dispersion-free' states, ie "states in which the values of all possible observables are simultaneously determined by physical parameters associated with the observed system" (Bohm 1952b p187). There is, however, a famous proof given by von Neumann (1932; 1955 Ch4 sec 1 & 2) to the effect that dispersion-free hv states cannot reproduce the results of QT and are therefore impossible. If this proof were rigorous in the sense that it had achieved what it had intended to achieve then it would appear that all possible kinds of hvts consistent with the results of QT are impossible. This impression is vindicated by the fact that von Neumann concludes his proof with the remark (1955 p324) that "we need not go any further into the mechanism of the 'hidden parameters' since we now know that the established results of QM can never be re-derived with their help." Thus, from the time of its publication in 1932 for the next twenty years it was generally believed in the scientific community with few exceptions (see eg Reichenbach 1944 p13; Blokhintsev 1952 pp95-193) that all discussion of hvs was pointless and had been shown to lead to contradictions. That this was far too sweeping a claim was shown by Bohm in two ways in his 1952 papers. Firstly he gave an example of a crude but consistent hvt which one would suppose impossible according to the conclusion of von Neumann, and secondly he gave explicit arguments indicating the "irrelevance" of von Neumann's proof as regards his own hvt. This reopened the possibility of speculations about hvs which had been stifled by the over-generalised claim made by von Neumann for the applicability of his proof.

Although it was recognised from 1952 that something was wrong with the claims which had been made concerning von Neumann's proof, it was not until 1966, as a result of the work of Bell, that exactly what was wrong became clear. Von Neumann's proof is mathematically valid (for a summary of the proof, see Ballentine 1970 pp374-5) but the difficulty was shown by Bell (1966 pp448-9) to lie in the relation between the mathematics and the physics. In particular, Bell showed that one of the essential assumptions made in the proof, although correct for quantum mechanical states, "is by no means a 'law of thought' and there is no a priori reason to exclude the possibility of states for which it is false" (Bell 1971 p174). This assumption is: Any real linear combination of any two Hermitian operators represents an observable, and the same linear combination of expectation values is the expectation value of the combination. The difficulty which arises when this condition is applied to an hvt arises because the hv states are assumed to be dispersion-free and therefore, as shown above in I, there is no distinction for these states between expectation values and eigenvalues. For a dispersion-free state the eigenvalue is unique and therefore always equal to the expectation value. Now consider the example of the spin components of a particle. The spin operators for the x, y and bisector directions are σ_x, σ_y and $2^{-½}(\sigma_x + \sigma_y)$ and all have eigenvalues ±1. It is clear that the eigenvalues are not additive in the way that the expectation values are additive according to the above assumption, ie the eigenvalues of $2^{-½}(\sigma_x + \sigma_y)$ are not the corresponding linear combinations $2^{-½}(\pm 1 + \pm 1)$ of eigenvalues of σ_x and σ_y. Thus the assumption that expectation values are additive excludes the possibility of dispersion-free states. Therefore, concerning the possibility of an hv interpretation, von Neumann concluded his proof (1955 p325) with the ambiguous (see Bub 1968 p203; Bohm 1971a p101) statement "it is therefore not, as is often assumed, a question of re-interpretation of QT - the present system of QM would have to be objectively false in order that another description of the elementary process than the statistical one be possible." However, the additivity assumption is not trivial, indeed it is not at all obvious why it should be true for expectation values of non-commuting observables such as the spin components above since a measurement of each different component requires a different experimental arrangement (eg a Stern-Gerlach magnet in a different orientation). Thus there is no way of relating a measurement of ($\sigma_x + \sigma_y$) to those of σ_x and σ_y since all require a different experimental arrangement. Further, the result of a measurement giving an eigenvalue of ($\sigma_x + \sigma_y$) will not be equal to the sum of an eigenvalue of σ_x plus an eigenvalue of σ_y. The fact that the average over a large number of measurements of ($\sigma_x + \sigma_y$) should be equal to the sum of the averages over a large number of the two other measurements involving different experimental arrangements is a peculiar and non-trivial property of quantum states. While this result must be reproduced in an hvt of the first kind when the hvs are averaged over, there is no reason to demand it of the individual dispersion-free states which (at least for the present) cannot be individually prepared. Indeed, since every observable of a dispersion-free state has a unique value equal to its eigenvalue, and since there is, in general, no linear relationship between the eigenvalues of non-commuting observables, von Neumann's assumption cannot possibly apply to dispersion-free states. Thus von Neumann's proof does not exclude every possible hvt as he claims, but only a certain very limited class of hvts which make the above (inappropriate) assumption. In fact, Scheibe (1973 p164 & p171) has pointed out that von Neumann's assumptions are so restrictive that CM, contrary to what is usually assumed, would not be an hvt in relation to CSN under von Neumann's interpretation of the programme. Omitting the additivity assumption, Bell (1966 p448) has shown, by means of an example, that it is possible to construct an hvt of the first kind which is realistic in the sense that its states are dispersion-free.

IV. A final requirement of an hvt to be considered here is that it be 'noncontextualistic'. A noncontextualistic hvt is one in which the specification of an hv, λ, corresponds to some intrinsic attribute of each observable of a system. In particular, the measurement of some observable of a system is independent of the apparatus used to make the measurement. Thus, for example, a measurement of the total angular momentum J^2 might be made using an instrument which simultaneously measures J_x of the system, or one which simultaneously measures J_y. According to a noncontextualistic hvt, the result of the measurement of J_2 is independent of the 'context' supplied by the additional measurements of J_x or J_y. Let us split such theories into two possible types, stochastic and deterministic noncontextualistic hvts. Both are independent of the context supplied by the particular measuring instrument used in the measurement of some observable A; in the stochastic theory the specification of λ yields a probability distribution $P_\lambda(\alpha)$ over the range of possible values α of the observable A where this probability distribution is regarded as an intrinsic attribute of the observable A of the system while in the deterministic theory the specification of λ attributes a definite value $A(\lambda)$ to each observable A of the system. Some definite conclusions can be reached regarding both these types of hvt. Concerning stochastic hvts, it has been shown by Bell (1971) that his inequality (equation 32) applies equally to stochastic and deterministic local hvts. Therefore, if the experiments of Clauser should prove that $\Delta > 0$ then stochastic local hvts would be disconfirmed. There is however no empirical evidence which refutes nonlocal stochastic hvts. But Shimony (1971 p192) has raised methodological objections against this family of theories since, at present, the family is too large and "some member of it seems to be compatible with any experimental data whatsoever." Until this family can be limited to some empirically accessible subfamily, it cannot be given detailed attention.

As regards deterministic noncontextual hvts the conclusions are more final. All such theories, except in the case of extremely simply systems, are inconsistent. This was first proved in a corollary of a theorem of Gleason (1957) and in a different way by Kochen and Specker (1967). This corollary implies that if the set of observables contains all the quantum mechanical observables of a system, then there exists no consistent noncontextualist deterministic hvt for the system, unless it is describable quantum mechanically with a Hilbert space of dimension 2 or less. Since, according to Shimony (1975, p595), "noncontextual deterministic hvts ... until recently were the only hvts ‘ considered in most discussions", it is clear that, as a result of Gleason's corollary, proponents of hvts must once again change the fundamental premises on which their hvt would be based. The necessity to consider contextualist hvts can, in fact, be viewed as bringing one closer to the basic postulates of Bohr than to those of classical physics. This point is argued by Bell (1966 p451) who indicates the similarity between contextualist theories and Bohr's view regarding "the impossibility of any sharp distinction between the behaviour of atomic objects and the interaction with the measuring instruments which serve to define the conditions under which the phenomena appear" (1949). Bell (ibid) concludes that "the result of an observation may reasonably depend not only on the state of the system (including hvs) but also on the complete disposition of the apparatus" as is required in contextualist hvts. Thus, by accepting Bohr's arguments to the effect that the measuring instrument is involved in a fundamental and inextricable way in the discussion of quantum systems, Bell justifies the reasonableness of contextual hvts and hence allows the search for hvts to continue despite the restrictions imposed by Gleason.

Let us consider now, in detail, the conclusions which can be drawn from the work of Gleason (1957) and Kochen and Specker (1967). (Gleason was not primarily concerned with hvs but Belinfante (1973 pp44-6) has shown how conclusions concerning hvs can be drawn from an unnamed corollary (1957 p889) to Gleason's main work. Bell (1966 pp450-1) has given a simplified proof of these conclusions. Simplified proofs of the arguments of Kochen and Specker have been given by Belinfante (1973 pp38-9 & pp63-4); see also Gardner (1972a pp14-5). Since Kochen and Speaker state (1967 p70) that their conclusions can be obtained from the corollary to Gleason's theorem, we shall consider their results to be equivalent to Gleason's (although they obtained the results in different ways). The mathematical result can be expressed very simply: equation 39 must be replaced, in the best possible hvt, by

$$n = n(\psi,\lambda,\{\psi_i\}) \qquad \ldots 40$$

where $\{\psi_i\}$ is the complete orthonormal set of eigenfunctions that are the possible results of the measurement. In other words, it would never be possible to devise an hvt which is able to determine the particular outcome ψ_n of a system initially in state ψ simply by a specification of ψ and the hvs λ. It is also necessary to specify all the eigenfunctions $\{\psi_i\}$ of which ψ_n is a member. Thus, for example, suppose one makes a measurement of observable A of a system initially in state ψ. If the measurement yields the value A_n thus leaving the system in state ψ_n then even if all the hvs λ of the system were known it would still be impossible to write a deterministic equation for the system $\psi_n \rightarrow \psi_n$ unless all the eigenfunctions of A were known uniquely. Given this requirement then the ψ and λ together may select from the set $\{\psi_i\}$ unambiguously the state ψ_n which represents the result of the measurement. However, difficulties arise because the set {ψ_i} is not necessarily uniquely defined by consideration of the system alone. Take, for example, the case considered by Kochen and Speaker of a j = 1 wave function where

$$J^2 = J_x^2 + J_y^2 + J_z^2 \qquad \ldots 41$$

and where J_x^2 J_y^2, J_z^2 and J^2 all commute, ie have simultaneous eigenfunctions. When equation 41 operates on such an eigenfunction, each term in the equation multiplies this eigenfunction by the corresponding eigenvalue, which is $2\hbar^2$ for J^2 and either $\hbar^2$ or 0 for each other term. Thus, if J_i^2 is 0 then the other two J_j^2 ($j \neq i$) must each be $\hbar^2$. Since all the J^2's commute, they can be simultaneously measured (however, it is only necessary to measure two of the J_i^2's and the third can be inferred). Suppose that J_z^2 and J_x^2 are measured and found to be both $\hbar^2$. Now if the x, y, z frame is oriented about the Z axis so that x ≠ x', y ≠ y' and z = z' then one might expect that a measurement of $J_{z'}^2$ would give the same result as that of J_z^2, while one would not be surprised if a measurement of $J_{x'}^2$ gave a different result from that of J_x^2 since here the measurement is being made along a different direction. However, Kochen and Specker proved that if one attempts to assign a definite value (either $\hbar^2$ or 0) to J_z^2 independent of the directions x and y in which other measurements are made simultaneously then one is led to a contradiction. Kochen and Specker concluded from this that hvts were all impossible, but this claim is too general. Rather, what they have shown is that the result of a particular measurement depends not only on this measurement but also on which other measurements are performed simultaneously since only together do all these measurements define uniquely the complete set of eigenfunctions {ψ_i} which is involved. Thus, in agreement with equation 40, no definite value could be assigned to J_z^2 without reference to a specific frame of orthogonal directions to be considered, and hence without reference to a specific set {ψ_i}.

A conclusion which can be immediately deduced from equation 40 is that if one assumes as a basic requirement of an hvt that observables have values before they are measured then measurements do not in general yield those values, since otherwise the value yielded for one observable would not depend on which other observables are being measured simultaneously. It is therefore impossible to retrodict from the results of measurements the value of the observable of the system prior to measurement. Consider then the consequences for a realistic hvt such as has been proposed by Fine (1972; 1973) and by Popper (1967) who, according to Gardner (1972a p21), is concerned "not with determinism but with realism. He simply wishes to claim that it is at least possible (ie, consistent with QT) to maintain that unmeasured observables have values." If realism is to be maintained in, for example, the Schrödinger cat experiment, thus enabling one to say that the cat is in some definite classical state or other and not in a superposition of states before being observed, then since retrodiction to the state prior to observation is impossible it is not possible to say that the state which the cat is observed to be in is the same as the state the cat was in prior to observation. The state of the cat is always definite (ie alive or dead, but not a superposition of both) but it may change discontinuously and unpredictably from one to the other possibility at the moment of observation. Thus, although the assumption of realism allows one to maintain ontological reducibility, and also cryptodeterminism for micro- objects, one cannot maintain reducibility of laws from macro- to micro-domains and further there is the serious ontological problem that macroscopic objects may change from one state to another macroscopically distinct state merely as the result of observation. Such an ontology is no more comprehensible than that in terms of superstates.

Another implication from equation 40 is that deterministic noncontextual hvts are all impossible since in order to obtain a deterministic equation for n one has to specify $\{\psi_i\}$ which, in general, requires one to specify the experimental 'context'. Belinfante (1973 p19) has concluded from this that an hvt cannot tell what the state of a system per se is but only what state would be found if the system is subjected to a particular experimental arrangement. For example, a measurement of an observable A may be made by an instrument which distinguishes between some specific complete orthonormal set of eigenfunctions of A, $\{\psi_i\}$. In this case the result of the measurement is A_n where $n = n(\psi,\lambda,\{\psi_i\})$. It may happen, however, that the measurement distinguishes between some different set $\{\varphi_j\}$ of A of which ψ_n is again a member. (This could happen if A has some degenerate eigenvalues.) In general, $n(\psi,\lambda,\{\psi_i\})$ will not be equal to $n(\psi,\lambda,\{\varphi_j\})$ and therefore it is possible that the two measurements will give different results A_g and A_h even though the system is described by the same function ψ and hvs λ. Thus an hvt could not tell what the values of certain observables are but only what values would be found if these observables were measured in some particular way. Now it might be objected that hvts are impossible because they would contradict the UP by simultaneously predicting the values of noncommuting observables, such as position and momentum. However, the best that an hvt can do is to predict the value which would be obtained in a particular experimental arrangement. Since there are no simultaneous eigenfunctions of x and p, $\{\psi_i\}$, it can be seen from equation 40 that no hvt can simultaneously predict the precise values of x and p. They could, however, predict the precise values of each which would be obtained given different experimental arrangements.

We thus see that there are a number of classical theoretical features which cannot possibly be restored to physics by means of an hvt, assuming that the impossibility proofs leave no loopholes (but see Jammer (1974 p312) and Belinfante (1973 p41 & p95)). Let us now look at Bohm's hvt which is a consistent, nonlocal, non-relativistic, contextual, cryptodeterministic hvt of the first kind to see what an hvt can possibly do. Since the theory is nonlocal, Bell (1966 p452) indicates that it resolves the EPR paradox "in the way which Einstein would have liked least." In a letter to Renninger, dated 3 May 1953, Einstein wrote, referring to Bohm's theory, "... I don't think this theory is tenable." And again, in a letter to Born dated 12 May 1952, Einstein wrote, "Have you noticed that Bohm believes (as de Broglie did, by the way, 25 years ago) that he is able to interpret the QT in deterministic terms? That way seems too cheap to me." The reason why the theory was 'too cheap' for Einstein to take seriously was, according to Ballentine (1972 p1770), because "Bohm's hvt did not establish any deeper connections between quantum mechanics and electromagnetism, gravitation, or any other theory." Other objections may have been that the theory assumes that the wave function describes the individual system rather than an ensemble of systems in disagreement with Einstein's SI, and also the theory supposes that a complete description of the individual system can be obtained merely by a completion of current QT (cf Ch4, I). It seems that the theory was not even taken very seriously by its author who writes concerning the original Bohm hvt (Bohm & Bub 1966a p462): "It was not seriously envisaged as an ideal theory and suffered from a lack of simplicity and elegance of structure." Nevertheless, the theory has been described by Jammer (1974 p277) as a "major breakthrough" since it supplied a counter-example to the long-held von Neumann impossibility proof. Since it was this proof that seemed to entail the ontological interpretation of QT, once this proof becomes questionable then so does the ontological interpretation which is at the basis of the CI. Thus the possibility of a reinterpretation of QT is opened up when a loophole is seen in von Neumann's proof; Bohm's hvt forms the basis of such a reinterpretation.

The essence of Bohm's theory is that it is possible to rewrite the Schrödinger equation in a form which, in the classical limit ($\hbar \rightarrow 0$), resembles the Hamilton-Jacobi equation of classical physics. This suggests the possibility of some sort of classical reinterpretation of QT. This had already been recognised by Madelung (1926), de Broglie (1926) and Rosen (1945). Certain objections to the theories of de Broglie and Rosen led these authors to abandon their attempts to reinterpret QT realistically, however Bohm (1952b Appendix B) argues that these objections could have been overcome if the ideas had been taken to their logical conclusion. Taking the above quantum mechanical equations, Bohm then proposes the introduction of hidden variables λ, for the position of a particle and $d\lambda/dt$ for the velocity of the particle. Thus a particle is assumed to have a definite position and velocity (or momentum) at all times. The theory is therefore realistic in the sense of III. Since, however, no methods are known for measuring λ without disturbing $\psi(\lambda)$, the λ are 'hidden'. Notice also that the momentum mv, $m d\lambda/dt$, of a particle as defined by Bohm is not the same as the quantum mechanical observable represented by the operator P. Now, following the analogy with the Hamilton-Jacobi equation and taking λ to be the usual particle position of QT, ie $x = \lambda$, Bohm concludes that the velocity of the particle can be written,

$$v = d\lambda/dt = (\hbar/2im)\, \nabla \ln(\psi/\psi^*) = (\hbar/2im)\, (\psi^*\nabla\psi - \psi\nabla\psi^*)/\psi\psi^* \qquad \ldots 42$$

The corresponding force is then

$$f = m\, d^2\lambda/dt^2 = -\nabla(V+U) \qquad \ldots 43$$

where V is the usual potential field occurring in the Schrödinger equation and U is a new potential, called the quantum mechanical potential by Bohm (1952a), given by

$$U = -(\hbar^2/4m)\, [\nabla^2(\psi^*\psi) - ½\nabla(\psi^*\psi)]/\psi^*\psi \qquad \ldots 44$$

The nonlocal character of this theory is now easily seen since U can transmit instantaneous disturbances from one particle to another. The realistic character of the theory is also clear since every particle is ascribed a definite position and momentum at all times, however, as a result of the quantum mechanical force $-\nabla U$, the trajectory of a particle will, in general be very complicated due to the complex form of U in equation 44 and its complicated dependence on the wave function ψ of the system.

There are a number of serious objections to this theory. The most obvious of these is that the hv λ is a purely hypothetical quantity since it does not show itself in any experiment yet performed and can certainly be given no operational meaning since it is not known how it could be measured. The theory is therefore unverifiable, and thus, at least for the positivist, meaningless. A second objection follows from the attempt to regard ψ as being like an electromagnetic field which also exerts a force on a (charged) particle, however the initial velocity of a charged particle does not depend on the electromagnetic field, so the assumption that it would depend on ψ, as seen by equation 44, is not satisfactory. In order to avoid this assumption Bohm (1952a pp178-9) has suggested an ad hoc modification to the theory which complicates the mathematical formalism and is, of course, experimentally unverified. Another more serious physical objection has been pointed out by Belinfante (1973 p121) who writes, "Bohm predicts accelerations of free particles with no experimental evidence to support such a contradiction in terms, there is no sense in upholding such a claim." Finally there is the objection that Bohm's theory (which, by the way, he explicitly introduces in the hope that it will prove experimentally adequate in the domain of distances of the order of 10^{-13} cms or less) cannot be generalised to include relativistic effects without beginning again from scratch (and it is just in the domain of 10^{-13} cms that relativistic effects begin to become important). It is also significant as regards a relativistic Bohm type theory that in the relativistic realm the very thing which Bohm attempts to retain of classical physics, viz a definite continuous trajectory of individual particles, becomes more and more difficult to achieve at higher energies and smaller distances due to pair production (see Belinfante 1973 p114).

Nevertheless, despite these objections, the theory is realistic and logically consistent. Prior to this theory (1952) it was thought that von Neumann had proved the impossibility of such a theory. The advantage of a realistic theory is that it allows one, in principle, to describe an individual system in terms of a single precisely defined conceptual model in conflict with the conclusions of the CI. Consider, for example, the two slit experiment. We have seen (Ch.1) how, according to the CI, it is impossible to find a single model to explain all the results of this experiment; in particular the interference pattern is seen to conflict with a particle model. According to Bohm's theory (1952a p173), however, the quantum mechanical force (which is admittedly nonlocal) can account for the interference pattern in terms of a particle model since this force allows the particle to go to only certain bands on the detector while other bands are forbidden due to the presence of an infinite negative quantum mechanical force. Thus the usual claim that a single model to account for the results of this experiment is inconceivable is refuted. The trajectory of each particle is, however, extremely complicated and depends not only on hvs of the particle but also on those in the measuring apparatus. Nevertheless the conclusion is reached that, in contrast to the CI, it is possible to say, at least theoretically, that each particle passes through a definite hole in the screen. This conclusion is obviously much closer to the usual classical conceptions than that of the CI and to that extent it is more acceptable. It should be stressed, however, that there are ways in which this theory differs from classical conceptions and agrees with those of the CI. Thus, for example, Bohm (1952a p188) declares: "In this point, we are in agreement with Bohr, who repeatedly stresses the fundamental role of the measuring apparatus as an inseparable part of the observed system." This follows from the fact, mentioned above, that, according to Bohm's theory, the trajectory of a particle does not just depend on the hvs of the particle but also on those of the measuring apparatus. Bohm's theory should therefore not be thought of as classical, but only, perhaps, semi-classical.

Having demonstrated the logical consistency of formulating a single unified model of quantum mechanical processes which may involve causal and continuous motions of particles and fields at a sub-quantum mechanical level (in close agreement with traditional methodology of physics), Bohm (1957a, b; 1962) goes on to elaborate a possible model of the sub-quantum level. He says (1962 p65) that it is "not a model that I necessarily believe in, but one that ... may be true." The model which he describes is analogous to classical Brownian motion, but not the usual type of Brownian motion wherein a particle undergoes random displacement of its centre of mass while continuing to exist as a particle. Instead, the type of Brownian motion which Bohm (1957a p35) considers more appropriate is that for a gas near critical point such that the particle itself may be constantly forming and dissolving. Bohm regards this picture as more appropriate since quantum field theory describes all particle motion in terms of the continual creation and destruction of elementary particles and thus no particle retains a fixed identity as a particle throughout its motion, particularly when very short dimensions are considered. The following picture of the sub-quantum mechanical level thus emerges (1957b p121): "... the particle-like concentrations are always forming and dissolving. Of course, if a particle in a certain place dissolves, it is very likely to reform nearby. Thus, on the large-scale level, the particle-like manifestation remains in a small region of space, following a fairly well-defined track, etc. On the other hand, at a lower level, the particle does not move as a permanently existing entity, but is formed in a random way by suitable concentrations of the field energy."

In terms of this picture the two slit experiment becomes conceptually unproblematic; an initially single particle need not pass through one and only one hole but can dissolve on one side and condense on the other side of the screen in such a way that the particle seems continuous but is in fact at the moment when it arrives at the screen a gas which passes through both holes. This model is not, however, worked out quantitatively and in detail but is only conjectured on the grounds that a realistic model of this sort is not logically impossible given Bohm's hv interpretation of QT. There is again no question of any experimental evidence to support such a model except in so far as the hvt is constructed in such a way as to reproduce all the usual predictions of QT (this is clear since the formalism of the theory is so close to that of ordinary QT - for this reason Jammer (1974 p255) does not wish to call this a hv theory, but rather an hv interpretation).

Since the appearance of Bohm's hvt in 1952, two other hvts of the first kind have been proposed which are improvements on the 1952 theory and which are intimately connected since both involve the same method (called the 'polychotomic algorithm') for determining the choice of the resulting wave function ψ_n in a measurement from the set $\{\psi_i\}$ of possible outcomes. The first such theory to appear was that by Wiener and Siegel (1953; 1955) in which cryptodeterministically behaving hvs determine the results of measurements by the polychotomic algorithm and in which Brownian motion again serves as a model. The most promising hvt of the first kind in existence is, however, the theory of Bohm and Bub (1966a) which makes use of the ideas of Wiener and Siegel but goes beyond them by attempting to describe the reduction of the wave packet in terms of a causal equation. This equation is constructed merely by adding to the usual Schrödinger equation another term, B(x,t), which is intended to deal in a causal and continuous way with what happens during a measurement. It is then proposed that between measurements B is zero, or very small, since between measurements the Schrödinger equation without modification seems to be perfectly adequate to describe the evolution of the wave function in a causal and continuous way. During measurements, on the other hand, it is proposed that B becomes very large, swamping other terms and explaining how ψ can change causally and continuously into the final state ψ_n. The Bohm-Bub equation of evolution of ψ is

$$d\psi(x,t)/dt = B(x,t) - (i/\hbar)\, H\, \psi(x,t) \qquad \dots 45$$

where H is the Hamiltonian operator. It turns out, as a result of the analysis by Tutsch (1971), that B must necessarily be a rather complicated and ugly looking term involving integrals in the denominator under a summation sign and therefore on aesthetic grounds equation 45 appears dubious (Bohm and Bub in fact choose the simplest possible form for B). One particular consequence of 45 is that, since B cannot be linear in ψ, the validity of the SP must necessarily be broken during measurements. A further consequence of the form of B is that a change in ψ at one point depends on the value of ψ at every other point in space. Thus the proposed equation of motion, 45, is not only nonlinear but also nonlocal (as must be expected from the arguments in section I). The theory is therefore non-relativistic, without hope of relativistic generalisation if the equation of motion is to be retained in its present form. Another consequence of the nonlocal character of 45 is that no measurement can be made on one part of a system without thereby affecting the rest of the system, thus Bohm and Bub (1966a p467) confess that "only complete measurements are now feasible." They then give the following example to explain this point. Consider two particles which have interacted in the past and which now separate. The state function ψ describes the state of the composite system as a whole, and on measurement this function, still describing both particles, 'collapses' according to 45 into an eigenvector of the system as a whole. It is thus conceptually impossible to perform an experiment on one single particle of the pair. The authors therefore conclude (ibid): "Consequently, the paradox of EPR does not arise in this theory, because a composite system must always be regarded as an indivisible totality which in principle cannot he subdivided into independently existing units. (This is really the essence of Bohr's refutation of the paradox.) Of course, this is a fault and not an advantage of the theory, which has been formulated in terms which exclude the framing of the paradox." For reasons such as this the authors regard the theory as only a step towards a serious tentative theory (see Bohm and Bub 1966a p462) and "it is definitely not proposed that the theory ... is likely to be a 'right' one" (1966a p454).

The value of the Bohm-Bub theory lies in the extent to which it shows that it may be possible to restore determinism to QT by means of the introduction of cryptodeterministically behaving hvs at a sub-quantum mechanical level which describe the reduction of ψ as a deterministic process. Let us see briefly how this attempt avoids the difficulties imposed by von Neumann and Gleason. Von Neumann's criticisms are avoided by abandoning the additivity requirement which was unnecessarily restrictive and replacing it by a nonlinear relation. Gleason's difficulties cannot be avoided in any way if determinism is to be restored (as pointed out by Bell 1966) and therefore Bohm and Bub accept that as a consequence of restoring determinism they must retain the "wholeness" feature of Bohr's interpretation. This point has been much emphasised in later writings by Bohm (1971a p97) and particularly Bub (1968). Bub (1968 p186) even goes to the extent of claiming (in retrospect?) that "the hv approach is an attempt to extend Bohr's basic insight into the revolutionary significance of the QT, and not a conflicting proposal to restore classical determinism"! Similarly, Bohm (1971a p97) now says that the "deep intention (of hvs) is actually the continuation and extension of what is most basic and novel in Bohr's insights." Notice, however, that this (new?) aim of an hvt is unavoidable as a result of Gleason's proof and any other conflicting aim is now known to be inconsistent. This suggests that the proponents of hvs are learning from QT while there is little to suggest that Bohr's insights have been extended by hvts in any useful way.

We have not yet mentioned how hvs enter into this theory. The hvs are here vectors in a dual Hilbert space, $\langle\lambda|$. Equation 45 then describes the measurement process in terms of a nonlinear differential equation which relates the components of $|\psi\rangle$ to those of $\langle\lambda|$, where $B = B(\lambda)$. The Bohm-Bub theory is a theory of the first kind since the authors have shown that all the predictions of QT are reproduced when one averages over the hvs λ. Therefore, since the predicted outcome of a measurement of some observable A is equal to a_n where

$$\langle\lambda|A|\psi\rangle = a_n \qquad \ldots 46$$

When the hvs are averaged over giving $\langle\lambda| \rightarrow \langle\bar{\lambda}|$ then the theory gives predictions which are exactly those of QT, ie

$$\langle\bar{\lambda}|A|\psi\rangle = \langle\psi|A|\psi\rangle \qquad ...47$$

Physically, this implies that since in all usual experimental tests of QT the predictions of ordinary QT are found to be valid, under these conditions the hvs must have been randomly distributed. However, since immediately after a measurement it is possible, according to the Bohm-Bub theory, to deduce which of a number of equations the components of $\langle\lambda|$ satisfy (by the polychotomic algorithm), the components of $\langle\lambda|$ will no longer be completely random. Therefore, if another measurement is made on this system immediately after the first measurement then the results of this experiment need not agree with the predictions of ordinary QT since the conditions of equation 47 do not hold. Thus we have, as is pointed out by Bohm and Bub (1966a p466), the possibility of an experimental test of the theory. All that is required is an estimate of the time taken for the hvs to randomise after a measurement. Once an estimate of this time has been given then all that needs to be done is to make two measurements on a system one after the other, the second following the first in a time less than the time taken for the hvs of the system to return to equilibrium. Bohm and Bub (ibid) give an estimate of this time by making the following assumption: "Since most systems are either in thermal equilibrium or have emerged from a source in equilibrium, it seems plausible to assume a random process resulting from this. The characteristic unit of time of thermal processes in relation to quantum mechanics is: $\tau = \hbar/kT \approx 10^{-13}$sec, for room temperatures. Typical measurements generally involve longer times, so that the usual results of quantum mechanics are to be expected." Thus Bohm and Bub, again following the analogy of Brownian motion, make the assumption that the cause of the rapid randomisation of hvs after a measurement is thermal vibration, although no detailed account of the randomisation mechanism is given. They conclude that in order to make a measurement on a system with nonrandom distribution of hvs one would have to perform a measurement on a system in a time less than 10^{-13}sec after a previous measurement which will produce the nonrandom distribution.

Since all hvts of the first kind are based on the assumption that the predictions of QT are necessarily valid only when certain hvs are randomised and since a measurement causes a momentary upset in the random distribution, it is clear that a double measurement of the above type, if the time between measurements is made sufficiently short, can afford a test for all hvts of the first kind. Of course, without making definite assumptions concerning the time τ, no such test can be conclusive since it can always be conjectured, if the predictions of QT are confirmed, that the time τ is actually shorter than that involved in the experiment. Nevertheless, since a definite estimate of τ has been given by Bohm and Bub, it is of great interest to conduct an experiment wherein two measurements are made in rapid succession, the time between them being less than 10^{-13} sec. Obviously, if such an experiment showed the predictions of QT to be wrong then this would be of great significance (although non-relativistic QT might not be adequate to deal with such experiments where times much less than 10^{-13}sec are involved). However, such an experiment has been performed by Papaliolios (1967) and his results confirm to within 1% the predictions of QT down to a time $\tau = 2.4 \times 10^{-14}$ sec. This shows that the estimate of τ given by Bohm and Bub is too large by at least a factor of 4. If one still wishes to maintain, for reasons of metaphysical principle, an hvt of the first kind and if one also wishes to maintain that the choice of $\hbar/kT$ as the relaxation time might still have some relevance then it would be of interest to repeat the experiment at lower temperatures since "it is unclear ... whether the relevant temperature is that of the apparatus or of the observed system (or of both, in some combination)" (Bohm 1971a p113). But Papaliolios (1967 p624) concludes that "the lack of theoretical understanding of this choice of τ ... does not at this time justify cooling the apparatus to liquid air (or lower) temperatures." Papaliolios has however proposed to refine his measurements in other ways (see Jammer 1974 p336). Since all experiments so far confirm the predictions of ordinary QT, it may be concluded that we may reject all hvts of the first kind on the grounds that they are needless complications of ordinary QT (at least in terms of its quantitative predictions).

III

In sections I and II of this chapter we have considered some of the usual considerations as regards the meaning, motivations, capabilities and limitations of so-called hvts. In particular we considered the possibility of restoring locality, causality and reducibility (ontological and lawlike) to fundamental physics. We concluded that local causality seems to be refuted by experimental evidence, while nonlocal causality and reducibility of laws would seem to be impossible if the predictions of QT are to be maintained. Thus we found a gradual shift in emphasis from hvts which were supposed to restore the basic philosophical assumptions of classical physics to those which are intended to 'extend Bohr's basic insight into the revolutionary significance of the QT'. This is obviously a dramatic change of viewpoint and must be reflected in a growing confusion as to what exactly an hvt is now supposed to be.

In this section we shall not be at pains to 'absolutely hold fast' to any specific classical philosophical principle such as locality or causality since it seems that such an approach is bound to be unsuccessful. Instead we shall return to Bohm's original conception of an hv in the sense of a variable which 'we cannot at present (ie 1952) measure' (see Ch4, I). In the search for such variables we shall not be concerned to find fault in the quantum mechanical formalism nor in any of the predictions of QT, ie we shall accept the quantum mechanical description of those variables which can at present be measured, such as position, momentum etc. Instead we look for new variables which may be essentially quantum mechanical (eg spin) but which have until now been hidden both theoretically and experimentally from the physicist's view. This type of theory we call an hvt of the third kind.

In the realm of modern elementary particle physics such a theory is, in fact, not hard to find. For example, a recent theory of nuclear matter introduces a new variable called 'charm' which is somewhat analogous to spin and which was, until the development of the theory, totally hidden from theoretical gaze. Experimental evidence for this theory appeared in 1974. The particular theory which we shall consider here involves other variables such as 'isotopic spin', 'hypercharge' and 'strangeness'. All these variables are totally new to the physicist's vocabulary but are well-defined and as experimentally meaningful as any other physical concept. Their novelty and unfamiliarity (obvious from the names) would seem to make them appropriate candidates for the name hvs, at least in the sense of Belinfante (1973 p8) who says that certain variables are hvs "because they kept themselves hidden from the eyes of those who invented QT." The variables of this theory are not, however, hidden in the sense of position and momentum for molecules vis-a-vis CSM since QT is assumed to be accurate in describing the domain of this theory (cf CSM in the domain of individual molecules). But it is a theory (probably unlike those of the first and second kinds) which does what Bub (1968) and Bohm (1971a) want an hvt to do, namely extend Bohr's basic insights. The motivation for this kind of theory (the third kind) is not the amendment of false interpretations of QT but the extension of, essentially, the orthodox interpretation into new and problematic realms which appear accessible to quantum mechanical treatment. This realm is the realm of the so-called elementary or fundamental particles.

The theory which we consider as an example of an hvt of the third kind is called the ‘eight-fold way’ since it introduces eight new variables which were once hidden from sight. (Remember Einstein's comment in Ch3, that ‘it is the theory which decides what we can observe’.) This theory was developed by Gell-Mann and independently by Ne'eman in 1961 as a unified theory of symmetries between these eight variables, or ‘quantum numbers’. The theory might be said to be based on a generalisation of the relation between x, y and z components of the angular momentum operator in QT:

$$[J_x, J_y] = -i\, J_z \qquad \dots 48$$

where J_x is the x component of the angular momentum operator, etc. In the eight-fold way this relation is generalised to the relation:

$$[J_i, J_j] = i\, \varepsilon_{ijk}\, J_k \qquad \dots 49$$

where ε_{ijk} is a permutation symbol such that

$$g_{ij} = \varepsilon_{ipq}\, \varepsilon_{jpq} = -\, 2\delta_{ij}$$

where g_{ij} defines the metric of the space (this introduces the possibility of relativistic generalisation of the theory).

Equation 49 forms one of the basic equations of Lie algebra which was developed in the 19th century by Sophus Lie and which has been known for a long time to be related to QT (see Weyl 1923). This algebra forms the basis of Lie groups which are used in the eight-fold way to classify the elementary particles into groups which contain particles that are symmetrically related to one another through the eight new variables. (The importance of symmetry in this theory of elementary particles has led Heisenberg (1976) to compare elementary particles with 'Platonic forms'.) The 'representation' of a Lie group is given by the operator

$$D = e^{-i\underline{\alpha}.\underline{I}} \qquad \dots 50$$

where the components of 1 satisfy the relation

$$[I_i, I_j] = i\varepsilon_{ijk}\, I_k \qquad \dots 51$$

These D operators correspond to transformations in the space of basis states (which are, for example, |j,m> in the case where $\underline{I}$ is angular momentum) and these transformations are characterised by the real parameter $\underline{\alpha}$.

The physics of the eight-fo1d way then emerges from the discovery that a number of operators can be found which represent observables associated with elementary particles and which satisfy equation 51. D operators can then be constructed which transform one basis state (signifying the wave function of an elementary particle) into another. In other words, one can transform algebraically from one elementary state (or particle) into another by suitable operations. The eight-fold way finds eight operators which satisfy 51 and which represent observables associated with elementary particles. Then certain groups of these particles can be found which transform into one another and not into any others outside the group by means of the transformation operators D. These groups can then be regarded as self-contained in the sense that every particle in the group is related to every other by way of a symmetry operation in the space of the eight variables of the eight-fold way.

The variables in the eight-fold way do not, however, enable every elementary particle to be grouped by the scheme, for example the electron has not been grouped in this way. Therefore Gell-Mann does not believe that all the hvs have necessarily yet been found; he writes in Chew et al (1964) "there might exist some undiscovered quantum number ..." (eg charm). The scheme is nevertheless extremely successful in grouping many of the strongly interacting particles which have been found in nature.

But the real test of a theory is in its predictive power. Here the theory proves its worth since it has predicted the existence of a number of elementary particles which have since been found to exist in nature. For example, the theory predicted in 1961 the existence of a particle (called the Ω^-) having a certain set of values for its various new quantum numbers and the theory even predicted the approximate mass of the particle. In 1964 a particle was found which satisfied the description of the Ω^- and had the correct mass. The theory has thus been confirmed by experiment unlike any hvt of the first or second kind. Further, the theory tells us how to simplify, in terms of classification or grouping, the very large set of elementary particles which have now been discovered. Since, throughout, the validity of QT is assumed, confirmation of this theory supplies further confirmation of QT itself, and in particular it supplies confirmation in the realm of the elementary particle which is exactly the realm where, for example, the 1952 theory of Bohm questions the validity of QT. Thus more doubt is cast on the validity of hvts of the first kind. Apart from the particular predictions of the theory (ie the existence and properties of individual particles), the theory also makes general predictions about the mass differences between the particles in each group and about the number of members in each group. These predictions are largely confirmed by the evidence.

If this theory is accepted as an hvt then it is clear that all the usual problem of the orthodox interpretation of QT remain in this kind of hvt. However, there is a sense in which this theory is "more nearly determinate than the QT" as has been demanded of an hvt by Bohm (1957a p33). This can best be seen by an analogy. Consider all the various life forms which are, according to biological theory, ‘complete individual beings’. Then let us distinguish between two approaches to these ‘live entities’; that of natural history, of classification and grouping of the various species, and that of evolution, of the dynamical and continuous change of one species into another through time. So far we have been concerned with the theory of dynamical determinism, ie with the theory of how a system changes its state in time and space either during a measurement or otherwise. We are now, in hvts of the third kind, concerned with 'existential determinism' which is the theory of what basic states can and do exist. The aim of the theory is then to predict all the possible states of the world, given that they evolve (ie are dynamically determined) by the Schrödinger equation. Thus, the eight-fold way is not concerned to more nearly determine the dynamics of quantum mechanical states but rather to determine which elementary quantum mechanical states can possibly exist in nature. The theory is therefore concerned with existential determinism and does, in fact, enable one to determine the existence and properties of many of the elementary particles found in nature. The purpose of this section has been fulfilled if it has only suggested that the pursuit of understanding of QT is much more likely to lead to success if one considers the unmodified formalism. That is, it is this theory (ie QT which has to be understood and not some hypothetical elaborate hvt of the first or second kinds.

Chapter Five: Beyond

I

The need for a change in the fundamental principles of classical physics was recognised for the first time as a result of the development by Einstein in 1905 of the theory of relativity. However, the elementary QT discussed in Chapters One to Four does not take into account the requirements of relativity. This can easily be seen from a glance at the form of the Schrödinger equation in which the space coordinate $\underline{x}$ appears as a second order differential, $d^2/d\underline{x}^2$, while time appears as a first order differential, d/dt. Thus space and time in elementary QT do not appear on an equal footing as would be required by relativity and therefore elementary QT cannot incorporate the conclusions of relativity theory. Since relativity theory is physically well established, the non-relativistic nature of elementary QT clearly indicates a major limitation in the elementary theory. In particular, the theory cannot be expected to work at particle velocities approaching the speed of light nor at energies approaching the creation energies of elementary particles. In this chapter we consider attempts to go beyond these limitations of elementary QT and examine some of the principal problems of quantum ontology which ensue. In the orthodox interpretation of relativistic QT (RQT) we find, in agreement with Bohr (1927 p90), that "in the adaption of the relativity requirement to the quantum postulate, we must ... be prepared to meet with a renunciation as to visualisation in the ordinary sense going still further than the formulation of the quantum laws considered here(unto)."

According to Dirac (1963), Schrödinger's first attempt to devise an equation which generalised the ideas of de Broglie, did, in fact incorporate the refinements required by relativity. However, when he applied this equation (which is now called the Klein-Gordon equation for spin zero particles) to the behaviour of electrons in the hydrogen atom he obtained results which did not agree with experiment. He therefore abandoned the work and only later realised that if the requirements of relativity are ignored, then, to this rough approximation, his work was in agreement with observation. He therefore first published his non-relativistic equation although he clearly realised from the beginning that his theory was only preliminary, as an intermediate stage to RQT. The reason why Schrödinger's original equation did not apply to electrons in the atom was because the theory did not take into account the fact that electrons have an intrinsic angular momentum, or spin, of $(\sqrt{3}/2)\hbar$. Electron spin is a fundamentally quantum mechanical concept without classical analogue since the spin component can be only be one of two discrete values, $+\frac{1}{2}\hbar$ or $-\frac{1}{2}\hbar$, there being no continuum of values between these two. Change in electron spin can then not be understood classically but only quantum mechanically by way of a 'quantum jump'. The first to suggest that an electron has an intrinsic spin of this nature were G E Uhlenbeck and S A Goudsmit in 1925. This hypothesis was made as a result of a phenomenological analysis of the fine structure spectra of atoms and was further supported by the results of the Stern-Gerlach experiment of 1922. However, although the hypothesis was well founded experimentally, theoretically it appeared as a rather ad hoc addition to QT. Only in 1928, when Dirac proposed his relativistic generalisation of Schrödinger's equation, did electron spin appear as a natural consequence of QT (even though Dirac had not set out to account for the spin of the electron). The discovery by Dirac (1928) of a quantum mechanical wave equation which incorporates the principles of special relativity represents an achievement of cardinal importance for the development of the physics of elementary particles. (For a discussion of the progress towards the unification of general relativity and quantum physics, see eg Komar (1973) and Peat (1973.) The fundamental discovery which underlies the whole of the modern theory of elementary particles is the prediction from the Dirac equation of the existence of antimatter. According to relativistic mechanics, the energy, E, momentum, p, and rest mass, m_0, of a free particle are related by the equation

$$E^2 = c^2p^2 + m_0^2c^4 \quad \text{... } 52$$

where c is the velocity of light. This relation permits two solutions

$$E = \pm|E|, \text{ where } |E| = \sqrt{(c^2p^2+m_0^2c^4)} \quad \text{… } 53$$

one of which corresponds to a positive energy and the other to a negative energy solution. Now, it is clear from equation 53 that a particle with positive energy will have energy $E \geq m_0c^2$ while a particle with negative energy will have energy $E \leq -m_0c^2$. Thus there is a discontinuous gap of width $\Delta E = 2m_0c^2$ in which no particle can exist. Thus, classically, if a particle initially has positive energy then there is no possibility of a transition to a state of negative energy since this would imply existence of the particle for some time in the forbidden zone $m_0c^2 > E > -m_0c^2$. Negative energy states can therefore, according to CM, never be reached from positive energy states and can therefore be ignored without any difficulties arising. In QT, however, transitions can take place discontinuously and thus an electron may 'jump' from one state of positive energy to another state of negative energy. Therefore in Dirac's RQT both positive and negative energy solutions of equation 52 must be considered. This raises a problem for RQT: why do observed electrons of positive energy not spontaneously 'jump' into negative energy states releasing a photon of the appropriate energy? One can easily calculate the transition rate of electrons from positive to negative energy states, assuming the negative energy states to be empty, and show that positive energy electrons should not have the observed stability.

This problem of RQT was not immediately recognised by most authors who suggested that, just as in classical theory, negative energy solutions should simply be rejected. However, in 1950 Dirac suggested an original and far- reaching solution to this problem. Making use of the exclusion principle, which was proposed by Pauli in 1925 in an attempt to formulate a rule concerning the arrangement of the extranuclear electrons in atoms, Dirac suggested that every possible negative energy quantum mechanical state is filled with one electron. Since, according to the exclusion principle, no two electrons can have the same amount of energy and angular momentum, and since, according to Dirac, all the possible negative energy states are occupied, it follows that there can be no transitions of electrons from positive to negative energy levels in agreement with observation. It would seem then that by postulating an infinite 'sea' of negative energy electrons Dirac has solved the problem of the observed lack of transitions of electrons from positive to negative energy levels in a way which is logically consistent with RQT of electrons.

If this were all that was to be said concerning Dirac's solution to the problem of the existence of negative energy solutions in his equation of RQT then one obvious criticism would be its extremely uneconomical invocation of an infinite number of negative energy electrons. One would be tempted to suppose that there must be some alternative solution to the problem whereby it would not be necessary to postulate the existence of an infinite sea of electrons. However, the infinite sea interpretation of the non-existence of transitions from positive to negative energy states does make a definite prediction which is experimentally verifiable. Although it is not possible for an electron to jump from a positive energy state to a negative energy state with the release of a photon since all the negative energy states are filled, it is theoretically possible to induce a transition from a negative energy state to a positive energy state by means of a photon of the appropriate energy ($E \geq 2m_0c^2$). This would leave a 'hole' in the infinite sea of negative energy electrons, which would appear as a particle of positive mass and charge, as well as an ordinary positive energy electron. Dirac originally thought that this positively charged 'hole' might be identifiable as a proton but it was soon shown by Oppenheimer (1930) and Weyl (1931) that the 'hole' had to have a mass equal to the mass of the electron. Therefore Dirac predicted in 1931 as a result of his infinite sea interpretation the existence of "a new kind of particle unknown to experimental physics, having the same mass as and opposite charge to an electron. We may call such a particle an anti-electron." Thus Dirac predicted from RQT the first conceived form of antimatter, the anti- electron.

It is interesting to find that at this time some of the leaders of the orthodox interpretation of QT did not take very seriously Dirac's prediction of the existence of antimatter. Bohr humorously suggested that in order to catch elephants for the Copenhagen zoo one might erect posters which summarise Dirac's theory at watering spots that are frequented by elephants. On reading the poster an elephant, who is a proverbially wise animal, will become spellbound and will easily be captured in this trance-like state (see Gamow 1966 p132). Also Pauli (1933) suggested in his article in the Handbuch der Physik which was written in 1931 but not published until 1933 that the prediction of antimatter was a shortcoming of the Dirac theory. At the end of 1932, however, C D Anderson, who was not familiar with Dirac's theory or his prediction of the anti-electron, discovered a track in cloud chamber photographs of cosmic rays which he interpreted as a particle of the same mass as an electron but with opposite charge. This particle he called a positron and soon it became clear that this was the anti-electron predicted by Dirac.

Must we then accept the idea that 'empty' space is a sea of infinite density consisting of electrons with every conceivable value of negative energy and of which we are even less aware than a fish of the water around it? The answer to this is no. We may, on the contrary, postulate that space is full of positrons in the negative energy states, except for some holes which are the ordinary negatively charged electrons which we observe. This picture would still account for the creation of electron-positron pairs by a photon of sufficient energy since this photon would now be regarded as causing a transition of a positron from a negative to a positive energy state, leaving a hole in the infinite sea which would manifest itself in this case as an additional negatively charged electron. Since the experimental situation with regard to electrons and positrons produced (destroyed) in pair creation (annihilation) is totally symmetrical, there is no way to distinguish between these two theories. Dirac's interpretation is therefore not the only possible one. However, both of the above interpretations suffer from the difficulty involved with the assumption of the enormous charge density contributed by the negative energy states, whether they are filled with electrons or positrons. In order to avoid the asymmetry in charge involved in both these theories, and to restore complete symmetry between positive and negative charges, the modern version of the theory, as developed by Heisenberg (1934), Kramers (1937) and others, assumes that the true result should be equal to the arithmetic mean of the results of both these theories. It then follows that the charge of the vacuum is reduced to zero and that electrons and positrons can be treated theoretically in a totally symmetrical way - the states of one electron or one positron are positive energy states. There is, however, a contingent both asymmetry since there appear to be more electrons than positrons in the universe

The modern theory of antimatter involves a great deal of mathematics and is much more difficult than the original Dirac theory to interpret in terms of a simple picture. Calculation of pair creation and annihilation is made in this theory without reference to the Dirac sea. Thus, only in this theory is there true creation and annihilation of particles, for in the Dirac theory the pair of particles which appears as a result of the absorption of a high energy photon by the vacuum is not created out of nothing but merely appears as a result of the transition of an already existing electron from a negative to a positive energy state. "Conservation of particles ... comes at the price of introducing an infinity of new particles," writes Audi (1973 Ch7). But in the modern theory there is no conservation of particles, either of particles of zero rest mass or of particles of non-zero rest mass (although there is conservation of charge, baryon number and lepton number). This conclusion is of consequence for the concept of wave-particle duality which arose in the CI of elementary QT. In elementary QT it is possible to create and annihilate photons for these have zero rest mass and therefore may be created without a great expenditure of energy which would take us into the relativistic domain. Thus for photons in elementary QT the wave model can be regarded as entirely 'physical' in the sense that emission and absorption of waves (ie creation and annihilation of waves) is an entirely natural feature of the classical wave model. The particle model for photons is also 'physical' in the sense that under the appropriate circumstances photons scatter in a way entirely analogous to classical particles. However, as far as particles with rest mass are concerned in elementary QT, the wave model is not entirely 'physical' since these 'waves' cannot be emitted or absorbed. Thus, in the elementary theory of massive particles, wave-particle duality becomes distorted. But in RQT where conservation of particles is not assumed, complete duality between waves and particles is restored and new symmetry between photons and particles transpires.

There are, in fact, difficulties in the original Dirac sea interpretation which have arisen since its formulation in 1930. Since protons cannot be regarded as the 'hole' in the infinite sea of negative energy electrons, they too must, according to RQT, have associated anti-protons. There should therefore also exist an infinite sea of negative energy protons in order to account for the existence of the anti-proton which was discovered by E Segre and O Chamberlain in 1955. Likewise neutrons, whose existence was predicted by Rutherford in 1920 and which were discovered by Chadwick in 1932, must have an associated infinite sea of filled negative energy states in order to account for the existence of the anti-neutron which was discovered in 1956 - and so on with all the baryons (see Glossary) which have been discovered since then. Thus the vacuum, on the infinite sea interpretation, is not only filled with an infinite number of electrons of negative energy but also with an infinite number of protons, neutrons, etc. This is obviously totally undesirable since none of these infinity of particles is observed, and also since the original concept of the vacuum as 'empty' becomes absurd. It is clearly desirable to eliminate all talk of an infinite sea and treat particles and their associated anti-particles symmetrically - this is what the modern theory attempts to achieve although it lacks a simple conceptual picture of the new situation.

A further, and even more difficult, problem for the original Dirac interpretation is the discovery of particles, and their associated anti- particles, which have finite rest mass and zero spin. The relativistic equation for these particles can be obtained by squaring the Dirac equation for spin ½ particles; this yields the Klein-Gordon solutions, one of positive energy and one of negative energy. The problem here though is that Pauli's exclusion principle does not apply to particles of zero spin. Therefore if one attempts the sea interpretation of the anti-particles of zero spin particles then one finds that one cannot fill the negative energy states since any number of these particles can have the same energy. An anti-particle as a 'hole' in the sea then becomes meaningless since the concept of 'hole' is no longer well defined. According to the modern theory, however, this presents no problem since no mention is made of a sea of filled negative energy states. These problems in Dirac's interpretation do not imply that his interpretation of spin ½ anti- particles is necessarily wrong, but show some conceptual difficulties which are involved in extending this interpretation beyond electrons. Nevertheless, Dirac's model of antimatter is consistent with all the observed behaviour of at least electrons and positrons.

Another model, which was originally due to J A Wheeler (see Feynman 1966 p32), was developed by Feynman in 1949. This model does not involve an infinite sea of filled negative energy states but still allows one to think of the electron and the positron simply as two aspects of one particle. The model is easily extended to other particles besides electrons. The basic idea can best be understood in terms of an important theorem due to Pauli (1955) and Lǜders which was essentially anticipated by Shell in 1948. This theorem is currently known as the TCP theorem.

Theorem 3. The laws of RQT are invariant under the combined operations of charge conjugation (ie changing the sign of all charges), spatial reflection (ie taking the mirror image) and time reversal.

This theorem follows as a logical consequence of the assumptions of RQT and is therefore assumed to be strictly correct. It is interesting to note, however, that prior to 1956 it was assumed that each of these operations individually left the laws of RQT invariant. Thus it was thought, for example, that the mirror image of any physically occurring process was also a process which was physically possible since it was thought that nature does not distinguish between left and right in elementary processes. However, it only follows from RQT that the laws are invariant under the combination of all three symmetry operations and therefore the question of whether they are invariant under these operations taken separately remained logically open. As a result of certain theoretical considerations concerning the decay of kappa mesons, T D Lee and C N Yang suggested in 1956 that in certain processes (ie weak interactions) nature does distinguish between left and right. This speculation was soon verified experimentally by C S Wu et al in 1957. Further, in 1964 it was found that certain processes which are forbidden if the laws of RQT are invariant under the combined operations of charge conjugation and spatial reflection do in fact occur. This implies, from the TCP theorem, that the laws of RQT are not, as was previously supposed, symmetric with respect to the direction of the flow of time. That is, one cannot assume that all processes which take place as time moves forward are also possible if time were to move backwards. It is now believed that all of the above three symmetries taken individually or in pairs may be broken in certain processes; only taken together do they give an unbroken symmetry.

It should now be clear that, as a result of theorem 3, which we make no attempt to prove, the concept of a particle, say an electron, moving from left to right across the page as time increases is exactly equivalent to the concept of a particle of equal and opposite charge moving from right to left across the page in reversed time. A further consequence of the theorem is that this new particle will have a mass exactly the same as that of the original particle. Therefore Feynman proposes to call this new particle a positron. A positron, then, according to Feynman, is an electron moving backwards in space and time. Pair creation might then be understood in the following way: an electron travelling backwards in time and towards some position in space, O, encounters a photon at O which drives the electron forward in time so that it moves away from O in the opposite spatial direction. Observed from our own point of view, that is in terms of a time which continually moves forward, this process will appear as the creation of an electron and positron at O with the disappearance of a photon at O. This picture, although difficult in that it involves the unfamiliar idea of particles moving backwards in time, is consistent with the observations of all processes involving matter and antimatter. Further, the fact that not all processes seem to obey the criterion of time reversal invariance of physical laws already implies discussion of processes going 'backwards in time’.

An advantage of this interpretation of antimatter over that of Dirac is that not only does it dispense with the concept of a vacuum filled with an infinite number of particles but also it enables one to regard an electron and positron (or any other particle and anti-particle) as essentially one and the same. Following 0ckham's principle of economy, Feynman's interpretation is therefore to be preferred to that of Dirac since it involves the postulated existence of the fewest number of particles. Nevertheless, as Feynman (1966 p44) himself points out concerning "the idea of the positron being a backward-moving electron; it was very convenient, but not strictly necessary for the theory because it is exactly equivalent to the negative-energy-sea point of view." Finally as regards this interpretation it should be noted that the question of whether or not there is, really, creation or annihilation of matter (or light since a photon is its own anti-particle) is a matter for debate since this depends on one's view of time: thus, for example, proponents of the 'block universe view' might conclude that nothing is created or destroyed during pair production while proponents of the 'myth of passage view' might conclude the opposite. In the following we shall refer to pair production as the "creation" of a particle and its anti-particle with the "annihilation" of a photon.

Taking the prediction by RQT of the possibility of pair production together with the time-energy uncertainty relation,

$$\Delta E \, \Delta t \geq \hbar \qquad \ldots 54$$

which is still valid in RQT, yields a further ontological consequence of the theory which is of great significance. This is the possibility of the existence of 'virtual' particles. One of the most basic ideas in physics is the conservation of energy. The UP, however, requires an uncertainty in any measured energy of an amount ΔE, and hence allows non-conservation within ΔE provided that this non-conservation lasts for a time interval no greater than Δt where $\Delta t = \hbar/\Delta E$. Thus, for example, a charged particle can continually emit photons of energy ΔE and reabsorb them without violating conservation of energy provided the whole process is over in time Δt. These photons are called 'virtual' since they are in principle unobservable. If they could be observed then conservation of energy would be violated but the UP ensures that these particles cannot be observed. Similarly, a photon may continually create and annihilate virtual electron-positron pairs by transformation of energy into mass and back again. While these virtual particles cannot be observed directly they do have experimentally observable consequences. For example, a particle with zero electric charge may be able to transform into two particles, one with a positive and one with a negative charge:

$$A^0 \rightarrow B^+ + C^- \qquad \ldots 55$$

$B^+ + C^-$ would then constitute a possible virtual pair for A^0. A^0, although itself uncharged, would then be capable of interacting with an electric field through the virtual charged particles B^+ and C^-. One could then calculate the effect of these particles and predict the expected magnetic moment of the uncharged particle A^0. This is precisely what was done with the neutron which was known to be electrically uncharged and yet to have a magnetic moment. The 'anomalous' magnetic moment of the neutron was calculated on the assumption of the presence of virtual protons and negative pi mesons giving results in extremely close agreement with observation. Thus, postulating the presence of virtual particles can lead to testable predictions even though the particles themselves are in principle unobservable. In this way the concept of virtual particles has become acceptable. Although the concept of virtual particles is characteristic of RQT since it would be impossible without the related concept of pair creation, the concept of virtual transitions is familiar from non-relativistic QT and is used, for example, in the explanation of the tunnel effect.

The question arises of the reality of virtual particles. Is there any need for an ontological distinction between 'virtual' and ‘real’ or 'observable' particles? If virtual particles are accorded real existence then this would conflict with the criterion for existence suggested by the logical positivists according to whom a thing exists only if it could in principle be observed. It might also be objected that virtual particles cannot be real because they violate conservation of energy which is 'impossible'. But if the UP is correct and ultimate then it would be impossible to know that energy is always conserved in detail. Further, virtual cannot be understood merely to mean possible because virtual states have observable consequences which would not exist if no actuality was ascribed to them. Since energy conservation cannot be used as an argument against the reality of virtual particles, it has been concluded that it is not necessary to make an ontological distinction between real and virtual particles (see Hesse 1965 p277). Note, however, that in the examples considered above we have only considered simple cases where the number of virtual particles postulated is small. It may be though that the particle B^+ in 55, for example, also can exist for some time in a virtual state containing the two particles D^+ and A^0. Thus,

$$B^+ \rightarrow D^+ + A^0 \qquad \ldots 56$$

The picture becomes further complicated, and the number of postulated 'real', entities increases beyond two. Worse still is the realisation that this multiplication of entities can go on indefinitely since the original particle A^0 has reappeared, and therefore 55 and 56 together can be reapplied in a cyclic fashion to generate an infinite number of virtual particles from the original particle A^0.

This poses obvious conceptual difficulties if one wishes to maintain that virtual particles are ontologically similar to real particles. It also exemplifies a difficulty for a hidden variables interpretation of RQT. If it is conjectured that the original particle A^0 has hidden variables which determine which of the infinite number of possible virtual states A^0 actually adopts in a given period of time then the number of hvs must, according to 55 and 56, in general be infinite since there are an infinite number of potential points at which an effective decision has to be made as to whether a virtual pair is there created (or annihilated) or not, and each of these point decisions is, at least in RQT, independent of all the others. A further difficulty arises for a deterministic hv interpretation of RQT since the vacuum itself can influence the state of A^0 by way of the creation of virtual particles in vacuum. Since, in this case, there seems to be nothing tangible initially to which the hvs could be ascribed, it becomes hard to understand how determinism can be applied in its usual sense to this situation. Of course, it is not necessary for an hvt to restore classical determinism but then hvts begin to look like a minor addition to QT rather than a saviour of a dying philosophy.

There is another phenomenon which suggests that one should believe in the physical reality of virtual particles. This phenomenon is the observed interaction between particles, such as the Coulomb force between charged particles. According to RQT, the force which arises between two charged particles, two electrons say, can be explained with great accuracy in terms of the emission and absorption of virtual photons passing between the two particles. Thus the simplified picture of the attractive force between oppositely charged particles emerges wherein attraction is produced as a result of the passage of virtual photons back and forth between the two particles binding them together. This is a simplified picture since much more complicated virtual states than single photons are also possible, according to the theory. Nevertheless, the great significance of the postulated existence of virtual particles can now be seen in that they enable one to treat interactions and forces between particles in terms of the theory of the particles themselves and without the introduction of further postulates. In this way, the theory of quantum electrodynamics (QED) can not only explain the electromagnetic interactions between charged particles (Coulomb scattering) but also the interactions occurring in the scattering of photons with charged particles (Compton scattering) and of photons with photons (which is not allowed in Maxwell's theory but is in Dirac's theory).

Although very successful in its own domain (the domain of leptons and photons), QED can in no way account for the interactions binding together the particles in the atomic nucleus. For example, all charged particles in the nucleus (protons) have the same charge (positive); they should therefore strongly repel one another since they are very close together (~10^{-13} cms). Instead they are held together by some other force which can overcome the electromagnetic repulsion. Further, when in 1919 Rutherford managed to split the nuclei of nitrogen atoms by bombarding them with high energy alpha particles (helium nuclei) it became clear that very high energies were required to cause the disintegration and thus it was shown that the forces holding the nuclear particles together are much greater than electromagnetic forces. Before Rutherford's experiment only alchemists believed in transmutation of the elements. Were the forces between nuclei substantially less (say the strength of electromagnetic forces) then transmutation would be commonplace. Another characteristic of the nuclear force (or strong interaction as it is now called) is its short range. From the experimental fact that protons only interact with the nucleus when they are of sufficiently high energy to overcome the electromagnetic potential barrier and come extremely close to the nucleus (~10^{-13} cms) the short range nature of the strong interactions soon became apparent. Finally, electromagnetic forces again seem unable to account for this interaction since nuclear forces are known to operate between the neutral neutron and proton.

As we have seen, according to QED, the Coulomb force results from exchange of photons between charged particles. A natural extension of this idea to the problem of nuclear forces is to assume the existence of a new particle which produces the nuclear potential. This idea was propounded in 1935 by Yukawa who hypothesised the existence of a so-called 'heavy quantum' to account for the observed properties of the nuclear force. He reasoned as follows when there is a charge at the origin, O, the Coulomb potential φ satisfies Laplace's equation,

$$\Delta\varphi = 0 \text{ except at the origin,} \qquad \text{.... } 57$$

where $\Delta \equiv d^2/dx^2 + d^2/dy^2 + d^2/dz^2$. Or one can look for the static spherically symmetric solution of the equation,

$$\Box\varphi = 0 \text{ except at the origin,} \qquad \text{... } 58$$

where $\Box \equiv c^{-2}d^2/dt^2 - d^2/dx^2 - d^2/dy^2 - d^2/dz^2$. The solution is given by the well known Coulomb potential

$$\varphi \propto 1/r \qquad \text{... } 59$$

where r is the distance from the origin, O. In order to obtain a short range force, Yukawa (1935) modified equation 58 to read,

$$\Box\varphi - K^2\varphi = 0 \text{ except at the origin,} \qquad \text{… } 60$$

where K is some constant. The static spherically symmetric solution is then given by

$$\varphi \propto e^{-Kr}/r \qquad \ldots 61$$

which describes a short range potential that falls off approximately exponentially with increasing distance from the origin rather than linearly as in the case of the Coulomb potential of equation 59. From the experimental information on nuclear forces it is possible to estimate the order of magnitude of K:

$$K \sim 5.10^{12}\ \mathrm{cm}^{-1} \qquad \ldots 62$$

This is a classical theory of nuclear forces.

In QT, however, one has to quantize the wave φ. One then finds a new kind of quantum particle associated with φ. This quantum was called the 'heavy quantum' by Yukawa as opposed to the 'light quantum' or photon of QED. The energy, E, and momentum, p, of a heavy quantum satisfy Einstein's relation,

$$E^2 - p^2c^2 = (Mc^2)^2 \qquad \ldots 65$$

where M is the rest mass of the new particle. Following; the usual rules, the wave equation for φ is obtained from equation 65 by the replacement

$$E \rightarrow -(\hbar/i)d/dt, \quad p_x \rightarrow (\hbar/i)d/dx, \text{ etc.}$$

This yields the wave equation (which is actually the Klein-Gordon equation for spin zero particles),

$$[\square - (Mc/\hbar)^2]\ \varphi = 0 \qquad \ldots 64$$

(Notice that putting M = 0 in 64 gives 58, the equation for the photon.) Now if we identify 60 with 64, we get

$$K = Mc/\hbar \quad \text{or} \quad M = K\hbar/c \qquad \ldots 65$$

From 62 and 65 we can compute the mass of the new particle giving

$$M \approx 200\ m_e \qquad \ldots 66$$

where m_e denotes the electron rest mass.

The resulting picture of the interaction between protons and neutrons in the nucleus, according to Yukawa's theory, is in terms of the exchange of virtual particles of mass approximately 200 times the electron rest mass between protons and neutrons. Since these particles are massive in contrast with the photon, much more energy is required to create them. Therefore, if they are to be the virtual particles accounting for the nuclear interactions, the UP requires that they can only exist for a time Δt which is much shorter than the corresponding time for virtual photons. Further, these particles have rest mass and therefore cannot travel at the speed of light. Nuclear interactions are therefore short range as opposed to electromagnetic interactions which are long range. Originally, Yukawa suggested that these particles have a charge (and the anti-particles an opposite charge). Thus a positively charged virtual particle might depart from a proton, leaving a neutron in its place, and arrive at a neutron which would then become a positively charged proton. Alternatively, the negatively charged anti-particle might depart from a neutron, leaving a positively charged particle - a proton - and arrive at a proton giving a neutral particle - a neutron. However, this does not explain the forces between neutron and neutron or between proton and proton. In order to explain this N Kemmer in 1938 postulated the existence of a further particle of the same mass but of zero charge.

Meanwhile, the search was on for particles which fit Yukawa's description. In 1937 a particle of approximately the correct mass was discovered by C D Anderson and S H Neddermeyer in cosmic radiation. For a number of years after this discovery it was generally thought that this particle together with its anti-particle were the particles predicted by Yukawa's theory. These particles were charged and were called mu mesons: they were called 'mesons' (from Greek μέσος - middle) since their mass was in between that of electrons which were called ‘leptons’ (from Greek λεπτός - light) and neutrons and protons $m \approx 2000\ m_e$ which were called 'baryons' (from Greek βαρύς - heavy). Later mu mesons were reclassified as leptons and renamed 'muons'. The attempt to equate muons with the particles predicted by the Yukawa theory led to a number of contradictions the most serious of which was that muons seemed able to pass right through atomic nuclei without anything drastic happening to them, whereas, if Yukawa's theory is correct, they ought to be absorbed very readily by nuclei, their mass being converted into energy in the process. These difficulties were resolved only in 1947 when G F Powell et al discovered the existence of another particle slightly heavier than the muon in high altitude cosmic rays. These new particles were called pi mesons; they interacted very strongly with nuclei and altogether behaved in accordance with ideas of Yukawa. There are three different pi mesons of this mass now known; there is a positively charged pi meson and its anti-particle, a negatively charged pi meson and its anti-particle and a neutral pi meson which is its own anti-particle. These discoveries confirm the soundness of the ideas and predictions of Yukawa and Kemmer. The muons discovered in 1937 are still not understood although it is now conjectured that they might be an excited state of an electron since all their properties except mass seem to be exactly the same as those of an electron, but this would imply that electrons have internal structure.

Thus the RQT of electromagnetic interactions as well as of strong interactions has been extremely successful in accounting for atomic and nuclear phenomena and, according to Schwinger (1966 p29), the theory "has failed no significant test, nor can any decisive confrontation be anticipated in the near future." Nevertheless, the theory faces grave logical difficulties. The trouble is that the theory has never been shown to be self-consistent - indeed around 1935 it was thought to be manifestly inconsistent. The physical origin of these difficulties centres round the existence of virtual particles. Even a free particle which is not interacting with any other has to be considered as being surrounded by a 'cloud' of virtual particles which the particle is constantly emitting and reabsorbing. This 'self-interaction' implies that, according to the theory, the real mass, m, and charge, e, of the particle differ from their observed values by some amounts Δm, and Δe. So a particle actually having a mass, m, and a charge, e, will appear to have a mass, $m + \Delta m$, and a charge, $e + \Delta e$, where the correction terms Δm and Δe are due to the surrounding cloud of virtual particles. But, according to the theory, Δm and Δe are infinitely great, suggesting that the original m and e are minus infinity. Now, obviously, with these infinities appearing out of the mathematics it is impossible to make any logical sense of the theory as it stands. For this reason, Feynman (1966 p43) says: "I don't think we have a completely satisfactory relativistic quantum-mechanical model, even one that doesn't agree with nature but, at least, agrees with ... logic."

However, despite this mathematical impasse, in 1947 H A Bethe surmised that if one substituted the experimental values of mass and charge for the corresponding infinite constants wherever they appear in the theory then the theory might yield convergent results which agree with experiment. This procedure developed principally by Feynman and Schwinger is called renormalisation of mass and charge, and has brought the possibility that the theory will lead to finite results by renormalisation even if it contains defects. Further, Dyson (1949) has shown that all the infinities appearing in QED can be treated by the renormalisation procedure to an arbitrarily high order of approximation. Surprisingly enough, this method of renormalising QED and removing the inherent infinities yields results which are in excellent agreement with the experimental facts. Although, of course, the theory cannot give predictions concerning the mass and charge of the electron, it does predict such things as the various scattering formulas for electrons and photons and the observed relativistic fine structure spectrum of the hydrogen atom.

Despite the immense practical successes of renormalised QED, the unsolved question still remains as to whether the mathematical basis of the theory is sound. So far renormalised QED has defied all attempts of the mathematician to make it sound. It is therefore sometimes conjectured, for example by Dirac and Feynman, that the theory is not sound and that its remarkable success should be looked on as a fluke. Thus Feinberg (1972 p42) writes: "Because of this mathematical impasse, it remains unclear whether the principles of relativistic quantum mechanics are sufficient to describe elementary particle systems, or whether some additions or modifications of them will be necessary to do this. Probably most physicists, including myself, tend toward the latter view, but there is hardly any indication yet as to the direction we must follow in order to reach the new theory." The inadequacy of renormalisation methods is particularly striking in the case of nuclear interactions. Renormalisation techniques applied to the RQT of nuclear interactions (cf QED) do not always yield convergent results. The reason for the greater difficulty in the strong interactions can be understood as follows. For electrons, according to QED, the probability that one virtual photon is emitted and reabsorbed is small and directly related to the value of the fine structure constant, $e^2/\hbar c = 1/137$. The probability of emitting two photons simultaneously is proportional to $(1/137)^2$, three to $(1/137)^3$, etc. Thus, only very rarely are two or more simultaneously present in the surrounding cloud; in most cases no photon at all is emitted. However, the corresponding constant for strong interactions is very much larger, probably greater than 1. As a result we must expect that a proton or neutron will just as readily produce two, three or more pi mesons as it will produce one. This obviously complicates a detailed description of what is going on. The description, in fact, gets so complicated that accurate predictions are not always possible, even by renormalisation methods. Thus the conventional RQT of the dynamics of strong interactions is not only possibly logically unsound but also is not generally successful in giving correct predictions. In the next section we therefore consider a new approach which has been developed for the strongly interacting particles, which are called hadrons and include all the mesons and baryons but not the leptons or the photon. This new approach has been developed in the last 20 years although it was first suggested by Heisenberg (1943) much earlier as a possible means of avoiding the divergence difficulties of quantum field theory, which at that time had not been solved by renormalisation.

II

The new approach to strong interactions is called the scattering-matrix or S-matrix approach. The S-matrix is defined as the operator which transforms the initial state, written in Dirac notation as $|in\rangle$, into the final state, written as |out>. Thus,

$$|out\rangle = S\,|in\rangle \qquad \text{... 67}$$

Let us first show a connection between equation 67 and those of elementary QT beginning from the time dependent Schrödinger equation,

$$i\hbar d\psi/dt = H\,\psi \qquad \text{… 68}$$

where H is the Hamiltonian operator. The differential equation 68 can be transformed into the integral equation,

$$\psi(t) = \psi(t = -\infty) - \frac{i}{\hbar}\int_{-\infty}^{t} dt_1\, H(t_1)\,\psi(t_1) \qquad \text{… 69}$$

Solving this equation by iteration then gives,

$$\psi(t = +\infty) = \left[\sum_{n=0}^{\infty}\left(-\frac{i}{\hbar}\right)^n \frac{1}{n!}\int_{-\infty}^{+\infty} dt_1 \ldots \int_{-\infty}^{\infty+} dt_n\,\left(H(t_1)\ldots H(t_n)\right)\right]\psi(t = -\infty)\text{… 70}$$

where ψ(t = -∞) is the wave function of the system at time t = -∞, and can therefore be interpreted as the initial state of the system while ψ(t = +∞) is the wave function of the system at time t = +∞ , and can therefore be interpreted as the final state of the system. Comparing now 67 with 70, we find the equivalent of the S-matrix in elementary QT is an infinite series of progressively more complicated terms (the term in square brackets). Thus, if the appropriate form is taken for S then the new approach is capable of yielding all the predictions of non-relativistic QT. Notice, however, that equation 67 deals exclusively with the asymptotic states, t = ±∞, while equation 68 deals with states at finite times, ψ(t). One might therefore suspect that elementary QT can give predictions at finite times which cannot be duplicated by the S-matrix approach. It has been shown, however, for example by Stapp (1971 Appendix A), that "one can calculate the transition probabilities corresponding to preparations and measurements that take place at finite times in terms of asymptotic states." Therefore the S-matrix and Schrödinger formulations give equivalent results and elementary QT makes no predictions beyond those made by the S-matrix approach.

But there is a sense in which the S-matrix is logically superior to the Schrödinger formulation. Since the prepared and measured systems are in principle fully distinguishable from the preparing and measuring devices only in the asymptotic limit, asymptotic states are the only ones which are in principle well defined.

In section I we found that the relativistic generalisation of QT leads to unsurmounted mathematical difficulties. In terms of the relativistic analogue of equation 70 these difficulties can be expressed as the fact that some of the terms in the infinite expansion turn out themselves to be infinite. Thus the technique of dealing with interactions in RQT by means of perturbations giving rise to infinite expansions becomes suspect. The question therefore arises; can one define an S-matrix in terms other than perturbation expansions? The S-matrix approach begins from the assumption that it might be possible to calculate the elements of the S-matrix directly, without the use of an underlying equation describing the detailed time-dependent evolution of the state of a system, by requiring them to have some general properties that ought to be valid. We shall now consider some of the properties which have been imposed on the S-matrix and then view some resulting conceptual implications of this approach. (We have refrained from calling the S-matrix approach a theory since this would suggest a degree of finality and mathematical rigour which is as yet lacking Nevertheless, much progress has been made in recent years towards a fully fledged theory by demanding that the S-matrix elements have a few well-defined properties.)

The first important property which is to be satisfied by the S-matrix is the superposition principle (SP) of QT. Consider, for example the initial state, written |in>, of two particles that subsequently come together, interact and separate. The SP then implies that the final state, |out> , can be written S |in> , where S is a linear operator. Thus, the SP imposes the condition that the operator S is linear.

A second important property which must logically be satisfied by the S-matrix is the condition that must be imposed in order to ensure that the total probability for an arbitrary final state to arise from some initial state is unity, ie if something; goes 'in', one of the possibilities must come out The condition which this imposes on the S-matrix can be derived as follows. Given that the set of states |n> is orthonormal and complete, ie

$$\langle m|n\rangle = \delta_{mn} \quad , \quad \sum_m |m\rangle\langle m| = 1 \qquad \text{… 71}$$

any state can be expressed as a superposition of the states |n>. Thus an arbitrary state |x> can be written as

$$|x\rangle = \sum_n a_n\, |n\rangle \qquad \text{… 72}$$

where $|a_n|^2$ is the probability that the system ends up in the state |n>. The condition that the total probability is equal to unity can then be written as,

$$\sum_n |a_n|^2 = 1 \qquad \text{… 73}$$

Now, given that |m> is equal to S|n>,'it follows from 72 that

$$a_m = \langle m|S|x\rangle \qquad \text{… 74}$$

Combining 75 and 71 gives

$$\begin{aligned} 1 &= \sum_m |\langle m|S|x\rangle|^2 \\ &= \sum_m \langle x|S^\dagger|m\rangle\langle m|S|x\rangle \\ &= \langle x|S^\dagger S|x\rangle \text{ , from 20} \\ &= \sum_{nn'} a_{n'}{}^* a_n \langle n'|S^\dagger S|n\rangle \text{ , from 72} \qquad \text{… 75} \end{aligned}$$

In order for 75 to hold for all a_n it is necessary that

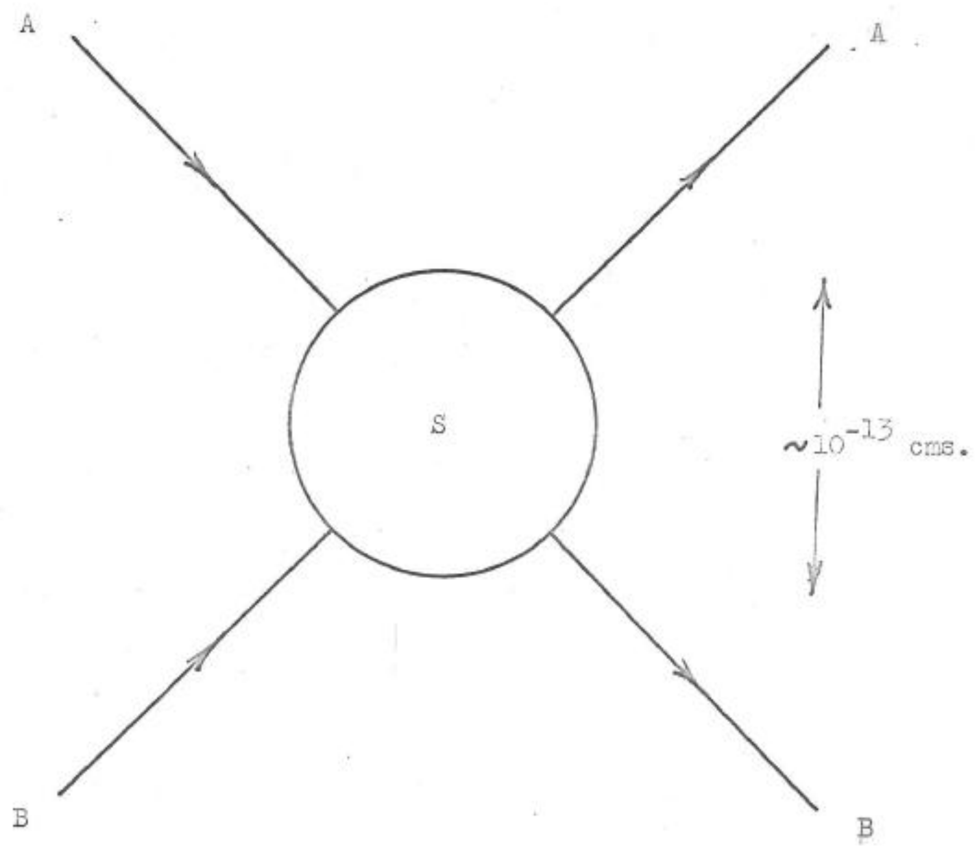
$$\langle n'|S^\dagger S|n\rangle = \delta n_{n'n}$$

which implies that

$$S^{\dagger}S = 1, \quad \text{or} \quad S^{\dagger} = S^{-1} \qquad \dots 76$$

Equation 76, then, is a mathematical property of S related to the conservation of probability, namely that S is unitary.

Figure 6



Strong Interaction: A + B → A + B

The next property which we wish to impose on the S-matrix is the requirement of special relativity. The principle of relativity, as applied to the transition probabilities determined from the elements of the S-matrix, asserts that these probabilities should be invariant for all uniformly moving observers. This again leads to further constraints on the general form of S; it also implies conservation of energy, momentum and angular momentum.

Another property which is imposed on the S-matrix is the property of causality. It is, of course, well known that QT, in contrast with classical mechanics, is only statistically determinate. However, QT is still causal in the sense that according to Schrödinger's equation probabilities evolve in a causal way in an undisturbed system. It has been found very difficult to derive rigorously the consequences of this property for the S-matrix, yet at the same time it is one of the most important for the constitution of the S-matrix. In view of the difficulty, one does not, in general, attempt to impose this property directly but instead one imposes a property which is believed to be a requirement of the causality property. This secondary property is that the transition amplitudes (whose square yields the transition probabilities) are analytic (ie smooth) functions of particle energies and momenta (analyticity can be regarded as the mathematical expression of causality).

Let us now consider an important property of scattering processes which it will be necessary to refer to later in this chapter and which follows from the analyticity and unitarity of the S-matrix. This property, which was first derived in the S-matrix approach in 1954 by M Gell-Mann and M L Goldberger, is called 'crossing symmetry' and is inherent in Dirac RQT as a result of theorem 3 and seems to be confirmed by experiment. According to crossing symmetry, scattering amplitudes have a certain symmetry with respect to the interchange of particles and anti-particles. We shall explain the concept of crossing symmetry by reference to the Dirac interpretation of antimatter. Consider the scattering reaction of two particles A and B where the initial state, |in>, can be regarded as the initial state of the two particles moving towards some region, O, and where the final state, |out>, can be regarded as the final state of A and B moving away from O. This situation is symbolised by the graph in Fig.6. Assuming the initial four-momentum of A to be p_1 and of B to be p_2, and the final four-momentum of A to be p_1' and of B to be p_2', we then have,

$$\begin{matrix} A & + & B & \rightarrow & A & + & B \\ p_1 & & p_2 & & p_1' & & p_2' \end{matrix} \quad \ldots \quad 77$$

where we have indicated the four-momenta of the particles. Now, according to the Dirac theory of antimatter, it is completely equivalent to consider the effects of a positron with momentum k, positive energy E, or that of the balancing electron with momentum -k, energy -E. One can also say that the electron of the Dirac sea with four-momentum (-E,-k) and its 'hole', the positron with four-momentum (E,k), are together equivalent to nothing (the vacuum). Thus, if we suppose that in the initial state of the above reaction 77 there is a pair consisting of particle B(p_2') and anti-particle $\underline{B}$(-p_2') and in the final state there is also a pair B(p_2), $\underline{B}$(-p_2), then, according to the Dirac theory, that does not amount to a real modification of these states. We thus get the reaction

$$\begin{matrix} A & + & B & + & B & + & \underline{B} & \rightarrow & A & + & B & + & B & + & \underline{B} \\ p_1 & & p_2 & & p_2' & & -p_2' & & p_1' & & p_2' & & p_2 & & -p_2 \end{matrix} \quad \ldots \quad 78$$

Now, since we find the same states of the two B's in the initial and final state, 78 is nothing but the reaction,

$$A + \underline{B} \rightarrow A + \underline{B} \qquad \ldots \qquad 79$$
$$p_1 \quad -p_2' \qquad p_1' \quad -p_2$$

Since the reactions 78 and 79, and hence the reactions 77 and 79, are equivalent, the amplitudes for these reactions are equal. This property of the elements of the S-matrix is called crossing symmetry. In fact, following the above argument, there are altogether six different reactions which are 'crossing symmetric' and therefore have the same amplitude. These reactions are,

$$A + B \rightarrow A + B\ ,\ A + \underline{B} \rightarrow A + \underline{B}\ ,\ \underline{A} + B \rightarrow \underline{A} + B\ ,$$
$$\underline{A} + \underline{B} \rightarrow \underline{A} + \underline{B}\ ,\ A + \underline{A} \rightarrow B + \underline{B}\ ,\ B + \underline{B} \rightarrow A + \underline{A} \qquad \ldots 80$$

In terms of the Feynman interpretation of antimatter, crossing symmetry is an expression of the simple rule that in any reaction it is possible to replace a particle on one side of the reaction by its anti-particle with opposite four-momentum on the other side which follows from the view that antimatter is matter going backwards in space-time. We have thus shown that crossing is a property inherent in conventional RQT. It is, however, also possible to derive this property in the S-matrix approach. This yields a theorem which is closely related to theorem 5.

Theorem 4. The amplitude for a process obtained by 'crossing over' a particle into its anti- particle is equal to the amplitude for the original process.

So far we have introduced three principal properties of the S-matrix: unitarity, which in fact incorporates the SP since S is taken to be linear, relativistic invariance and analyticity. While each of these properties would seem to be necessary ingredients of any theory of comparable power to the equations of RQT, unfortunately, as yet, no general S-matrix theory has been constructed which possesses all three properties. Various reasons for this failure have been suggested. For example Aronov (1971) argues that the principles of QT and of relativity are contradictory and therefore one could not expect to construct a consistent theory involving both these principles. W Chew (1970 p23; 1971a p2333) on the other hand argues that there might only exist one S-matrix in equation 67 which is consistent with the observed phenomena (ie there exists only one transformation in strong interactions from input to output) and if this was the case then the failure of physicists to construct a consistent S-matrix theory would be understandable on the grounds that the observed phenomena "are too complex to be encompassed by any explicit construction of human imagination" if there is only one logically possible S. We shall consider Chew's point of view presently.

One final constraint on the S-matrix approach which we shall mention is that the approach is only applicable to short range forces, ie strong interactions but not electromagnetic interactions. The reason why electromagnetic interactions are not capable of being treated by an S-matrix approach is, perhaps, not immediately obvious since, as we have seen, QED can be expressed in terms of an S-matrix which takes the form of an infinite power series. Further, since the expansion coefficient of this series is small (= 1/137) all but the first few terms of the series can be ignored giving results which are in excellent agreement with experiment. The problem, however, for the S-matrix approach, which attempts to avoid power series expansions that introduce infinities and have to be renormalised, centres round the zero mass of the photon. Troubles arise in discussions of the unitarity and analyticity of the S-matrix when the concept of zero mass particles is introduced. Even the concept of an asymptotic state becomes imprecise in this connection because the long range nature of electromagnetic forces means that they cannot be neglected even at times approaching $\pm\infty$. Consequently:

(i) From a thoroughly S-matrix point of view it is entirely baffling why the superficial S-matrix description of electromagnetic interactions involving only the first few terms in the infinite expansion should yield results which are of such amazing accuracy.

(ii) The only particles which can be properly treated by the S-matrix approach are hadrons (there exist no zero mass hadrons).

(iii) In practice, hadron interactions are treated by means of an S-matrix approach while electromagnetic interactions are treated by means of QED.

Having considered a number of principal properties of the S-matrix, we now ask: what are the conceptual implications of this approach? The predictions of the S-matrix approach are tested by way of experiments on the asymptotic states of hadrons involved in localised scattering reactions. In terms of the tetrode in Fig.6, where, for simplicity, we consider two particle initial and final states, these asymptotic states are, to good approximation, represented by the lines outside the circle. Inside the circle the localised scattering process takes place which is described by the S-matrix - the process is localised since we are dealing with short range interactions (the diameter of the circle might therefore be considered to be to good approximation of the order of nuclear dimensions, $\sim 10^{-15}$cms). Since no measuring instrument based entirely on nuclear forces has ever been constructed, if indeed one can be conceived, all measurements can only be made on approximately asymptotic states of nuclear scattering processes. Thus the region inside the tetrode circle is, in principle, out of reach of experimental observation. That is, the details of any process involving short range forces can only be determined in terms of asymptotic states of reacting particles, the region outside the circle, while the region inside the circle cannot be subjected to direct observation.

This situation can be clarified by way of a simple example. Suppose one is in a helicopter following a bank robber in a red car on a road below. Another similar car, which we assume is indistinguishable from the first from the helicopter's point of view, enters from a side road. So long as the movements of the first car are traced with sufficient care, one will never become confused with regard to which car one is supposed to be following. I shall call this a simple illustration of the 'principle of hot pursuit' which states that as long as one has some means for following the exact movements of an object one can describe with certainty all its actions. But now consider the case where, after the second car has entered the main road, both cars pass into a tunnel under a hill. Eventually they reappear and part on different routes. Since the two cars might have changed places in the tunnel (one may overtake the other) it is not known which car one should follow in order to catch the robber when they part on different routes. In order to be certain that the correct car is being followed two helicopters are required since both cars must be followed, not knowing which one is wanted. I shall call this a simple example of the 'principle of cold pursuit' which states that if at some stage one cannot have the means to follow exactly the movements of an object at all times, then when one recovers the means one must take into account every possible occurrence which could take place during the interruption.

Applying these principles to physics, we find that classical mechanics is an example of a theory that satisfies the principle of hot pursuit since there is no situation where in principle one cannot follow the movements of objects in detail. In hadron physics, however, the principle of cold pursuit comes into play in the region inside the tetrode circle where strong interactions are involved. In this region the detailed movements of hadrons cannot be observed since no means of observation exists, according to the S-matrix approach. Thus the theoretical description of hadrons leaving this region must take into account every possible occurrence which might have taken place within the region.

Some authors go further (eg Chew 1963; 1971b) and argue that in the S-matrix approach the very concept of microscopic space-time should be given up entirely (cf inside the tunnel) since it cannot be given a meaning within the tetrode circle. In this sense the region within the tetrode circle in the S-matrix formulation might be compared with the ‘absolute elsewhere' of Minkowski space-time wherein limits are set to the physical meaning of spatio-temporal concepts. In relation to the view that the concept of microscopic space-time should be abandoned in the S-matrix approach it is of interest to note that the philosophy which led Heisenberg to the initial consideration of the S-matrix can be traced to his belief (see Heisenberg (1938)) that the divergence difficulty inherent in RQT might be avoided by the introduction of a new fundamental constant having the dimensions of space (ie quantising space). In this way the usual conception of microscopic space-time would doubtless have to be abandoned in favour of some fundamentally different conception. However, no such theory has yet been constructed in detail. Nevertheless, the concept of the S-matrix introduced by Heisenberg connects the input and output of a scattering experiment without seeking to give a localised description of the intervening events. In this way microscopic space-time does lose its meaning (by becoming redundant) in the S-matrix approach (see also Stapp 1971 p1315).

Another concept which can be abandoned in the S-matrix approach is the concept of a ‘master equation of motion’ that describes in detail the movements and interactions of hadrons. In conventional RQT the equation playing this role is the Dirac equation. But now the Dirac equation is abandoned in favour of the S-matrix which is a collection of all possible particle reaction amplitudes and thus encompasses all conceivable hadron experiments. In the theory of mechanics the Hamiltonian is the generator of the system motion with time. In his original paper, Heisenberg (1943) argued that the assumption of a Hamiltonian implies the possibility of a continuous time displacement of the wave function which seems to contradict the existence of a fundamental length. Heisenberg therefore abandoned the notion of a Schrödinger equation and of a Hamiltonian. He believed that in a future theory the S-matrix would take over the role played by the Hamiltonian. Thus, in the S-matrix formulation, there is no equation of motion that determines the detailed time development of the initial state into the final state. Instead, the transition amplitude of asymptotic input to output is determined by the S-matrix which is derived directly from its general properties. According to Chew (1971a p2334): "The recognition that equations of motion are unnecessary for predicting hadron behaviour has been a powerful spur to the bootstrap idea." We therefore now turn to an examination of the 'bootstrap' interpretation of hadrons which arises from the S-matrix approach to strong interactions.

III

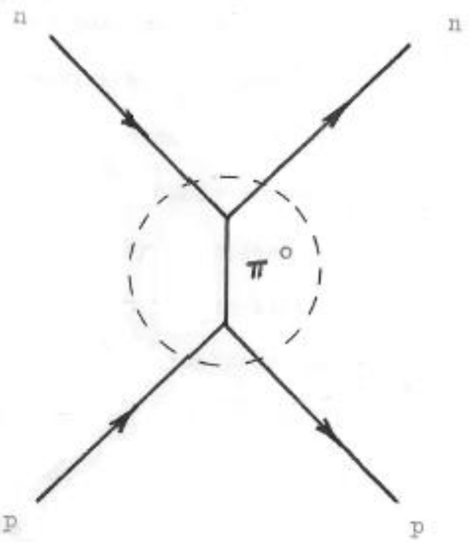
In 1930 only three fundamental particles, the electron, the proton and the photon, had been observed in nature. In the next few years this number increased due primarily to a study of the particles produced in cosmic radiation. However, after the end of the Second World War a number of large accelerators were built and this led to the discovery of a profusion of new and unexpected particles so that by 1957 the number of known particles had increased to about 30. During the next six years many high energy accelerators were built and put into operation in Europe, Russia and America. Thus by 1964 the number of known particles had increased to about 100. The reason why none of these new particles had been observed earlier is because they are, for the most part, more massive than protons and therefore require large amounts of energy to produce them. Energies of this amount were not available before the construction of large accelerators. Also, most of these particles are highly unstable, decaying into lighter particles in a time of the order of 10^{-23}sec. Techniques used to detect particles which only last for such a short time have only recently been developed.

Around this time, therefore, it became clear that not all of these particles could be regarded as fundamental. Attempts were therefore made to treat some of these particles as composites of others and thus reduce the number of fundamental (non-composite) particles to a manageable number. An example of a particle which was thought to be a composite is the hydrogen atom. By assuming it to be composed of one proton and one electron its observed properties had been predicted with great accuracy. However, according to RQT, a hydrogen atom is so composed only some of the time. For a small fraction of time it might be composed of a proton, an electron and an electron-positron pair. This possibility has been taken into account and yields the observed fine structure of the hydrogen spectrum which had been inexplicable in terms of the simpler composite model. The concept of a composite is therefore somewhat complicated by RQT. Another particle that is generally regarded as a composite is the deuteron which is supposed to consist of one proton and one neutron. Due to the fact that nuclear forces hold the proton-neutron composite together, this simple model is less accurate than that for the hydrogen atom since large numbers of virtual pi mesons may be involved (the infinite expansion for strong interactions cannot be approximated to the first few terms as can be done with electromagnetic interactions due to the large value of the nuclear expansion coefficient). Nevertheless, predictions made on the basis of this simple model yielded qualitative success which it was assumed could be improved upon by considering more complex virtual states. Taking the operational definition that a particle is not fundamental if all its properties can be calculated in principle by treating it as a composite, the deuteron and indeed all the nuclei of atoms heavier than hydrogen were therefore considered as non-fundamental composites.

However, a neutron can be considered as being composed for part of the time of a proton and negative pi meson, and part of the time as more complicated compositions. A calculation on this basis yields the same qualitative agreement as with the deuteron. Thus the neutron too might be regarded as a non- fundamental composite. This type of argument can be applied in principle to all the strongly interacting particles (although in practice the problem of including all significant possible composite structures is in most cases still too difficult). Thus none of the hadrons need be considered as more fundamental than any other since all of them can be treated as composites. (Electrons cannot be treated in this way since there are no known forces strong enough to hold any other known particles together with enough of their mass converted into binding energy leaving a total mass as small as that of the electron.)

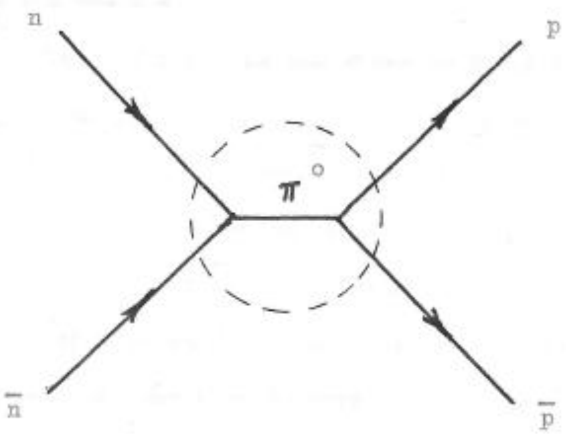
All of this would be a consequence of conventional RQT if it were possible to apply the mathematics rigorously to every particle and hence accurately deduce all their properties by considering them to be composites of other particles. However, in terms of the infinite expansions involved here, the number of possible composite structures becomes too large to be mathematically manageable. We therefore turn to the S-matrix approach which avoids perturbation techniques. In particular we consider the implications of theorem 4 for the concept of a particle.

Figure 7



Interaction $n + p \rightarrow n + p$ involving exchange of a pion.

Figure 8



Interaction $n + \bar{n} \rightarrow p + \bar{p}$ involving a resonance pion.

Consider the reaction involved in the scattering of a neutron and a proton,

$$n + p \rightarrow n + p \qquad \ldots\ 81$$

To a first approximation this scattering can be regarded as taking place by way of the exchange of a virtual pi meson. Thus all the possible reactions taking place in the tetrode circle can in this case be simplified to the single reaction symbolised by Fig.7. According to theorem 4, this same figure is also applicable to the reaction in which particles are crossed over into anti-particles. One such crossed reaction is the following

$$n + \underline{n} \rightarrow p + \underline{p} \qquad \ldots\ 82$$

The crossed equivalent of Fig 7 is now given by Fig.8 for reaction 82. In this case the pi meson is no longer a particle exchanged between the two particles but is instead an intermediate composite state consisting of both initial particles together. Thus an exchanged particle in one reaction becomes an intermediate particle in the crossed reaction.

From Yukawa theory we know that a force can be considered as the exchange of a particle. We thus find that a force in one reaction is related by crossing to an intermediate particle in another reaction. Now, if one assumes that the constraints of relativistic invariance, unitarity and analyticity, lead to a unique S-matrix then it follows that a force of given definite characteristics will entail the existence of an intermediate particle in the crossed reaction whose characteristics are determined uniquely. Similarly, the existence of an intermediate particle in a given reaction will entail a force (and its related particle) of definite characteristics in the crossed reaction. Only one particular force is consistent with the existence of a particular intermediate particle in the crossed reaction and vice versa. This situation leads to the possibility of predicting the properties of certain hadrons by means of self- consistency arguments alone. Such a calculation was first made by G F Chew and S Mandelstam in 1960 for the case of the rho meson.

Assume that the universe consists of only two kinds of particles, the pi and rho mesons (ie the effects of all other particles can be ignored). The force between two pi mesons can be regarded as a result of the exchange of virtual rho mesons in the same way as the force between nucleons is, according to Yukawa theory, the result of the exchange of virtual pi mesons of Fig.7. However, the force thus produced is strong enough to produce another particle which is a virtually bound two pi meson state and this state has the quantum numbers of the rho meson cf Fig.8. Thus the quantum numbers associated with the force between two pi mesons are related by means of an S-matrix which makes the appropriate transformations to the quantities in the crossed reaction to the quantum numbers associated with the intermediate particle in the crossed reaction. Since these sets of quantum numbers are identical because the same particle is involved in both situations, this leads to a self-consistency problem in which the properties of rho must be such that it produces an interaction between pis which leads to a rho. Writing the quantum numbers of the rho as λ one could then in principle determine λ simply by looking for self-consistent solutions of the equation

$$\lambda = S\,\lambda \qquad \ldots\ 85$$

where S is the appropriate S-matrix. By this method Chew and Mandelstam predicted the mass of the rho meson to be 685 Mev. Experimentally its mass is known to be 765 Kev. Considering the initial approximation of assuming the existence of only two types of particle, this prediction can be regarded as modestly successful.

A second, more complicated application of self-consistency arguments has been given by Chew in 1962 to predict the mass of the delta baryon. Assume the universe consists of only three types of particle, the pi meson, the nucleon and the delta baryon. It is possible to understand some of the properties of the delta baryon by treating it as a pi meson-nucleon composite which is held together to a first approximation by the exchange of a nucleon. Chew then went on to show that the nucleon in turn can be treated as a pi meson-nucleon composite which is held together to a first approximation by the exchange of a delta baryon. This conceptual situation again leads to an equation of the form of 85 from which the properties of the delta baryon can be predicted through self-consistency. This has led to reasonably good predictions of the mass and lifetime of the delta baryon.

From this example we now begin to see a conceptual difficulty in the classical conception of a composite. According to classical ideas, if an object A is composed of other objects B, C, etc, then in a sense A 'contains' B, C, ... In the above example, however, if we use the containment model then we find that it is possible to have two different objects, the delta baryon and the nucleon, which contain exactly the same objects, a pi meson and a nucleon, in the same proportions. In terms of classical ideas it is hard to see how these two composite objects can be different. Worse still, we find that a nucleon is composed of a nucleon plus something else. In terms of the containment model this is incomprehensible. Yet another criticism can be found for the containment model of hadrons. It can sometimes happen that the following two reactions can both occur:

$$A \rightarrow B + C \qquad \ldots 84$$

and

$$B \rightarrow A + C \qquad \ldots 85$$

In terms of the containment model this would imply that A contains B and simultaneously B contains A. This makes no sense classically unless A and B are identical which they are not. The conceptual problem arises because the situation is inherently relativistic. The energies required in order to divide A or B into their component parts is of the order of magnitude of the energy required to create those parts because the binding energy is of the order of the rest energy of the virtual particle of the nuclear force. Therefore it becomes unnecessary to insist that the parts were already contained in the object prior to their release.

As a result of this partial quantitative success in regarding a number of hadrons as being composites of other hadrons, in 1962 G F Chew and S C Frautschi made the following generalisation:

Bootstrap Hypothesis: All hadrons are composites of other hadrons, none of them being any more elementary than any other.

This view of the nature of nuclear matter is inherently relativistic for the reason explained above. It is also inherently quantum mechanical since there may be a number of different possible constitutions of a given hadron, for example a positively charged sigma particle might be composed of a proton plus neutral pi meson or a neutron plus positively charged pi meson with different probabilities. All possible states of each hadron are taken into account by way of the SP. Thus, according to the bootstrap hypothesis, each hadron helps to generate other hadrons which in turn help to generate the first one. In this way the whole set of hadrons 'pulls itself up by its own bootstraps' (cf frontispiece by N C Escher). The world is then not seen as a construction built out of a number of unanalysable basic entities but rather as a self- generating network or relationships between entities which can play any of three roles:

(a) Each entity can play the role of a composite structure;

(b) It can be a constituent of another composite structure;

(c) It can be exchanged between constituents and thus constitute part of the force holding the structure together.

Thus, in contrast to the classical conception of ontological reductionism, it is argued that all hadrons are constructed out of other hadrons - there is no hadron aristocracy. The bootstrap hypothesis is therefore sometimes referred to as 'nuclear democracy'.

A further additional assumption made by Chew and Frautschi is that the S-matrix is uniquely defined by the above three constraints of unitarity, etc. This implies that S in equation 83 is unique. Given that S is unique and non-trivial, the solution, λ , of 83 is also unique. Thus, if the assumption of a unique S-matrix is correct then it follows that when the assumption is combined with the bootstrap hypothesis, not only are the properties of every hadron dependent on the properties of every other but also the properties of each hadron are uniquely determined by the properties of all the others. Consider, for example, the effect of changing the mass of the pi meson. Since the mass of the proton is partly determined by the creation of virtual pi mesons, this change would lead to a change in the mass of the proton. But the proton, in turn, affects the mass of the pi meson on account of the creation of virtual proton-antiproton pairs. Hence the mass of the pi meson would be changed further, and so on. Similar relationships hold among the other properties of the different particles. However, if the S-matrix is unique then this non-linear chain of causation is so restrictive that there is only one set of properties for each particle which would be consistent with the properties of all the others, that is the requirement of self-consistency of the properties of the whole set of hadrons determines uniquely the properties (eg the masses) of all the hadrons. This implies that there is only one possible ie self-consistent set of hadrons which is the one found in nature.

Let us now look at the justification for this conclusion. The examples given above of the rho meson and the delta baryon strongly suggest that it might be possible to derive the properties of all hadrons by considering them to be composites of other hadrons but this conjecture has never been proved although some progress in this direction has been achieved - see Dolan et al (1968). It seems, however, that in the case of the hadrons we are, at any rate, dealing with a set of particles whose properties are very deeply interconnected and which seem therefore to resist the simplification of classical reductionist principles. As regards the conjecture that general principles determine a unigue S-matrix, the situation is also still vague. It has never been shown that the relativistic invariance, unitarity and analyticity constraints on S determine S uniquely. Indeed, it has never even been shown that these three constraints are consistent with one another. There is here, however, the possibility of a disproof of the conclusion that the set of properties of the hadrons is self-determining. If it could be shown that the above three constraints allow more than one S-matrix to exist, then, given the correctness of the constraints and the existence of no others except those which can be derived from the original three, eg crossing, the hadrons found in nature could not be said to be the only possible consistent set. A similar conclusion would also follow if it could be proved mathematically that the existence of one S-matrix implies the existence of another also satisfying the above conditions. The uniqueness postulate cannot be directly tested experimentally if there only exists one S-matrix in nature. On the other hand, ability to derive certain other supplementary properties from self-consistency principles constitutes justification for the bootstrap hypothesis. Not all such properties have yet been derived in this way, for example, conservation of baryon number. If this property could be shown to follow in a self-consistent way from the general S-matrix then this would give impressive support to the bootstrap hypothesis. If, however, it could be shown that a principle could not be so derived this would imply either that the bootstrap hypothesis incorporating the uniqueness postulate is inadequate or that the principle does not in fact apply in nature.

The basis of the bootstrap idea is that physics should follow from the requirement of self-consistency among its components. Taken to its logical conclusion, a complete bootstrap hypothesis asserts that "nature is as it is because it is the only possible nature consistent with itself" (Chew 1963 p762). Thus in a complete bootstrap theory everything should follow from the requirement of self-consistency alone. Since it is hard to see how one could begin to construct a physical theory without the initial input of some unquestioned physical assumptions, every useful bootstrap theory can only be partial in the sense that it is only an approximation to a complete bootstrap theory, because it begins with certain unquestioned assumptions. In the case of the hadron bootstrap these assumptions are the three general constraints on the S-matrix together with the restriction that only hadrons are being considered. The bootstrap approach is then to gradually extend the generality of each partial theory by self-consistency arguments and thus reduce the number of initial assumptions. For example, it is hoped that leptons might eventually be included in the scheme.

Since every bootstrap theory is partial, covering a limited domain and making initial assumptions, the theory is inherently an approximation to a complete theory. One might then ask: why should such approximations work - even approximately? Since everything is interconnected, complete understanding of one aspect of nature requires complete understanding of all of nature. This question cannot be answered except to say that it has been a crucial discovery of scientific methodology that it is possible to ‘understand’ individual aspects of nature in an approximate way without having to understand everything all at once. This inevitably involves error, since everything is interconnected, but the error is often found to be small. If it were not possible to approximate then conventional science could not begin. Similarly, in the bootstrap approach, without the discovery that partial theories work tolerably well, this approach could not begin.

Let us conclude with some ontological implications of the attempts to combine relativity theory with QT.

1. According to classical physics (CP), matter is indestructible and cannot be created. This leads to the principles of conservation of (rest) mass and particle number. According to RQT, however, both of these principles must be abandoned in favour of the evidence that matter can be created and destroyed unless one admits the existence of an infinite sea of negative energy particles or that antimatter is matter moving backwards in time and adopt a block universe point of view.

2. According to CP, matter and vacuum are conceptually opposed. In RQT, however, due to the possibility of the creation of virtual particles out of nothing in vacuum, the distinction between matter and vacuum loses clarity.

3. In CP the elements or atoms from which material objects are constructed are assumed to be inert and passive. In RQT these particles are regarded as centres of continuous activity due to the surrounding cloud of virtual particles. Even elementary matter is not dead but very much alive.

4. Classically a particle is thought to transform and move continuously according to an equation of motion which governs its state at all times. According to the S-matrix approach, however, in general

only the initial and final states of the particle can be so described, the intermediate state being subject to the principle of cold pursuit and not a continuous equation of motion.

5. According to the bootstrap approach, matter need not be considered as being constructed out of 'irreducible basic building blocks' whose properties are arbitrarily ascribed by nature. The particles might have the properties that they do because this is the only way which they can be consistently ascribed. Further, no particle need be irreducible in principle - there is no need to postulate basic building blocks to account for physical matter.

Abbreviations

CFT: Classical Field Theory
CI: Copenhagen Interpretation
CM: Classical Mechanics
CP: Classical Physics
CPM: Classical Particle Mechanics
CT: Classical Theory
EPR: Einstein, Podolsky and Rosen
ETP: Einstein, Tolman and Podolsky
hv: Hidden Variable
hvs: Hidden Variables
hvt: Hidden Variables Theory
MM: Matrix Mechanics
QED: Quantum Electrodynamics
QM: Quantum Mechanics
QT: Quantum Theory
RQT: Relativistic Quantum Theory
SI: Statistical Interpretation
SP: Superposition Principle
UP: Uncertainty Principle
WM: Wave Mechanics

Glossary

Classification of Particle Types

Baryons: All strongly interacting, half-integral spin particles. This type of particle includes neutrons, protons and most of the recently discovered particles such as the lambda, sigma, delta and omega particles.

Hadrons: All strongly interacting particles. This includes all the baryons and mesons, but not the leptons or the photon,

Leptons: The known leptons are electrons, muons and neutrinos. These particles interact through the weak interactions but not the strong interactions.

Mesons: All strongly interacting, integral spin particles. This type of particle includes, for example, pi, rho, kappa and eta mesons.

Classification of Interaction Types

Electromagnetic Interactions: This interaction is responsible for atomic and molecular binding and hence for most forces of the everyday world. The particle associated with this force is the photon.

Gravitational Interactions: This interaction is responsible for the force between massive objects. The quantum particle associated with the force of gravity is believed to be the graviton, a spin 2 particle.

Strong Interactions: These are short range, charge independent interactions between nuclear particles. For example, the force between a neutron and a proton is the result of a strong interaction; in this case the particle associated with this force is the pi meson.

Weak Interactions: These interactions are $l0^{-13}$ times weaker than strong interactions and are responsible for β-decay radioactivity and particle decays taking longer than 10^{-15} sec. The associated particle is called the Intermediate Vector Boson.

Note: Parapsychology

In 1968 the first of a number of experiments of a similar and novel type was conducted by two French scientists (who felt it necessary to write their paper under pseudonyms), 'Duval' and 'Montredon' (1968). In their experiment a binary number generator was used to apply electric shocks randomly at suitable intervals to one or other of the two halves of a cage in which a mouse was confined. The movements of the animal were automatically monitored by a photocell device and the experiment was entirely self-operating thus eliminating human influence on the animal. It was found that the mice were able to avoid going to that half of the cage which was about to receive the shock, to an extent producing odds against pure chance of about 1000 to 1. This experiment was repeated by Levy (who, according to Costa De Beauregard, has now been exposed as a cheat) and McRay (1971) using a random number generator as described by Schmidt (1970a) whose functioning is based on a simple quantum mechanical process (the decay of radioactive strontium-90 nuclei). A number of similar experiments using a quantum mechanical source of randomness have now been reported using mice, gerbils, and humans. All of these experiments have given statistically significant results which seem inexplicable on the basis of pure chance alone - see eg Schmidt (1969).

These experiments can be conceptually simplified by a simple variation of the Schrödinger cat experiment described in Chapter Two. Assume both boxes A and B are attached to an electrocuting device and that there is a trap door between the two which allows the cat to pass between either box at will. If an electron happens to pass into box A through slit A of the screen and the cat happens to be in that box then it will be killed. If on the other hand it is in box B on this occasion, it will be saved. Assuming that the cat wishes to survive, this arrangement can in principle be used as a test for precognition on the part of the cat. If it is able to determine before the event ie precognise which slit the electron will pass through then it can arrange to be in the 'safe' box. If, after a large number of experiments on an ensemble of such arrangements, it is found that significantly more cats are alive than dead, one has grounds for saying that the cats could predict the outcome of events which are essentially quantum mechanical random events. While this is intended purely as a thought experiment which exemplifies the possibility in principle of detecting precognition in animals, it is horrifying to find that experiments have been conducted which attempt to detect precognition in rats wherein it was felt necessary to kill them randomly according to whether they ran across an odd or even number of floor squares (see Morris 1970). One would think there were more subtle ways to test for poison in a substance than to eat it!

The relevance of these experiments to a discussion of quantum mechanical principles should now be clear. According to the physicist, within its domain of application no conclusive refutation of the principles of QT has ever been given. The above experiments, however, can be interpreted as an unconventional test of the UP. According to the UP, it is impossible in principle to predict with more than 50% accuracy which slit, A or B, an electron will pass through in a double slit arrangement. Using basically this arrangement, Schmidt (1969) found in an experiment involving 20,000 trials with two 'gifted' subjects that they were able to predict the outcome of theoretically unpredictable sub-atomic processes with a probability against chance of 10^{10} to 1! If these results are accepted then it would seem that parapsychologists have found a serious, although unconventional, objection to the basic principles of QT One might object that the machine cannot have been truly random. However this argument is hard to support since not only do the principles of QT entail that the source of randomness (Sr^{90}) is random but also the randomness was computer tested in control experiments and found to have no significant bias.

Another series of experiments, very closely related to the above precognitive experiments, involve the concept of psychokinesis (PK) which is "used of alleged supernormal movements of objects, not due to any known force" (Myers 1907). In terms of the original Schrödinger cat experiment, one might imagine that a cat situated in box A might be able, by psychokinetic powers, to force the electron to pass through slit B rather than slit A. Thus, in an ensemble of cat experiments, if significantly more cats end up alive than dead then in this alternative interpretation one has grounds for saying that psychokinesis in cats had been demonstrated. Somewhat less gruesome experiments on this principle have recently been performed. One cold day in 1970 Schmidt placed a cat in an unheated garden shed. The only warmth in the shed came from a 200 watt lamp which was coupled to one of the outputs of a binary random number generator based on the random emission of electrons from a strontium-90 source. Whenever the generator produced a +1 pulse the lamp was turned on, and with a -1 pulse it was turned off. According to the UP, the machine should generate approximately equal numbers of +1 and -1 pulses. However, Schmidt (1970b) found that whenever the cat was in the shed (but not otherwise) the machine generated significantly more 'ons' than 'offs'. The overall effect gave odds of about 60 to 1 against chance, suggesting that the animal, in order to keep warm, was exerting a psychokinetic effect on the machine. Similar experiments involving a suitable control experiment to test randomness yet with much more significant results have now been performed on cockroaches, lizards, chickens, etc (see eg Watkins 1971; Levy (but see earlier note) and Andre 1970).

Schmidt (1970 p175) has also performed PK tests with human subjects. In this case the random number generator was connected to a display panel having 9 lamps in a circle. A +1 pulse from the generator caused the lamps to light up one at a time in a clockwise direction, while a -1 pulse caused them to light up in an anticlockwise direction. The idea of the experiment was then to try by PK to make the lamps light up in some predetermined direction. In his experiments Schmidt found, surprisingly enough, that the lamps tended to light up in a direction opposite to that willed by the subjects giving a significant negative deviation with odds against chance of 1000 to 1. When left unattended the machine showed no bias. In 1972 in the Cambridge Society for Psychical Research A D Cornell (President), B Carr and I decided to attempt to repeat Schmidt's experiment using basically the same experimental arrangement. A machine was designed by T Hooley which triggered randomly on high energy noise fluctuations of essentially quantum mechanical origin and built by Hooley and me. No consistent set of significant results has yet been obtained with this machine although a debate has arisen as to what it would mean for a machine to be 'random' and yet be ‘PK inf1uenced'. We note however that some other experimenters in this field have reported insignificant deviations from chance results using a Schmidt machine with schoolboys (see eg Randall 1974 p491).

We conclude that while parapsychology might prove to have interesting consequences for quantum physics, it is still too early to judge whether quantum principles are seriously endangered from this direction.

Bibliography

Aronov R A (1971) in Zeman ed. (1971 pp265-274).

Audi M (1973) The Interpretation of Quantum Mechanics, University of Chicago Press.

Bacon F (1620) Novum Organum.

Ballentine L E (1970) Rev Mod Phys 42, 358-381.
(1972) Am J Phys 40, 1763-1771.

Baracca A, Bohm D J, Hiley B J & Stuart A E G (1975) Nuovo Cimento 28B, 453-465.

Bastin T ed. (1971) Quantum Theory and Beyond, CUP.

Bedford D & Wang D (1975) Nuovo Cimento 26B, 313-325.

Bell J (1964) Physics 1, 195-200.
(1966) Rev Mod Phys 38, 447-452.
(1971) in d'Espagnat ed. (1971b pp171-181).

Belinfante F J (1973) A Survey of Hidden Variable Theories, Pergamon Press.
(1975) Measurement and Time Reversal in Objective Quantum Theory, Pergamon Press.

Bergstein T (1972) Quantum Physics and Ordinary Language, Macmillan.

Berkeley G (1710) A Treatise Concerning the Principles of Human Knowledge.

Birkhoff G & von Neumann J (1936) Annals of Maths 37, 823-843.

Bleuler E & Bradt H L (1948) Phys Rev 73, 1398.

Blokhintsev D I (1952) Critique de la Conception Idealiste de Theorie Quantique, in Questions Scientifiques (Editions de la Nouvelle Critique, Paris). “

Bohm D J (1951) Quantum Theory, Prentice-Hall.
(1952a) Phys Rev 85, 166-179.
(1952b) Phys Rev 85, 180-193.
(1957a) in Korner ed. (1957 pp33-40)
(1957b) Causality and Chance in Modern Physics, RKP.
(1962) in Quanta and Reality, Hutchinson.
(1971a) in Bastin ed. (1971 pp95-116).
(1971b) in d'Espagnat ed. (1971b pp412-469).

Bohm D J & Aharonov Y (1957) Phys Rev 108, p1070.
(1960) Nuovo Cimento 17, 961.

Bohm D J & Bub J (1966a) Rev Mod Phys 38, 453-469
(1966b) Rev Mod Phys 3; 470-475.

Bohr N (1925; 1927; 1929) in Bohr (1934)
(1939) Atomic Theory and the Description of Nature, CUP.
(1935) Phys Rev 48, 696-702.
(1937) Philosophy of Science 4, 289.
(1938; 1949; 1954; 1955; 1957) in Bohr (1958).
(1939) in New Theories in Physics, International Institute of Intellectual Cooperation, Paris.
(1948) Dialectica 2, 312.
(1958) Atomic Physics and Human Knowledge, Wiley.
(1960; 1961; 1962) in Bohr (1963).
(1963) Essays 1958-1962 on Atomic Physics and Human Knowledge, Interscience.

Born M (1926) Z Physik 38, 803.
(1943) Experiment and Theory in Physics, CUP.
(1955) Die Naturwissenschaften 11, 425-431.
(1971) The Born-Einstein Letters, Macmillan.

Born M & Jordan P (1930) Elementary Quantum Mechanics, Springer.

Bub J (1968) BJPS 19, 185-210.
(1973) in Hooker ed (1973 pp1-65).
(1974) The Interpretation of Quantum Mechanics, Reidel.

Bunge M (1956) Am J Phys 24, 272.
ed. (1967) Quantum Theory and Reality, Springer.

Campbell N R (1928) An Account of the Principles of Measurement and Calculation, Longmans & Green. I

Chew G F (1963) Science Progress 51, 529.
(1968) Science 161, 762-765.
(1970) Phys Today October pp23-28.
(1971a) Phys Rev D4, 2330-2335.
(1971b) in Bastin ed. (1971).

Chew G F, Gell-Mann M & Rosenfeld A (1964) Scientific American, February.

Clark R W (1973) Einstein, The Life and Times, Hodder & Stoughton.

Clauser J F (1969) Bulletin Am Phys Soc 14, 578.

Clauser J F, Horne M A, Shimony A & Holt R A (1969) Phys Rev Letts 23, 880.

Colodny R G ed (1962) Frontiers in Science & Philosophy, Pittsburgh Vol 1, University of Pittsburgh Press.
(1965) Beyond the Edge of Certainty, Pittsburgh Vol 2, University of Pittsburgh Press.
(1972) Paradigms & Paradoxes, Pittsburgh Vol 5, University of Pittsburgh Press.

De Broglie L (1926) C. R. Acad Sci Paris 183, 447.
(1927) J de Physique et du Radium 8, 225-241.
(1939) Matter & Light, New York.

D'.Espagnat B (1971a) Conceptual Foundations of Quantum Mechanics, Benjemin.
ed. (1971b) Proc Int School of Phys “Enrico Fermi”, Course IL, Academic Press.

De Witt B (1971) in d'Espagnat ed. (1971b pp21l-262).

Dirac P A M (1928) Proc Roy Soc (A) 117, 610; 118, 351.
(1950) The Principles of Quantum Mechanics, Clarendon Press.
(1965) Scientific American, May.

Dolen R, Horn D & Schmid C (1968) Phys Rev 166, 1768.

Duval P & Montredon E (1968) J Parapsychol 32, 153-166.

Dyson F (1949) Phys Rev 75, 486; 1736.

Einstein A (1936) J Franklin Inst 221, 349-382.
(1948) Dialectica 2, 320.
(1949) in Schilpp ed. (1949 pp665-688).

Einstein A & lnfeld L (1938) The Evolution of Physics, CUP.

Einstein A, Podolsky B & Rosen N (1935) Phys Rev 47, 777-780.

Einstein A, Tolman R & Podolsky B (1931) Phys Rev 37, 780-781.

Ellis G F R; (1971) Proc Int School of Phys "Enrico Fermi", Course XLVII, pp104-182, Academic Press.

Everett H (1957) Rev Mod Phys 29, 454.
(1975) in The Many-Worlds Interpretation of Quantum Mechanics, (De Witt B & Graham N eds.), Princeton University Press.

Faget J & Fert C (1957) Cahiers de Physique 11, 285-296.

Feinberg G (1972) in Colodny ed. (1972 pp33-46).

Feyerabend P K (1958) Proc Aris Soc Supp Vol 32, 75-104.
(1962) in Colodny ed. (1962 pp191-253).
(1968) Philosophy of Science 35, 309-331.
(1969) Philosophy of Science 36, 82-105.

Feynman R P (1949) Phys Rev 76, 749; 769.
(1966) Phys Today, August pp31-44.

Feynman R P & Hibbs A R (1965) Quantum Mechanics and Path Integrals, McGraw-Hill.

Fine A (1970) Phys Rev D2, 2783-2787.
(1972) in Colodny ed. (1972 pp3-31).

(1973a) BJPS 24, 1-37.
(1973b) in Suppes et al eds. (1973 pp567-592).

Fock V A (1951) Uspekhi Fisiceskich Nank 45, 3-14.

Frank P (1956) Erkenntnis 6, 303-317

Freedman S J & Clauser J F (1972) Phys Rev Letts 28, 938.

Furry W H (1936) Phys Rev 49; 393; 476.

Gamow G (1958) Scientific American, January.
(1966) Thirty Years that Shook Physics, New York.

Gardner M R (1972a) BJPS 23, 13-23
(1972b) BJPS 23, 89-109.

Gleason A M (1957) J Maths & Mechanics 6, 885-893.

Good I J ed. (1961) The Scientist Speculates, Heinemann.

Groenwold H J (1974) Physics Reports (C) 11, 327-354.

Gurney E, Myers F W H & Podmore F (1886) Phantasms of the Living, London.

Hanson N R (1962) in Quanta and Reality, Hutchinson pp85-93.
(1963) The Concept of the Positron, CUP.

Heisenberg W (1927) Z Physik 43, 172.
(1930) The Physical Principles of the Quantum Theory, University of Chicago Press.
(1934) Z Physik 90, 209; 92, 692.
(1938) Ann der Physik (Leipzig) 32, 20.
(1943) Z Physik 120, 513; 673.
(1955) in Pauli ed. (1955).
(1958) Physics and Philosophy, Harper & Row
(1959) 2nd ed. of (1958).
(1961) in On Modern Physics, Polter C N, New York pp9-10.
(1967) in Rozenthal ed. (1967).
(1971) Physics and Beyond, Harper & Row.
(1976) Phys Today, March pp32-39.

Hesse M B (1965) Forces and Fields, Littlefield-Adams.
(1966) Models and Analogies in Science, University of Notre Dame Press.
(1974) Camb Rev, February 95, 62-65

Hooker C A (1970) Am J Phys 38, 851-857
(1972) in Colodny ed. (1972 pp67-302).
ed. (1973) Contemporary Research in the Foundations and Philosophy of Quantum Theory, Reidel.
ed. (1975) The Logico-Algebraic Approach to Quantum Theory, Reidel.

Hume D (1759) Treatise of Human Nature.

James W (1890) The Principles of Psychology, Holt.

Jammer M (1966) The Conceptual Development of Quantum Mechanics, McGraw-Hill
(1974) The Philosophy of Quantum Mechanics, Wiley.

Jauch J M (1968) The Foundations of Quantum Mechanics, Addison-Wesley.

Jönsson C (1961) Z Physik 161, 454-474.

Jung C (1948) Preface to White S E, Das Uneingeschrankte Weltall, Zurich.
(1960) The Structure and Dynamics of the Psyche, Collected Works Vol VIII London.

Kammerer P (1919) Das Gesetz der Serie, Deutsche Verlags-Anstalt, Berlin

Kant I (1787) Critique of Pure Reason.

Kasday L (1971) in d'Espagnat ed. (1971b pp195-210).

Kasday L, Ullman J & Wu C S (1970) Bulletin Am Phys Soc 15, 586.

Kochen S & Specker E P (1967) J Maths & Mechanics 17, 59-87.

Komar A (1973) in Hooker ed. (1973 pp305-327).

Korner S ed. (1957) Observation and Interpretation, Butterworths.

Kramers H A (1937) Proc Kgl Ned Akad Wet 40, 814.

Landé A (1965) New Foundations of Quantum Mechanics, CUP.

Langhoff H (1960) Z Physik 160, 186.

Levy W J & Andre E (1970) Abstract J Parapsychol 34, 303.

Levy W J, McRay A et al (1971) J Parapsychol 35, 1-17; 120-131.

London F &: Bauer E (1939) La Theorie de l'Observation en Mecanique Quantique Hermann & Cie.

MacKay D M (1957) Mind 66, 390-394.
(1958) Proc Aris Soc Supp Vol 32, 105-122.

Madelung E (1926) Z Physik 40, 322.

Margenau H (1931) Monist 41, 1-36.
(1932) Monist 42, 161-188.
(1937) Philosophy of Science 4, 337.
(1950) The Nature of Physical Reality, McGraw-Hill
(1954) Phys Today 7, 6-13.

Matthews P T (1968) Introduction to Quantum Mechanics, McGraw-Hill.

Meier C A (1935) in Der Kulturelle Bedeutung der Komplexen Psychologie, Berlin.

Messiah A (1961) Quantum Mechanics, Amsterdam.

Mill J S (1879) A System of Logic, London.

Morris R L (1970) J Am Soc Psych Res 64, 242-260.

Myers F W H (1907) Human Personality and its Survival of Bodily Death,
Longmans-Green.

Nagel E (1961) The Structure of Science, Hardcourt.

Oppenheimer J R (1930) Phys Rev 35, 959.

Papaliolios C (1967) Phys Rev Letts 18, 622-625.

Park J L (1966) Philosophy of Science 35, 205-231.

Pauli W (1933) Handbuch der Physik, Springer pp83-272.
ed. (1955) Niels Bohr and the Development of Physics, Macmillan.
(1961) Aufsatze and Vortrage uber Physik und Erkenntnis-Theorie, Verlag
Vieweg Braunschweig.

Pauli W & Jung C (1952) Natureklarung und Psyche, Zurich.

Peat F D (1973) in Hooker ed. (1973 pp328-345).

Petersen A (1963) Bulletin of the Atomic Scientist 19, 8-14.
(1968) Quantum Theory and the Philosophical Tradition, MIT Press.

Pfleegor R & Mandel L (1967) Phys Rev 159, 1084-1088.

Popper K R (1935) Logik der Forschung, Springer.
(1959) English translation of (1935).
(1967) in Bunge ed. (1967 pp7-44).

Przibram K (1926) Paul Kammerer als Biolge, Monistische Monatshefte, November.
(1967) Letters on Wave Mechanics, Vision Press

Putnam H (1965) in Colodny ed. (1965 pp75-101).
(1967) in Bunge ed. (1967 p7).
(1969) Boston Studies in Phil of Sci 5, 216-241.

Randall J L (1974) J Soc Psych Res 47, 485-494.

Reichenbach H (1944) Philosophic Foundations of Quantum Mechanics, University
of California Press.
(1948) Dialectica 2, 337-350.

Robertson H P (1929) Phys Rev 34, 163-164.

Rosen N (1945) J Elisha Mitchell Sci Soc 61, 67-73.

Rosenfeld L (1957) in Korner ed. (1957 pp41-45).
(1963) Z Physik 171, 242-245.
(1967) in Rozenthal ed. (1967 pp114-137).

Rozenthal S ed. (1967) Niels Bohr, His Life and Work as seen by His Friends and Colleagues, New York.

Scheibe E (1973) The Logical Analysis of Quantum Mechanics, Pergamon Press

Schilpp P A ed. (1949) Albert Einstein: Philosopher-Scientist, Evanston.

Schmidt H (1969) J Parapsychol 33, 99.
(1970a) J Parapsychol 34, 219-224.
(1970b) J Parapsychol 34, 255-261.

Schrödinger E (1935a) Die Naturwissenschaften 23, 807-812; 824-828; 844-849.
(1935b) Proc Camb Phil Soc 31, 555-563.
(1936) Proc Camb Phil Soc 32, 446-452.

Schwinger J (1966) Phys Today, June pp27-34.

Sharp D H (1961) Philosophy of Science 28, 225-233.

Shimony A (1963) Am J Phys 31, 755-773.
(1971) in d'Espagnat ed. (1971b pp182-194).
(1973) in Suppes et al eds. (1973 pp593-601).

Siege1 A (1966) in Wiener et a1 (1966 Ch5).

Sinclair U (1930) Mental Radio, Springfield.

Stapp H P (1971) Phys Rev D3, 1303-1321.

Suppes P et al eds. (1973) Logic Methodo1ogy& Philosophy of Science IV.

Szilard L (1929) Z Physik 53, 840-856.

Tutsch J H (1971) J Math Phys 12, 1711.

Von Neumann J (1932) Mathematische Grundlagen der Quantenmechanic, Springer.
(1955) English translation of (1932).

Walker E H (1970) Mathematical Biosciences 7, 131.

Wassermann G D (1956) CIBA Foundation Symposium on ESP, London p53.

Watkins G (1971) Abstract J Parapsychol 35, 62.

Weingarten D (1974) Nuovo Cimento 22B, 293-327.

Weyl H (1928) Gruppentheorie und Quantunmechanik, Leipzig.
(1931) 2nd edition of Weyl (1928).

Wheeler J A (1946) Ann NY Acad Sci 48, 219.

Wiener N el al (1966) Differential Space, Quantum Systems & Prediction
(B Rankin ed.), MIT Press.

Wiener N & Siegel A (1953) Phys Rev 91, 1551.
(1955) Nuovo Cimento Supp 2 (4), 982.

Wigner E (1961) in Good ed. (1961).
(1970) Am J Phys 38, 1005-1009.
(1971) in d'Espagnat ed. (1971b pp1-19).
(1973) in Hooker ed. (1973 pp369-385).

Witmer E (1967) J Phys 35, 40-52.

Wittgenstein L (1921) Tractatus Logico-Philosophicus, RKP.

Wu C S & Shaknov J (1950) Phys Rev 77,136.

Yukawa H (1935) Proc Roy Maths Soc Japan 17, 48.

Zeman J ed. (1971) Time in Science & Philosophy, Elsevier.