

~~██████████~~ in Synthese 1979

The Mathematical Foundations of Quantum Mechanics

by

David A. Edwards  
Department of Mathematics  
University of Georgia

## Contents

### Preface

### §1. A Formal Schema For Science.

- 1.1. The Empirical Situation.
- 1.2. The Theoretical Situation.
- 1.3. Category Theory.
- 1.4. Symmetry Principles.

### §2. Classical Theories

- 2.1. Statics.
- 2.2. Dynamics.
- 2.3. Classical Statistical Mechanics.
- 2.4. Field Theory.

### §3. Quantum Theories

- 3.1. The Logic of Quantum Mechanics.
- 3.2. Dynamics.
- 3.3. The Quantum Theory of Particles.
- 3.4. Examples.
- 3.5. The Quantum Theory of Interacting Particles.
- 3.6. Quantum Statistical Mechanics.
- 3.7. Quantum Field Theory.
- 3.8. Final Remarks.

## Preface.

Classical mechanics was first envisaged by Newton, formed into a powerful tool by Euler, and brought to perfection by Lagrange and Laplace. It has served as the paradigm of science ever since. Even the great revolutions of 19th century physics--namely, the Faraday-Maxwell electro-magnetic theory and the kinetic theory--were viewed as further support for the complete adequacy of the mechanistic world view. The physicist at the end of the 19th century had a coherent conceptual ~~scheme~~ which, in principle at least, answered all his questions about the world. The only work left to be done was the computing of the next decimal.

This consensus began to unravel at the beginning of the 20th century. The work of Planck, Einstein, and Bohr simply could not be made to fit. The series of ad hoc moves by Bohr, Eherenfest, et al., now called the old quantum theory, was viewed by all as, at best, a stopgap. In the period 1925-27 a new synthesis was formed by Heisenberg, Schrodinger, Dirac and others. This new synthesis was so successful that even today, fifty years later, physicists still teach quantum mechanics as it was formulated by these men.

Nevertheless, two foundational tasks remained: that of providing a rigorous mathematical formulation of the theory, and that of providing a systematic comparison with classical mechanics so that the full ramifications of the quantum revolution could be clearly revealed. These tasks are, of course, related, and a possible fringe benefit of the second task might be the pointing of the way "beyond quantum

theory". These tasks were taken up by von Neumann as a consequence of a seminar on the foundations of quantum mechanics conducted by Hilbert in the fall of 1926. In papers published in 1927 and in his book, The Mathematical Foundations of Quantum Mechanics, von Neumann provided the first completely rigorous treatment of quantum mechanics. von Neumann first realized that the essential mathematical reason for the equivalence of the matrix and wave mechanical formulations of quantum mechanics was the isomorphism between the sequence space  $\ell^2$  and the function space  $L^2$ . He was thus led to formulate the notion of an abstract Hilbert space and to develop its mathematical theory. The main motivation for the axioms for quantum mechanics chosen by von Neumann in his book was that they yielded a rigorous theory, notationally similar to and computationally as powerful as that developed by the physicists. In his book he also alluded to a deeper, more profound approach. He states on page 253:

As can be seen, the relation between the properties of a physical system on the one hand, and the projections on the other, makes possible a sort of logical calculus with these. However, in contrast to the concepts of ordinary logic, this system is extended by the concept of "simultaneous decidability" which is characteristic for quantum mechanics.

von Neumann and G. Birkhoff developed this approach in 1936 in a paper entitled "The Logic of Quantum Mechanics". They introduced the notion of the logic  $\mathcal{L}$  of a physical theory. For classical theories  $\mathcal{L}$  is a Boolean algebra; but for quantum theories  $\mathcal{L}$  is assumed to be isomorphic to the logic  $\text{Proj}(H)$  of projections in a complex separable Hilbert space. In fact, it is now possible to derive von Neumann's original formulation of quantum mechanics from the single axiom that  $\mathcal{L} \cong \text{Proj}(H)$ .

This program will be discussed in §3. Other assumptions on  $\mathcal{L}$  are known to be appropriate in quantum statistical mechanics and in quantum field theory and these will also be discussed in §3.

A natural question is: "What is the appropriate general abstract structure for our logics, i.e. in what category do they live?". When von Neumann and Birkhoff first considered this problem in the mid 1930's, the natural thing to do - given the mathematics of the time - was to assume that  $\mathcal{L}$  was a set on which the classical operations were defined, but that they satisfied non-classical identities. This approach leads one to assume that  $\mathcal{L}$  is an ortho-complemented lattice, or an ortho-complemented partially ordered set, or a partial Boolean algebra, or some other related structure. This choice of syntax leads to certain semantical viewpoints which have been explored by Finkelstein, Putnam, Kochen and Specker, and others. Mathematically, our approach, which will be based upon the concept of a diagram of Boolean algebras, may be viewed as the natural change in perspective brought about by the move from Birkhoff and MacLane's A Survey of Modern Algebra, to MacLane and Birkhoff's Algebra. The natural semantics associated to our new syntax seems to be identical to Bohr's philosophy of complementarity, i.e. the world must be viewed through classical (Boolean) frames which interlock in a non-trivial fashion. In §1 we will try to motivate this choice of structure. In §1.3 we will provide a summary of the category theory required for this work. §1 will also discuss a general formal schema for science. Such a schema is necessary in order to understand the formalism and how it is to be used. In §1 we will also discuss various questions relating to the nature of probability; the problem of

reductionism and the possibility of hidden variables theories; and the possibility of "realistic" interpretations of the formalism. §2 will discuss classical theories and §3 quantum theories.

This work is not meant to be a first introduction to the foundations of quantum mechanics (for that see [38]), nor a systematic treatise including the detailed development of the requisite mathematics (for that see [81]), but, instead, a systematic outline of a coherent approach to quantum mechanics with all the assumptions carefully scrutinized for their physical meaning.

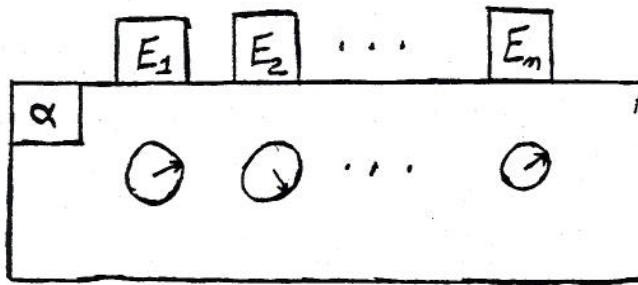
### §1. A Formal Schema for Science.

In this section we shall develop a formal schema which is meant to capture certain aspects of the scientific process. The relationship between our schema and the actual scientific process should be thought of as analogous to the relationship between the notion of a formal mathematical system and the actual mathematical process. Namely, we are neglecting human passions, social institutions and many other complexities of the "real" world and focusing only on certain central and formalizable aspects of the process. Our goal is to provide a single schema which will encompass both classical and quantum theories as special cases. This will provide a clear comparison of the similarities and differences between these two theories, while at the same time pointing the way "beyond" quantum mechanics.

### 1.1. The Empirical Situation.

We will assume that each experiment that one is interested in performing has a finite formal description  $E$ , and that performing  $E$  yields a result in a finite set  $R_E$ . The experiments can be thought of as being on a fixed "system" which may be in various internal states  $\alpha \in \mathcal{A}$ , depending upon how we prepared the system. The breakup of the description of an experiment into the pair  $(\alpha, E)$ , consisting of an initial "preparation" followed by an actual "observation", though somewhat conventional, is very important for the development of our formalism. On the other hand, the notion of the "system" itself is often quite opaque and misleading. It plays a role similar to Kant's ding and sich. There is little trouble when one is talking about billiard balls, tables, and the like, where one can simply "point" to the system itself which one is trying to study. But how does one "point" to an electron? Our experiments should really be thought of as "modes of percieving", "ways of looking", and the results of our experiments should be thought of as "what was seen". Careless use of the notion of the "system itself" inevitably leads to the notorious measurement problem which itself is reminiscent of the classical problem of the relationship between phenomena and numina. In our approach one deals strictly with phenomena and never even supposes the existence of numina. Of course, by invoking Berkeley's universal observer one can connect our approach with that of the traditional realist position.

A picturesque way of thinking about our view of the empirical situation is as follows. Assume we are given a large collection of black boxes such that each box has a label  $\alpha \in \text{Lab} = \mathcal{A}$  and various buttons  $E_i \in \text{Exp}$ . Below each button is a dial which may read a value in a finite set  $R_{E_i}$ . When a button is pressed, a result is recorded on the dial and then the box blows up. Only one button can be pressed.



We will treat all boxes with the same label as identical. Hence, in order to be able to "repeat" an experiment, it is necessary to have many boxes with the same label. Given the label  $\alpha$ , we would like to be able to predict the result obtained when the experiment  $E$  is performed. It is possible that knowledge of  $\alpha$  is insufficient to predict the result of performing  $E$ . For example, if our experiment consists of tossing a coin, we usually do not pack enough information into our description of its state to determine a unique outcome. But one is certainly free (in this case, at least) to believe that the reason for this lack of determinism is purely practical, and not fundamental, and that a fuller description of the states of the system is possible (at least in principle), which would yield a deterministic system. This belief leads to a research strategy of attempting to find a more complete, more basic, more fundamental description of the

system. Whether such a reductionistic program is always possible and/or useful is a perennial question which we will face many times in this work. In any case, we will always assume that the state  $\alpha$  at least determines a unique probability distribution  $p(\alpha)$  on each of the sets  $R_E$ . This is a highly non-trivial assumption. It implies that one has done a good job in describing the aspects of the system under consideration which affect the results of the experiments to be performed. For example, if we have two types of weighted coins, one type of which always comes up heads and the other always comes up tails, and if we cannot distinguish the types of coins, then it is possible that in tossing a large collection of such coins no stable frequency distributions will be discovered.

To each set  $R_E$  we associate its power set  $\mathcal{P}(R_E)$  (= the set of all subsets of  $R_E$ ) and think of  $\mathcal{P}(R_E)$  as the Boolean algebra of questions associated to the experiment  $E$ . That is, each  $Q \subset R_E$  is to be thought of as the question: When  $E$  was performed, was the result in  $Q$ ? The diagram (see §1.3) of sets  $C^S = \{R_E\}$  will be called the surface configuration space of our system. The associated diagram of Boolean algebras  $\mathcal{L}^S = \{\mathcal{P}(R_E)\}$  will be called the surface logic of our system. Our basic assumption so far is the existence of a map  $p: \text{Lab} \rightarrow \text{Prob}(\mathcal{L}^S)$ , from the set of labels to the set of probability measures on  $\mathcal{L}^S$  (see §1.3). The set Lab can always be assumed to have a natural convex structure as follows. If  $\alpha_1, \dots, \alpha_n$  are labels and  $p_1, \dots, p_n$  are weights such that  $p_i \geq 0$  and  $\sum p_i = 1$ , then choose a large natural number  $N$  and form a new ensemble of  $N$  boxes with label  $\alpha$  by choosing for each  $i$ ,  $p_i N$

boxes with label  $\alpha_i$  and stamping the new label  $\alpha$  over the old label  $\alpha_i$ . Knowledge of the label  $\alpha$  is certainly not maximal knowledge of our system.  $\alpha$  is called a classical mixture of the states  $\alpha_i$  with weights  $p_i$ . The map  $p: \text{Lab} \rightarrow \text{Prob}(\mathcal{L}^S)$  is easily seen to be a convexity-preserving morphism from the above described convex structure on Lab to the natural convex structure on  $\text{Prob}(\mathcal{L}^S)$ . But, while a label which is a classical mixture certainly represents ignorance, there is no a priori reason to attribute any ignorance to a label  $\alpha$  whose associated probability measure  $p_\alpha$  on  $\mathcal{L}^S$  is not pure! The deterministic research program, if successful, will redescribe all our present labels which do not go to pure probability measures as classical mixtures of a new set of pure labels all of which do determine pure probability measures. But, until such a program is successfully carried out, we will assume that the process may be "intrinsically stochastic". This seems to be the most reasonable and commonsensical attitude towards a coin toss, for example.

How does one determine  $p$ ? One way, of course, is to do lots of experiments. This direct approach can only yield an estimate of  $p$ , may be very expensive, and, in any case, is usually not very enlightening. What one really needs at this point is a good theory. The theoretical situation will be discussed in the next section.

There are some people who just can't resist trying to look inside a black box. They will insist on trying to open the box with a screw driver, or, at least, look into the box with an x-ray machine. From our point of view, what they are doing is, in effect, putting

the original black box into a larger black box which contains more buttons, i.e. they are insisting on using a larger repertoire of modes of perceiving. Sometimes such a viewpoint is useful, sometimes not. Why not simply allow all possible modes of perceiving? The reason is simple. It has to do with the chances of developing a good theory. The world as a whole is a mess! Only by narrowing our sights and focusing in on a small part of the universe can we have any reasonable hope of attaining a deep understanding. This has been the method of science, and it has proved very successful in practice. Thus, we have Skinner's program for understanding animal behavior and developing an art of behavioral engineering without invoking scalpels, electrodes or modern chemistry. Similarly, in physics it is (usually) inappropriate to look into the experimenter's sex life. Both the use of a scalpel and a view of the experimenter's sex life may be very revealing. But they are both banned from the allowed techniques of the above sciences. Other sciences do allow the scalpel and/or a view of the experimenter's sex life, namely, neurophysiology and Freudian psychology. Thus, the basic method is, first, to focus in on a few basic modes of perceiving and to develop a good theory and deep insight into this restricted domain of discourse; and then later to attempt to expand one's repertoire and broaden the domain of discourse.

One might also wonder how it is possible to develop good theories of what we see before we have reasonable theories of how we see. In other words, how can we use instruments without knowing how they work? Of course, the fact is that we do so all the time! Man has seen the

world with his eyes and formed many successful theories about the world long before he understood how his eyes worked. (For that matter, do we even today understand how the eyes work or what light is?) In fact, one needs a good theory about what was supposed to have been seen before one can say that an instrument malfunctioned. A good example of this is the fact that only fairly recently has the existence and prevalence of color blindness been discovered. Only by having an external standard of what he should have seen can a person who is color blind actually realize it. Another good example is that Newtonian celestial mechanics was successfully developed using telescopes (or at least the eye) as the basic observing instrument before it was known that light had a finite velocity of propagation; it was, in fact, a disagreement between what Newtonian mechanics said should be seen and what was actually seen that led to the first reasonably accurate estimate of the velocity of light. Of course, pointing a familiar instrument in a new direction will usually yield new perspectives which are rapidly assimilatable. Thus, once we feel that we understand how a microscope works, very rapid progress can be made by simply looking at everything under the microscope. This has been the case for light and electron microscopes and probably will be the case for quark microscopes, if they are ever made. Having a theory of how an instrument works clearly involves looking at the instrument with other - presumably more familiar - instruments.

Now, it may happen that we discover, through a lot of experimenting, that the map  $p: \text{Lab} \rightarrow \text{Prob}(\mathcal{L}^S)$  is not surjective. We may believe this to

be due to our lack of technical knowhow in providing for a sufficiently diverse range of preparations, i.e. if we were cleverer, we could enlarge the set Lab and thus obtain a surjective map. This attitude leads to a research strategy of attempting to enlarge Lab. If this strategy fails in significantly enlarging the image of  $p$ , then we might change our attitude towards this failure of surjectivity from that of being our fault to that of being "nature's" fault, i.e. we may decide to attribute the lack of surjectivity of  $p$  to a "law of nature". Thus, that objects when released fall and that the light goes on when the switch is flicked, would not simply be viewed as summaries of high correlation coefficients, but as something more significant. Thus, when it was discovered that parity was not always conserved, this was not simply viewed as a minor correction in estimated correlation coefficients, but as a major change in perspective and of how we expect the world to work. Physicists were surprised by Nature's violation of parity. After the discovered violation of parity, physicists then also expected and later discovered violations of C and T invariance. But most of them do not expect violation of CPT invariance. This is because CPT invariance is a theorem of a very general axiomatization of quantum field theory and its violation would require a drastic revision of the physicists' conception of nature. Only time will tell whether such a revision will, in fact, be necessary. But, in any case, a position has been taken which is certainly not simply a summary of past correlation coefficients. What we have been implicitly criticizing is Hume's attitude towards causality. With Hume's attitude one would never be surprised by any

occurrence. If the sun didn't rise tomorrow, one would simply and calmly revise one's estimate of the probability of the sun's rising on the next day. Such an attitude certainly does not describe the way people often behave. Positions have to be taken and theories have to be developed even though they open us up to the possibility of being surprised!

If the map  $p: \text{Lab} \rightarrow \text{Prob}(\mathcal{L}^S)$  is not injective one often describes the situation by saying that many superficially distinct ways of preparing the system actually put it into the same internal state. We will have more to say about this in the next section.

Summarizing, we have in this section introduced a formalization of the experimental process which includes: a set **Lab** of recipes for preparing the "system" under consideration; a surface logic  $\mathcal{L}^S$  of experimental questions; and a priori map  $p: \text{Lab} \rightarrow \text{Prob}(\mathcal{L}^S)$ . So far, the only way discussed of determining  $p$  is to do lots of experiments. This formalization is not too far removed from the description of his work that an experimentalist would give in a publication. Of course, the experimentalist would usually consider the most creative part of his work to be the process whereby he decides what to measure, how to measure it, and how to display the results. The actual carrying out of the experiment and the ensuing paperwork he may very well leave to his lab assistants. We have no idea how to formalize this very complicated decision process. (such a task might be as difficult as writing the program for HAL 9000). Luckily, it won't be necessary for our program of understanding the similarities and differences between classical and quantum mechanics. These differences reside, from our

point of view, entirely in the structure of the deep logic,  $\mathcal{L}^d$ , to be introduced shortly.

It has been suggested that our formalism leads to a picture of the world as a huge room filled with buttons and dials. This is certainly a bleak and colorless world view! But, by simply deciding to have the output of our instruments presented in the form of color photographs or put on cathode ray screens with accompanying sound tracks, we obtain a much more vivid and colorful world view.

### 1.2. The Theoretical Situation.

In this section we will introduce the notion of the deep configuration space,  $\mathcal{C}^d$ , and the deep logic,  $\mathcal{L}^d$ , of a system. (While we have already warned the reader in §1.1 of the dangers of the use of the notion of the "system", we find it linguistically necessary to keep invoking it in order to convey our ideas without inordinate circumlocutions.) Consider the problem of measuring the length of a rod. There are many different ways to measure the length of a rod, and, on the empirical level, they are all *a priori* independent, and thus determine disjoint sub-algebras of  $\mathcal{L}^S$ . In fact, on the empirical level, we cannot really talk about the length of a rod, but only about the value obtained by some length-measuring operation. The notion of the length of a rod is a theoretical concept, where all the various ways of measuring the length are conjoined into a single concept. In mathematical terms, we have a homomorphism  $\varphi: \mathcal{L}^S \rightarrow \mathcal{L}^d$  from our surface logic  $\mathcal{L}^S$  to a deep logic  $\mathcal{L}^d$  which describes the "true" nature of the system. The actual situation is even more complicated, since the various instruments are usually viewed as not to be actually recording the value of the true length but instead to be outputting a message which, besides containing the true signal, also contains noise. Thus, the empirical instruments are viewed as "noisy observables" from whose output one can only probabilistically infer the true value, i.e. one must attempt to filter out the noise in order to obtain the true signal. The reader may detect an echo of Plato's view that the senses reveal only a "corrupted copy" of the "true". Note that once we have introduced our deep logic and its associated

notion of what our instruments were supposed to have recorded, we can begin to talk about how well our instruments are working and how noisy they are. If  $\phi$  is not 1-1, then our theory tells us that certain observables, which on the empirical level are a priori different, are, in "reality", measuring the same thing. In terms of our box model, two different buttons (or, at least, certain parts of their outputs) are equivalent. One way this might happen is if we had a box in which more than one button could be pressed at the same time. Pressing  $E_1$  or pressing both  $E_1$  and  $E_2$  but recording only the dial reading of  $E_1$  are a priori completely different experiments, and, in fact, it is easy to design boxes where they are totally unrelated. But we may have reason to believe that the two experiments are really measuring the same thing (for example, if the box comes in two halves with no visible connections). Thus, choosing a  $\phi$  which is not 1-1 is a way of identifying certain empirical observables.

If  $\phi$  is not onto, and we really believe in the correctness of  $\mathcal{L}^d$ , then we are led to re-examine the empirical situation in order to discover means of experimentally measuring the qualities of the system which our theory,  $\mathcal{L}^d$ , says it has. If successful, this would yield a richer surface  $\log \mathcal{Z}^S$  and an extended homomorphism  $\bar{\phi}: \mathcal{Z}^S \rightarrow \mathcal{L}_d$  such that  $\bar{\phi}$  is onto. In fact, it is even reasonable to look for a  $\bar{\phi}$  which is surjective on Boolean subalgebras, i.e. if  $A_1, \dots \in \mathcal{L}^d$  are compatible, that is, are contained in a Boolean subalgebra of  $\mathcal{L}^d$ , then there should exist experimentally compatible  $\bar{A}_1, \dots \in \mathcal{Z}^S$  such that  $\bar{\phi}(\bar{A}_i) = A_i$ . In more informal terms, if  $\{A_i\}$  are theoretically compatible, then they are simultaneously observable. For example, suppose our experiments

consist of taking single photographs of a cup from various possible angles. Our experiments yield only the two dimensional projections of the actual cup in three space, i.e. we really believe in a deep logic consisting of the Borel algebra of the configuration space of closed subsets of  $R^3$ , while our surface logic is a disjoint collection of logics each of which is isomorphic to the Borel algebra of the configuration space of closed subsets of  $R^2$ . In this case, we can use holograms to extend the surface logic to one that will be surjective. On the other hand, if we are attempting to model the sun and wish to include in our model the "atoms at the center of the sun", then these "atoms" will appear in our deep logic, but certainly won't (at least at present) appear anywhere in our surface logic. An even more extreme example occurs when one wishes to discuss atoms in a star  $10^{10}$  light years away. According to present theory (special relativity), we will never be able to lift these theoretical observables up from the "twilight zone" of the deep logic to the "real world" of our surface logic! We thus realize that there will be many situations where we will be interested in a deep logic which is not entirely covered by a surface logic. One thus talks about theoretical or "potential" observables in contrast to the actual empirical (noisy) observables. Kant, also, had found it necessary to introduce the notion of "potential phenomena". We are thus led to a form of platonic dualism consisting of an ideal world of theoretical observables in contrast to an empirical world of corrupted observables. The map  $\varphi: \mathcal{L}^S \rightarrow \mathcal{L}^D$  induces a map on probability measures  $\varphi^*: \text{Prob}(\mathcal{L}^D) \rightarrow \text{Prob}(\mathcal{L}^S)$  and, in order to be able to use  $\mathcal{L}^D$  for making predictions, it is necessary to assume that we are

also given a map  $p_t : \text{Lab} \rightarrow \text{Prob}(\mathcal{L}^d)$  such that the diagram

$$(1.2.1) \quad \begin{array}{ccc} \text{Lab} & \xrightarrow{p} & \text{Prob}(\mathcal{L}^S) \\ & \searrow p_t & \uparrow \varphi^* \\ & & \text{Prob}(\mathcal{L}^d) \end{array}$$

commutes. Once we have committed ourselves to a particular choice of  $\mathcal{L}^d$ , no amount of experimental observations can force us to relinquish it! One can always invoke more "noise" in our instruments to explain away a seeming lack of commutativity in diagram (1.2.1). If the noise level gets very high, then the intensity of our commitment to this particular logic tends to decrease. We may then refer to it as a "model", which may be "good", possibly only "approximate", maybe only "heuristic". We may even go so far as to suspect that a different choice of  $\mathcal{L}^d$  would be "better". To talk about a logic  $\mathcal{L}$  as a model for some physical phenomena means that when we investigate the properties of  $\mathcal{L}$  we find them to behave in an analogous fashion to certain physical phenomena. Recall the physicists use of such models as the harmonic oscillator, the Ising model, and the  $P(\varphi)_2$  quantum field theory model. The first two models are usually called "approximate"; the third model isn't actually thought to correspond to any physically real system since we currently believe space-time to be four- and not two-dimensional. Yet all three models are thought to provide insight into the true workings of the world. What exactly is the relationship between the theoretician's computation of the spectrum of a certain Hermitian operator and the experimentalist's burning of hydrogen and measuring

the spectrum of the emitted light? Certainly, one usually describes the situation by invoking "informal" correspondence rules which, if formalized, would yield diagram (1.2.1). But we suspect that most--if not all--serious uses of theories involve the tying together of theory and experiment through the use of a vague appeal to analogy: Behold! the World and the Model behave similarly. We would like to know the "true" deep logic  $\mathcal{L}^d$ , but will settle for "model" logics which yield insight. In any case, we feel that the diagram (1.2.1) is crucial for understanding the formalism of quantum mechanics.

The factoring of  $p$  through  $\text{Prob}(\mathcal{L}^d)$  may allow us to "explain" why  $p$  is not surjective; namely, it isn't because  $\phi^*$  isn't. On the other hand, we would still want to understand why  $p_t$  isn't surjective. For example, a relativistic model would not contain a probability measure corresponding to a preparation of a billiard ball with a velocity  $10^{10}$  miles per second. Hence, any failure on our part to prepare such a system is looked upon as a success of the theory and not as a failure of our ingenuity. On the other hand, the model may contain a probability measure corresponding to a preparation of the billiard ball at a location  $10^{10}$  light years away. We don't know how to prepare such a system. But, by viewing ourselves as part of a larger relativistic model, we understand our failure as, again, not due to a lack of ingenuity on our part. We thus require any failure of  $p_t$  to be surjective to be "explained". Failure of  $p$  to be surjective should be explained by the lack of surjectivity of  $\phi^*$ .

The deterministic research program described in §1.1 suffers a huge setback with the requirement that the map  $p: \text{Lab} \rightarrow \text{Prob}(\mathcal{L}^S)$  factor

through the map  $\phi^* : \text{Prob}(\mathcal{L}^d) \rightarrow \text{Prob}(\mathcal{L}^S)$ . This is because  $\phi^*$  may have no pure probability measures in its image. This is, in fact, the situation for the standard quantum logic  $\mathcal{L}^d \simeq \text{Proj}(H)$ . The usual move adopted by physicists is to assume a basic collection of labels  $\text{Lab}_0$  which represent maximal knowledge of individual boxes and which go via  $p_t$  to extreme probability measures on  $\mathcal{L}^d$ . All other labels are to be described as classical mixtures of these basic labels. But is it fair to be prejudiced against a fair coin? We prefer not to require the image of  $\text{Lab}_0$  under  $p_t$  to consist of extreme probability measures. Thus, we may want to represent the state of a quantum system by a von Neumann density matrix and yet still claim no ignorance on our part! Similarly, we may want to describe a single sample of a "classical" gas by a probability measure on phase space. By assuming that the dynamics is governed by a Markov process instead of a Hamiltonian system, one may readily explain the macroscopic phenomena without having to invoke either ensembles or ergodic theorems (see §2). (This example really shows that what we are calling "classical mechanics" is really "quantum mechanics with a Boolean depth logic")!

With this attitude, one not only doesn't expect  $p_t$  to be injective on  $\text{Lab}$ , but one isn't even sure that it is reasonable to seek for a  $\text{Lab}_0$  on which it is injective. There may be many labels which go to the same probability measure on  $\mathcal{L}^d$ , but none of which can be physically described as classical mixtures of more basic labels. In any case, ignorance should only be ascribed to those labels which can be "redescribed" as classical mixtures. Ignorance should never be ascribed to probability measures on the deep logic. Yet, if one has done one's

work well, then it would seem reasonable to say of two labels  $\alpha, \beta \in \text{Lab}_0$ , such that  $p_t(\alpha) = p_t(\beta)$ , that they yield the same "state" of the individual boxes. Not to say so could only be justified by the belief in the existence of a larger theory containing more observables where  $\alpha$  and  $\beta$  would be seen to be empirically different. So long as our belief in a particular  $\mathcal{L}^d$  is firm, we will accept the equality  $p_t(\alpha) = p_t(\beta)$  as any "explanation" for the empirical observation that  $p(\alpha) = p(\beta)$ .

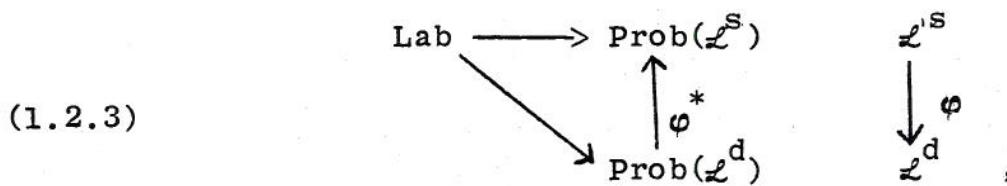
It may happen that several different theories yield reasonably accurate results. Then the decision between them may be based upon various non-scientific criteria-- historical, religious, personal, aesthetic, etc. One also often takes into account the pragmatic criteria of ease, speed, and cost of computation. It is these pragmatic criteria which induce astronomers to continue to use Newtonian mechanics instead of the much prettier general relativity of Einstein for orbital computations.

Given a theory  $T_1$ , one may seek to "reduce" it to a more basic, more fundamental, more expressive theory  $T_2$ . By a "reduction" we mean to imply that  $T_1$  is not found to be incorrect, only shallow. Thus, a reduction should include a redescription of the labels in  $T_1$  as classical mixtures of labels in  $T_2$ , i.e. a map  $\text{Lab}_1 \rightarrow \text{Lab}_2$ , and a redescription of the observables in  $T_1$  as observables in  $T_2$ , i.e. a map  $\mathcal{L}_1^d \rightarrow \mathcal{L}_2^d$ . One, of course, also requires that the diagram

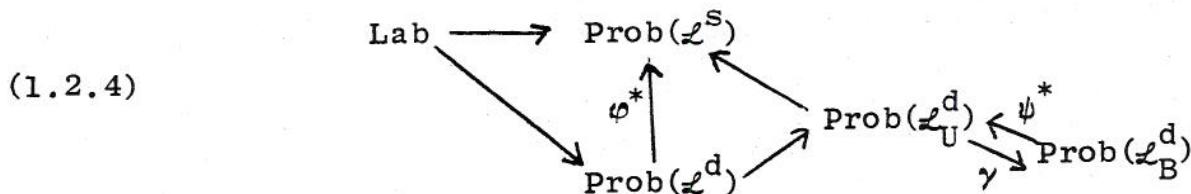
$$(1.2.2) \quad \begin{array}{ccc} \text{Lab}_1 & \xrightarrow{\hspace{2cm}} & \text{Lab}_2 \\ \downarrow & & \downarrow \\ \text{Prob}(\mathcal{L}_1^d) & \xleftarrow{\hspace{2cm}} & \text{Prob}(\mathcal{L}_2^d) \end{array}$$

commutes. Thus, instead of seeing two competing theories, the reduction lets us see two compatible theories-- the latter being the more profound one. Of course, this last value judgment would become confused if it were possible both to reduce  $T_1$  to  $T_2$  and also to reduce  $T_2$  to  $T_1$ ! In practice, reductions as defined above rarely occur since one of the main motivations of a reductionist is to obtain a more accurate-- "truer"-- theory. Thus, he will not expect diagram (1.2.2) to strictly commute. In other, more extreme situations, he will really be obtaining a totally different description of the world, and there might be virtually no basis for comparison of the theories-- neither on the depth level nor on the surface level. The two theories would really be about two entirely different types of boxes, i.e. two different worlds or realms of discourse.

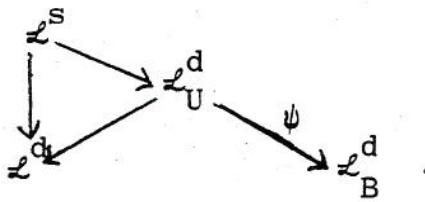
We can now begin to understand how it is possible to simultaneously have proofs that quantum mechanics admits of no hidden variables theories, while at the same time to have proofs that hidden variables theories always exist. The second proof is by far the simpler. Its essence is to simply observe that, given the situation



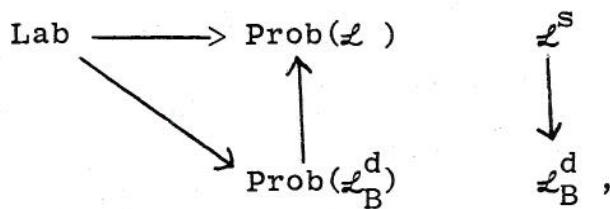
one can always factor it into



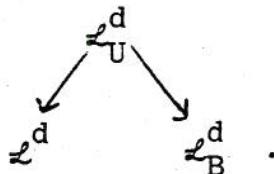
(1.2.5)



Here,  $\mathcal{L}_U^d$  is the "discrete" diagram of Boolean subalgebras associated to  $\mathcal{L}^d$  (i.e. forget the morphisms in  $\mathcal{L}^d$ ) and  $\mathcal{L}_B^d$  is the product of the Boolean algebras in  $\mathcal{L}_U^d$ .  $\psi$  in (1.2.3) is a right inverse to  $\psi^*$  obtained by forming the product measure. Thus, starting from (1.2.3) with a possibly non-Boolean  $\mathcal{L}^d$ , we can always obtain a diagram



with  $\mathcal{L}_B^d$  Boolean, and, furthermore, a redescription of all the probability measures on  $\mathcal{L}^d$  as classical mixtures of pure probability measures on  $\mathcal{L}_B^d$ . Note that we do not have a homomorphism  $\mathcal{L}^d \longrightarrow \mathcal{L}_B^d$ , but only the weaker relationship of a "correspondence"



The essence of this move is the "unpacking" of the original deep logic, i.e. we, in effect, say that  $\mathcal{L}^d$  identified to many surface observables which are really different. Thus, this move does not really yield a reduction, but, instead, a competitor. Our construction is, of course, too ad hoc to be mistaken for a "serious" competitor. But more specialized, less ad hoc models can always be found which will bear

a similar relationship to the original theory. This is the situation with the hidden variables models constructed by Bohm, Nelson and others. These theories are not really quantum theories, but, instead, competitor theories attempting - somewhat successfully - to redescribe the surface phenomena. We do not expect any of the present competitors of quantum theory to supplant it, but tolerance should be practiced since these alternatives often shed great light on the standard quantum mechanical description of phenomena, make us think more deeply, and may even suggest interesting experiments. In truth, the real motivation behind most hidden variables theorists is, not a love of Boolean logic, but, instead, the desire for a "realist" world view. They want to know how the "world" really works, not how it appears to work. In Einstein's words: "In the final analysis, the world only happens once". The realist would not care for our whole mode of discourse. His ideal would be a theory such as general relativity (interpreted realistically, of course). Putnam and, later, Bub have tried to interpret the quantum logic "realistically" as implying a revolution in "logic". This may be possible, but we don't feel that they have done the work necessary to make adequate sense of and to evaluate their suggestion. (For example, would one have to develop a quantum set theory?)

The fact that quantum mechanics doesn't admit a strict Boolean reduction follows from a deep theorem of Gleason (see §3). Gleason's theorem implies that the standard quantum logic,  $\text{Proj}(H)$ , admits no homomorphisms into any Boolean algebra. Similarly, Greechie has shown that one can construct logics which admit of no homomorphisms into either Boolean or (standard) quantum logics.

In order to obtain a more concrete understanding of what these theorems say, we shall now consider their implications for our "box" model of the scientific process. The theorem that hidden variables theories always exist implies that no matter how the boxes behave on the surface level, we can always fill in the boxes with Boolean devices so as to capture the same ensemble behavior, i.e. we can always design alternative boxes out of gears, circuits, and other classical Boolean devices, so that the external, surface behavior of our designed boxes is identical with that of "nature's" boxes. Thus, a Kelvin-Tait type program is always possible although one may doubt its usefulness. (Their original program for Maxwell's theory eventually 'fell into disrepute.) Hertz's statement that "Maxwell's theory is about Maxwell's theory" implied that one should consider one's tool box as containing more than just gears, i.e. electro-magnetic phenomena should be considered basic phenomena not requiring further explanation. In fact, today one is often more likely to attempt to explain the behavior of gears by invoking electro-magnetic phenomena than vice versa. If more than one button can be pressed at one time, then the Boolean move might require us to consider the pressing of several buttons simultaneously unrelated to the pressing of them individually, i.e. we might have a contextualistic theory. Thus, even if our boxes come in two halves, with one half on the moon and the other half on the earth, and they have no visible connections, we might still have to say that the results obtained by pressing a button on the part of the box on the earth may depend upon what buttons are simultaneously pressed on the moon. These "non-local" possibilities have been stressed by Bell. They require us to have "non-loca

Boolean devices if we want to build such boxes. Classical theory had plenty of non-local devices, e.g. the gravitational field, but these non-local objects were always viewed suspiciously by classical scientists. For example, Newton felt that anyone who really believed in action at a distance was not philosophically minded. Presumably, Einstein felt similarly. Bell has remarked that he had resolved the E-P-R paradox in a way that Einstein would have liked least. Thus, a high price might have to be paid in order to allow for a Boolean depth logic; for a realist, the price seems to be worth paying.

If one has a theory  $\mathcal{L}^S \xrightarrow{\phi} \mathcal{L}^D$ , and one requires that in the design of our artificial boxes we must make two buttons that  $\phi$  says are the same actually be the same (i.e. pressing either button yields the same result, not just the same statistics), then Gleason's theorem tells us that we cannot design such a box for a quantum mechanical system out of Boolean devices. We can, of course, design such a box using quantum mechanical devices such as lasers, electron beams, etc. If Greechie logics actually occurred, then one would need an even larger tool box of devices, since a Greechie box could not be built out of quantum mechanical devices.

In the aftermath of the downfall of the Kelvin-Tait program, one runs the risk of having to have a very heavy tool box of "fundamental devices". One tries to steer a course between viewing all devices as fundamental and viewing no device as fundamental.

### 1.3. Category Theory.

One of the main motivations for introducing the depth logic is that it is a way of identifying two surface observables, or, at least, parts of their outputs. A contemporary mathematician, attempting to formalize this idea, will naturally be led to the concept of a "diagram" of Boolean algebras. Furthermore, this assumption on the structure of  $\mathcal{L}^d$  is the weakest one still adequate to obtain all the results of quantum theory. In this section we will provide a short review of category theory. (See [50] or [51] for a systematic introduction to category theory.)

A category  $\mathcal{C}$  consists of objects  $X, Y, \dots \in \text{Ob}(\mathcal{C})$  and morphisms  $f, g, \dots \in \text{Mor}(\mathcal{C})$ . Every morphism  $f$  has a domain and a range,  $\text{dom}(f), \text{range}(f) \in \text{Ob}(\mathcal{C})$ . If the range of  $f$  is equal to the domain of  $g$ , then  $f$  and  $g$  can be composed to yield a morphism  $h = g \circ f$  from  $\text{dom}(f)$  to  $\text{range}(g)$ . Composition is associative and every object  $X$  has a unique identity morphism  $1_X$  from it to itself which is an identity for the operation of composition. (Actually, one can follow Lawvere, [43], and give an axiomatization of category theory which is independent of set theory, and then use category theory as a new foundation for mathematics.) Typical examples of categories are the category  $\mathcal{S}$  of sets and set mappings; the category  $\mathcal{G}$  of groups and group homomorphism; and the category  $\mathcal{T}$  of topological spaces and continuous mappings.

Given two categories  $\mathcal{C}_1$  and  $\mathcal{C}_2$ , a functor  $F$  from  $\mathcal{C}_1$  to  $\mathcal{C}_2$  is a pair of mappings  $F': \text{Ob}(\mathcal{C}_1) \rightarrow \text{Ob}(\mathcal{C}_2)$  and  $F'': \text{Mor}(\mathcal{C}_1) \rightarrow \text{Mor}(\mathcal{C}_2)$ , such that the category structure is preserved, i.e. domains, ranges, identity

mappings, and composition. Typical examples of functors are the "forgetful" functors  $N$  from  $\mathcal{G}$  or  $\mathcal{T}$  to  $\mathcal{A}$  which associate to every group or topological space its underlying set and to every homomorphism or continuous map its underlying set mapping. Given two functors  $F:\mathcal{C}_1 \rightarrow \mathcal{C}_2$  and  $G:\mathcal{C}_2 \rightarrow \mathcal{C}_3$ , one can form a composite functor  $H = G \circ F:\mathcal{C}_1 \rightarrow \mathcal{C}_3$ ; one thus obtains the category,  $\text{Cat}$ , of categories and functors.

Given two functors  $F, G:\mathcal{C}_1 \rightarrow \mathcal{C}_2$ , a natural transformation  $\eta:F \rightarrow G$  is a map  $\eta:\text{Ob}(\mathcal{C}_1) \rightarrow \text{Mor}(\mathcal{C}_2)$  such that for every  $X \in \text{Ob}(\mathcal{C}_1)$ ,  $\eta(X):F'(X) \rightarrow G'(X)$ , and if  $X \xrightarrow{f} Y$ , then the diagram

$$\begin{array}{ccc} F'(X) & \xrightarrow{\eta(X)} & G'(X) \\ \downarrow F''(f) & & \downarrow G''(f) \\ F'(Y) & \xrightarrow{\eta(Y)} & G'(Y) \end{array}$$

commutes. The construction which associates to a set the free group generated by that set determines a functor  $F:\mathcal{A} \rightarrow \mathcal{G}$ . The natural inclusion of a set  $X$  into (the underlying set of) the free group generated by  $X$  determines a natural transformation  $\eta:1_{\mathcal{A}} \rightarrow N \circ G$ . Similarly, one has a natural transformation  $\gamma:F \circ N \rightarrow 1_{\mathcal{G}}$ . One can also show that there is a natural bijection

$$(1.3.1) \quad \mathcal{A}(X, NG) \cong \mathcal{G}(FX, G) ,$$

where  $\mathcal{C}(X, Y)$  is the set of morphisms in  $\mathcal{C}$  from  $X$  to  $Y$ . Under these circumstances, one says that  $N$  and  $F$  are adjoint functors (the name

coming from the formal similarity of (1.3.1) with the formula defining adjoint linear operators on Hilbert space). The forgetful functor  $N:\mathcal{J} \rightarrow \mathcal{A}$  also has an adjoint (in fact, both a left and a right adjoint).

Given two categories  $\mathcal{C}_1$  and  $\mathcal{C}_2$ , we can form the functor category  $\mathcal{C}_2^{\mathcal{C}_1}$  whose objects are functors from  $\mathcal{C}_1$  to  $\mathcal{C}_2$  and whose morphisms are natural transformations of functors. Given a category  $\mathcal{C}$ , there is a natural associated "opposite" category,  $\mathcal{C}^{\text{op}}$ , obtained from  $\mathcal{C}$  by reversing all the arrows, i.e.  $\text{Ob}(\mathcal{C}^{\text{op}}) = \text{Ob}(\mathcal{C})$  and  $\mathcal{C}^{\text{op}}(X, Y) = \mathcal{C}(Y, X)$ . A functor  $F:\mathcal{C}_1^{\text{op}} \rightarrow \mathcal{C}_2$  is said to be a contravariant functor from  $\mathcal{C}_1$  to  $\mathcal{C}_2$ .

If  $\mathcal{C}$  is a category, then a diagram over  $\mathcal{C}$  is a functor  $X:I \rightarrow \mathcal{C}$ . We will usually use the notation  $\{X_{\alpha}\}$  to denote the diagram  $X$ , where  $X_{\alpha} = X(\alpha)$  for  $\alpha \in \text{Ob}(I)$ , the bonding morphisms being suppressed in the notation. If  $\{X_{\alpha}\}$  is a diagram over  $\mathcal{C}$ , then the direct limit  $\lim_{\rightarrow} \mathcal{C}(-, X_{\alpha})$  defines a contravariant functor from  $\mathcal{C}$  to  $\mathcal{A}$ , i.e. an object of  $\mathcal{A}^{\mathcal{C}^{\text{op}}}$ . Hence, one can form a category  $\text{ginj } - \mathcal{C}$ , whose objects are diagrams over  $\mathcal{C}$  and whose morphisms from  $\{X_{\alpha}\}$  to  $\{Y_{\beta}\}$  are natural transformations from  $\lim_{\rightarrow} (-, X_{\alpha})$  to  $\lim_{\rightarrow} (-, Y_{\beta})$ .  $\text{ginj } - \mathcal{C}$  is a full-subcategory of  $\mathcal{A}^{\mathcal{C}^{\text{op}}}$ .

We will now give a direct way of representing the morphisms of  $\text{ginj } - \mathcal{C}$ . Let  $\{X_{\alpha}\}_I$  and  $\{Y_{\beta}\}_J$  be diagrams over  $\mathcal{C}$ . Suppose given for each  $\alpha \in \text{Ob}(I)$  a  $\beta(\alpha) \in \text{Ob}(J)$  and a morphism  $f_{\beta(\alpha)}^{\alpha}: X_{\alpha} \rightarrow X_{\beta(\alpha)}$  in  $\mathcal{C}$  such that if we have a (not necessarily commutative) diagram

$$\begin{array}{ccc}
 & f^{\alpha'} & \\
 X_{\alpha}, & \xrightarrow{\beta(\alpha')} & Y_{\beta(\alpha')} \\
 \uparrow & X_{\alpha}^{\alpha}, & \uparrow Y_{\beta(\alpha')}^{\beta(\alpha)} \\
 & f^{\alpha} & \\
 X_{\alpha} & \xrightarrow{\beta(\alpha)} & Y_{\beta(\alpha)}
 \end{array}
 ,$$

where  $X_{\alpha}^{\alpha}$ , and  $Y_{\beta(\alpha')}^{\beta(\alpha)}$  are bonding morphisms, then there exists a  $\beta$  and a bonding morphism

$$Y_{\beta(\alpha')} \xrightarrow{\beta} Y_{\beta}$$

such that the compositions  $X_{\alpha}^{\alpha} \circ f_{\beta(\alpha')}^{\alpha'} \circ Y_{\beta}^{\beta(\alpha')}$  and  $f_{\beta(\alpha)}^{\alpha} \circ Y_{\beta(\alpha')}^{\beta(\alpha)} \circ Y_{\beta}^{\beta(\alpha')}$  are equal. Such a collection of maps  $\{f_{\beta(\alpha)}^{\alpha}\}$  determines a natural transformation

$$f : \varinjlim \mathcal{C}(-, X_{\alpha}) \rightarrow \varinjlim \mathcal{C}(-, Y_{\beta}),$$

and, hence, a morphism  $f$  in  $\text{ginj-}\mathcal{C}$  from  $X$  to  $Y$ . Every morphism in  $\text{ginj-}\mathcal{C}$  is so representable and it is even possible to define an equivalence relation on such representatives so as to give a direct definition of  $\text{ginj-}\mathcal{C}$  without thinking of it as a full sub-category of  $\mathcal{C}^{\text{op}}$ . This alternative definition is quite useful, e.g. it shows immediately that any functor  $F : \mathcal{C}_1 \rightarrow \mathcal{C}_2$  extends to a functor  $F : \text{ginj-}\mathcal{C}_1 \rightarrow \text{ginj-}\mathcal{C}_2$ .  $\mathcal{C}$  is a full sub-category of  $\text{ginj-}\mathcal{C}$  and the natural full embedding  $\text{ginj-}\mathcal{C} \rightarrow \mathcal{C}^{\text{op}}$  admits an adjoint. Given a diagram  $\{X_{\alpha}\}$  we can also form the functor  $\varinjlim \mathcal{C}(X_{\alpha}, -)$  and thus obtain a full sub-category  $\text{gpro-}\mathcal{C}$  of  $\mathcal{C}$ . The morphisms of  $\text{gpro-}\mathcal{C}$  also admit a direct definition. These categories are natural generalizations of the categories

$\text{inj-}\mathcal{C}$  and  $\text{pro-}\mathcal{C}$  introduced by Grothendieck [31]. We realize that the above brief discussion will not be very enlightening to most readers; our only excuse for not giving a longer, more leisurely discussion is that we were skeptical that it would be any more enlightening to the reader not already initiated into the mysteries of category theory.

In §1.1 we associated to each empirical observable  $E$  its finite set  $R_E$  of possible results and the associated finite Boolean algebra of "questions"  $\mathcal{Q}(R_E)$ . We also assumed that each state (i.e. label)  $\alpha$  determined a probability measure on  $\mathcal{Q}(R_E)$ . Since everything was finite, there were no technical complications. But, when one moves to the theoretical situation, one is often (in fact, almost always) motivated to allow the deep configuration space  $\mathcal{C}^d$  to be an infinite topological space (e.g., the real numbers  $R$ ) and to take for one's deep logic  $\mathcal{L}^d$  the Borel  $\sigma$ -algebra of  $\mathcal{C}^d$ ,  $\mathcal{B}(\mathcal{C}^d)$  (e.g., the Borel  $\sigma$ -algebra of  $R$ ), and to consider  $\sigma$ -additive probability measures on  $\mathcal{L}^d$  (e.g., Gaussian measure on  $\mathcal{B}(R)$ ). An historical example of this move is the relationship between the binomial and normal distributions. The normal distribution was originally viewed as simply a curve fit to the binomial distribution; only later was it viewed as defining its own real-valued random variable. Real-valued random variables are conceptually much more complicated than discrete-valued random variables though, as in the case of the normal distribution vis à vis the binomial distribution, often much simpler from a computational viewpoint. Thus, we will follow tradition and take as our basic (classical) logical category,

the category  $\mathcal{B}$  of Boolean  $\sigma$ -algebras and  $\sigma$ -homomorphisms. Probability measures will be understood to be countably additive. Usually our logics will arise as the Borel algebras of nice topological spaces. The general (non-classical) logical category we will use will be  $\text{ginj-}\mathcal{B}$ , i.e., the category whose objects are diagrams of Boolean  $\sigma$ -algebras. If  $\mathcal{L} = \{B_\alpha\}$  is a logic, then a probability measure  $p$  on  $\mathcal{L}$  is a family  $\{p_\alpha\}$  of probability measures on the  $B_\alpha$  such that if  $f_\beta^\alpha : B_\alpha \rightarrow B_\beta$  is a bond in  $\mathcal{L}$ , then  $p_\alpha = p_\beta \circ f_\beta^\alpha$ . The convex set of all such probability measures on  $\mathcal{L}$  will be denoted  $\text{Prob}(\mathcal{L})$ . It is easy to check that a homomorphism  $\mathcal{L} \rightarrow \mathcal{L}'$  in  $\text{ginj-}\mathcal{B}$  induces a convexity preserving map  $\text{Prob}(\mathcal{L}') \rightarrow \text{Prob}(\mathcal{L})$ . A homomorphism  $\mathcal{L} \rightarrow \mathcal{L}'$  will also be called an  $\mathcal{L}$ -valued random variable on  $\mathcal{L}'$ . A map  $\{\text{Prob}(B'_\alpha)\} \rightarrow \{\text{Prob}(B_\beta)\}$  in  $\text{gpro-}\mathcal{C}$ , where  $\mathcal{C}$  is the category of convex sets and convexity-preserving maps, will be called a noisy homomorphism or a noisy  $\mathcal{L}$ -valued random variable on  $\mathcal{L}'$ , and it will be denoted by  $\mathcal{L} \dashrightarrow \mathcal{L}'$ . Since  $\text{Prob}(\mathcal{L}) = \varprojlim \{\text{Prob}(B_\alpha)\}$ , a noisy homomorphism  $\mathcal{L} \dashrightarrow \mathcal{L}'$  also induces a map  $\text{Prob}(\mathcal{L}') \rightarrow \text{Prob}(\mathcal{L})$ . For  $\mathcal{L} = \mathcal{B}(R)$ , such homomorphisms will be called (noisy) real-valued random variables. (One could even motivate going so far as to change one's basic category from  $\mathcal{B}$  to the category of Boolean algebras and noisy homomorphisms. Addition and multiplication on  $R$  induce on the set  $\text{ginj } \mathcal{B}(\mathcal{B}(R), \mathcal{L})$  the structure of a partial commutative algebra over  $R$  (i.e. operations are only sometimes defined, but when they are, everything is classical (commutative)). This partial algebra is also equal to  $\varinjlim \mathcal{B}(\mathcal{B}(R), \mathcal{B}_\alpha)$ , where the direct limit is taken in the category of partial algebras. From this viewpoint, the move from a classical to a quantum mechanical system is not a move from

a commutative to a non-commutative algebra  $\mathcal{A}$  of real-valued observables, but, instead, a move from a commutative algebra to a partial commutative algebra of observables. Of course, every non-commutative algebra determines an underlying partial commutative algebra and also its diagram of commutative subalgebras. The fact that assuming the structure of a non-commutative algebra is the wrong assumption has already been observed in the literature (see, for example, [19]), but it is often replaced by another wrong assumption, namely that of assuming the structure of a Jordan algebra. These differing assumptions on the structure of  $\mathcal{A}$  affect the size of its automorphism group and, hence, of the allowable symmetries of the system (the weaker the assumed structure on  $\mathcal{A}$ , the larger is its automorphism group). In §3 we will discuss these issues more fully.

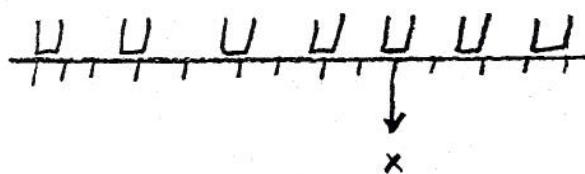
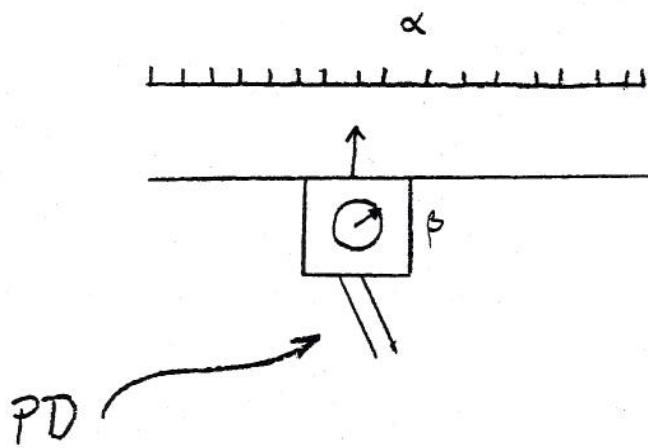
The contravariant functor  $\beta$  which associates to every topological space  $X$  its  $\sigma$ -algebra of Borel subsets extends to a functor  $\beta: \text{gpro-}\mathcal{T} \rightarrow \text{ginj-}\mathcal{B}$ . Thus, if we start by assuming the existence of a deep configuration space  $\mathcal{C}^d = \{\mathcal{C}_\alpha\} \in \text{gpro-}\mathcal{T}$  together with a (measurable) morphism  $\mathcal{C}^d \xrightarrow{x} \mathcal{C}^s$ , application of  $\beta$  yields a deep logic  $\mathcal{L}^d = \mathcal{B}\mathcal{C}^d$  together with a morphism  $\mathcal{L}^s = \mathcal{B}\mathcal{C}^s \xrightarrow{\beta x} \mathcal{B}\mathcal{C}^d = \mathcal{L}^d$ . One advantage of starting with  $\mathcal{C}^d$  is that using it one can define the partial algebra of continuous real-valued observables  $\mathcal{A}_c$  and not just the partial algebra of measurable observables  $\mathcal{A}_m$ . When one notes that all classical systems with manifold phase spaces have isomorphic algebras of measurable observables, but non-isomorphic algebras of continuous observables, one sees that  $\mathcal{A}_c$  captures a lot more of the physics of a classical system than does  $\mathcal{A}_m$ . Thus, I. E. Segal's approach to quantum mechanics via

assuming that  $\mathcal{A}_c$  is isomorphic to the partial algebra of Hermitian elements in a  $C^*$ -algebra, besides being a much simpler assumption than would be required on  $\mathcal{A}_m$ , is also a physically much richer assumption, since many non-isomorphic  $\mathcal{A}_c$ 's correspond to isomorphic  $\mathcal{A}_m$ 's.

#### 1.4 Symmetry Principles.

It should be reasonably clear to the reader that much stronger assumptions will have to be made before the formalism described in §1.3 can be made to yield interesting results. Symmetry principles are among the most powerful and beautiful ways of recovering the standard results of both classical and quantum mechanics. In this section we will introduce and try to motivate certain general symmetry assumptions. These ideas will then be further pursued in §§2 and 3.

Consider the situation pictured below:



(1, 4, 1)

We have a preparing device (PD) capable of being placed in various positions  $\alpha \in R$  and of being set at various settings  $\beta \in B$ ; thus, our set of basic labels is just  $\text{Lab}_0 = R \times B$ . We also have an array of detectors (D) placed at various positions  $x \in R$ ; thus, for instance, our surface logic might be  $\mathcal{L}(R)$ .  $R$  acts naturally on  $\text{Lab}$  and we may want to assume that besides this surface action on  $\text{Lab}$ , there is also a deep action of  $R$  on  $\text{Prob}(\mathcal{L}^d)$  such that the natural map  $p_t : \text{Lab} \rightarrow \text{Prob}(\mathcal{L}^d)$  is equivariant. This is a strong association! Of course, there is nothing sacred in the above about the group  $R$ --any group  $G$  may be appropriate; for example, the Euclidean group, the Galilean group, and the Lorentz group are often used.

In the situation pictured in (1.4.1) it will often be discovered that translating the preparing device by  $y \in R$  has the same effect as translating the array of detectors by  $-y$ . This type of observation often leads one to even assume that the symmetry group  $G$  also acts on  $\mathcal{L}^d$  in such a way that the natural map  $\varphi : \mathcal{L}^S \rightarrow \mathcal{L}^d$  is equivalent and the induced action on  $\text{Prob}(\mathcal{L}^d)$  is just the original action.

The group occurring most frequently is the time translation group  $T = R^1$ . Consider loaded coins whose probability of heads is  $\alpha(t) = \alpha + t \pmod{1}$ , where  $\alpha \in [0,1]$ .  $T$  is certainly acting nicely on the set  $\{(\alpha, t)\} = \text{Lab}_0$ ; but unless we pack a great deal into  $\mathcal{L}^d$ , it won't act in a compatible convexity-preserving way on  $\text{Prob}(\mathcal{L}^d)$ . This is because many different classical mixtures of these coins have the same ensemble statistics if our only allowable experiments are a single coin toss. In fact, by considering weighted perishable coins (after one toss they disappear), one is led to introduce a set  $\mathcal{J}$  of theoretical states and mappings  $\text{Lab}_0 \rightarrow \mathcal{J} \rightarrow \text{Prob}(\mathcal{L}^d)$ . Group actions may only extend from  $\text{Lab}_0$

to  $\mathcal{L}$  and not all the way to  $\text{Prob}(\mathcal{L}^d)$ . This example should be kept in mind to remind the reader of the strength of the assumption that  $G$  acts on  $\text{Prob}(\mathcal{L}^d)$  or even on  $\mathcal{L}^d$  itself. In the case of "irreversible" systems, one can only assume that the semi-group  $T^+ = R^+$  acts on  $\text{Prob}(\mathcal{L}^d)$ . If one starts with a basic map  $\text{Lab}_0 \rightarrow \text{Prob}(\mathcal{L}^d)$ , then even if the action of  $G$  extends from  $\text{Lab}_0$  to  $\text{Prob}(\mathcal{L}^d)$  there is no a priori reason to assume that the action of  $G$  on  $\text{Prob}(\mathcal{L}^d)$  is convexity preserving.

If we go all the way and assume that  $G$  acts on  $\mathcal{L}^d$ , then, of course, it also acts by convexity-preserving automorphisms on  $\text{Prob}(\mathcal{L}^d)$ . These subtle distinctions will soon be very important.

We've already mentioned in §1.1, that the breakup of the description of an experiment into the two parts  $(\alpha, E)$  has a conventional aspect. In our discussion of time, above, we've taken a "Schrodinger" viewpoint and had the time group act directly on the states. An alternative approach is to assume that time acts directly on one's observables.

If  $\mathcal{L}^S$  is the surface logic of the Schrodinger approach, then  $\{\mathcal{L}_t^S\}_{t \in T}$  will be the surface logic in the Heisenberg approach, i.e. one copy of  $\mathcal{L}^S$  for each time  $t \in T$ . A similar move can be made for other groups and semi-groups. Thus, if a Schrodinger approach used a deep logic  $\mathcal{L}^d$ , then the Heisenberg approach might use a deep logic  $\{\mathcal{L}_t^d\}$ .  $T$  acts in a compatible way on  $\{\mathcal{L}_t^S\}$  and  $\{\mathcal{L}_t^d\}$ , and, hence, induces a convexity-preserving action on  $\text{Prob}(\{\mathcal{L}_t^d\})$ . At this point, there is no real difference between these two approaches. But the Heisenberg approach suggests using a more interesting deep logic  $\mathcal{L}_H^d$  which contains  $\{\mathcal{L}_t^d\}$ . Then the two approaches become distinct!

Consider, for example, the case of a "particle". In this case it is reasonable to start from a basic position observable which yields a result in  $R^3$  and to assume that our system is invariant under the full Euclidean group  $\delta_3$ . Thus, we are looking for systems of imprimitivity for  $(\beta(R^3), \delta_3)$  in a deep logic  $\mathcal{L}^d$ . We may now add time to our picture by assuming that  $T$  acts on  $\text{Prob}(\mathcal{L}^d)$  (Schrodinger picture), or by assuming that we are actually looking for a system of imprimitivity for  $(\{\beta(R^3)_t\}, \delta_3 \times T)$  (Heisenberg picture). In fact, one should go further and replace the Newtonian space-time group  $\delta_3 \times T$  by either the Galilean or Lorentz group. Two other possibilities also occur as one contemplates the space-time point of view: namely, a theory of "events" and a theory of "trajectories". For a theory of events one should look for systems of imprimitivity for  $(\beta(R^4), G)$ , where  $G$  is an appropriate space-time group acting on  $R^4$  (e.g., the Lorentz group). For a theory of trajectories one might look for systems of imprimitivity for  $(\beta(L), G)$ , where  $L$  is the 6-dimensional homogeneous space of time-like lines in  $R^4$  and  $G$  is the (inhomogeneous) Lorentz group. One can, of course, go even further and consider more complex hyperspaces of subsets of  $R^4$ . Also, one can consider diagrams of logics  $\{\mathcal{L}_A\}$  indexed by subsets of  $R^4$ . Such objects are needed in field theory and the statistical mechanics of infinite systems. These ideas will be further pursued in §§2 and 3.

## §2. Classical Theories.

In this section we shall discuss various systems whose depth logic  $\mathcal{L}^d$  is either Boolean or a disjoint union of Boolean logics.  $\mathcal{L}^d$  will be obtained here as the logic associated to a deep configuration space  $\mathcal{C}^d$ .

### §2.1. Statics

The simplest non-trivial example is where  $\mathcal{C}^d = Z_2 = \{0,1\}$ .  $Z_2$  might be the appropriate configuration space for describing a coin tossing experiment. A probability measure  $\mu$  on  $Z_2$  is completely determined by its value on 0,  $\mu(0) \in [0,1]$ . If our experiment consisted of having our hand over a coin and then looking to see whether heads or tails was up, then it would be reasonable to use the probability measure with  $\mu(0) = 1/2$  and to say that the probability was fifty-fifty for a head or a tail. Clearly, our use of probability here is due to our ignorance of the true state of the coin until we lift our hand. On the other hand, if our experiment consisted of tossing a coin and then seeing whether it landed up heads or tails, then, again, it would be reasonable to say that the probability was fifty-fifty for a head or a tail to appear. But it would no longer be clear that our use of probability language was due to any ignorance on our part. One could easily consider tossing a coin to be an intrinsically stochastic process; and, in fact, this is the common sense point of view. On the other hand, a true believer in the classical world view could argue that the

result of a coin toss was as determined by the initial conditions of its environment as was the result of our lifting our hand, the "only" difference being one of computational difficulty. Thus, while in practice it is usual - even in classical probability theory - to often consider intrinsically stochastic processes, it is always possible in a classical theory to assume that our use of probability is only a reflection of our ignorance and/or computational weakness.

Moving on, we next consider a configuration space appropriate to describing a particle. If our experiment consisted of taking a flash picture of a single billiard ball on a pool table in a dark room, then the appropriate configuration space would be  $R^2$  or even a rectangular subset of  $R^2$ . Clearly, any use of probability language here is due to our ignorance of the actual location of the billiard ball. On the other hand, if the experiment consisted of tossing a rock over our shoulder into a lake and having a friend record the location of the splash, then, again,  $R^2$  would be an appropriate configuration space for our system, but the system could easily be conceived of as being intrinsically stochastic. We will see in §2.3 that thinking of atomic models more along the lines of the above rock tossing experiment than of the billiard ball experiment allows one to carry out Boltzmann's program for classical statistical mechanics without having to invoke either ensembles or ergodic theorems. Suppose now that we throw two rocks, one colored red and the other blue. If our friend is not color blind, then he could record the location of the splash of the red rock and the location of the splash of the blue rock. Hence,  $R^2 \times R^2$  would be

an appropriate configuration space for this system. If our friend is color-blind, then he would only be capable of recording the locations of the two splashes, but not which was caused by the red rock. In this situation, the surface configuration space  $\mathcal{C}^S$  could be chosen to be the space of unordered pairs of points (repetitions allowed) in  $R^2$  (this space is isomorphic to  $R^2 \times R^2 / Z_2$ , i.e. the quotient space of  $R^2 \times R^2$  by the natural interchange action of  $Z_2$  on it). Clearly, the use of this  $\mathcal{C}^S$  instead of  $R^2 \times R^2$  is entirely due to our friend's ignorance, i.e. his being color-blind, and we could choose  $R^2 \times R^2$  for our deep configuration space  $\mathcal{C}^D$  and  $\psi: \mathcal{C}^D \rightarrow \mathcal{C}^S$  to be the natural quotient map. On the other hand, if our experiment consisted of photographing the location of bubbles on a turbulent lake, then an appropriate surface configuration space might be the space of finite subsets of a region in  $R^2$ . There is no natural way to label bubbles or to talk about "interchanging" a pair of bubbles. Bubbles are momentary, unstable, localized structures in the lake. A move analogous to that made above of using  $R^2 \times R^2$  as a deep configuration space, while possible here, would be unnatural and inappropriate here. The use of the configuration space of finite subsets of a region in  $R^2$  in these circumstances should not be attributed to our ignorance, but to the system under consideration. If one were interested in describing the water itself, or some other continuous medium such as a deformed string or membrane, then the appropriate configuration space would be some infinite-dimensional function space.

§2.2. Dynamics,

So far we have said nothing about time and the dynamics of our systems. Suppose now that our billiard ball is rolling on the billiard table and that our camera has a timing device which allows us to take a single flash photograph at any of a finite number  $T$  of possible time delays. Then the appropriate surface configuration space would consist of  $T$  disjoint copies of a rectangle:

$$\mathcal{C}^S = \{R_t\}_{t \in T} = \left\{ \begin{matrix} \square \\ t=1 \end{matrix} \quad \begin{matrix} \square \\ t=2 \end{matrix} \quad \dots \quad \begin{matrix} \square \\ t=T \end{matrix} \right\} .$$

Of course, we really believe that the billiard ball is a "lasting particular" and that one could take the function space  $R^T$  as the deep configuration space  $\mathcal{C}^d$  with the evaluations at various times determining the map  $\psi: R^T \rightarrow \{R_t\}$  in gpro-Top. Taking  $R^T$  as the deep configuration space of this system leads to a research program of attempting to define empirical observables whose associated surface configuration space would include a copy of  $R^T$ . So, for example, one would attempt to develop a camera which could take  $T$  pictures after one setting. Success in this venture would lift  $R^T$  from being the deep configuration space up to its also being (a part of) the surface configuration space.

On the other hand, if our system consists of setting a timing device which then - at the appropriate time - flings a rock into water and outputs a photograph of the splash, then, again,  $\{R_t\}$  would be

the appropriate surface configuration space, but we would no longer consider  $R^T$  as an appropriate deep configuration space since there is no "splash" which is lasting over time.  $R^T$  would be the appropriate configuration space for a system which flings rocks at times  $t = 1, 2, \dots, T$  and outputs  $T$  photographs of their splashes (with the time  $t$  printed on the bottom of the photograph). Here, there is no lasting particular, but there is a path!

If time is modeled as the continuum  $R^1$ , then there is again no problem in defining the surface configuration space to be  $\{\mathcal{C}_t\}_{t \in R^1}$ , where  $\mathcal{C}_t$  is the configuration space of possible results at time  $t$ . If we now want to model a lasting particular, such as a billiard ball, we must choose some appropriate function space which we will denote by  $\mathcal{C}^T$ .  $\psi: \mathcal{C}^T \rightarrow \{\mathcal{C}_t\}$  is again determined by evaluation at the various times. A theory will tell us which probability measures on  $\mathcal{C}^T$  are actually empirically realizable. For describing a billiard ball, one would naturally tend only to include continuous or even piecewise smooth paths in  $\mathcal{C}^T$ .

One simple way of generating an interesting dynamical theory is to assume that one is given a continuous one-parameter group  $U_t$  of homeomorphisms of a space  $\mathcal{C}$ . The association  $c \rightarrow U_t c$  determines a map  $\mathcal{C} \rightarrow \mathcal{C}^T$  and hence a map  $\text{Prob}(\mathcal{C}) \rightarrow \text{Prob}(\mathcal{C}^T)$ . If  $U_t$  is a smooth action, then it is completely determined by its infinitesimal generator  $X$  which is a smooth vector field on  $\mathcal{C}$ . Thus, in Aristotlean mechanics one might try to describe the behavior of a billiard ball by giving a vector field on  $R^2$ . If the surface on which the ball moved was sufficiently rough, then this description might even be reasonably

accurate. But most of the time it would turn out to be very poor, as was first fully realized by Galileo. A billiard ball seems to have other intrinsic qualities beyond mere position. This idea can be formalized by assuming that the actual deep configuration space appropriate for a billiard ball is a space  $\mathcal{C}'$  together with a continuous map  $\pi: \mathcal{C}' \rightarrow \mathbb{R}^2$ . The simplest way to obtain such a pair  $(\mathcal{C}', \pi)$  is to assume that the internal qualities of the billiard ball are described by a space  $F$  and take  $\mathbb{R}^2 \times F$  for  $\mathcal{C}'$  and the natural projection map for  $\pi$ . More generally, one might describe the qualities of a system by a space  $T(\mathcal{C})$  which fibers over the space  $\mathcal{C}$  of external qualities.

The symmetry properties of  $\mathcal{C}$  and  $T(\mathcal{C})$  play an important role in the development of many theories. For example, for a particle the appropriate external configuration space would be  $\mathbb{R}^3$  and the natural action of the Euclidean group  $E_3$  on  $\mathbb{R}^3$  plays a very important role in most theories of particles. It is natural to suppose in this case that the natural action of  $E_3$  on  $\mathbb{R}^3$  lifts to a compatible action on the total configuration space  $T(\mathbb{R}^3)$ . More generally, if  $G$  is a symmetry group acting on the external configuration space  $\mathcal{C}$ , then we will assume that the action of  $G$  lifts to  $T(\mathcal{C})$ . If we take for  $G$  the full diffeomorphism group of  $\mathcal{C}$ , then the possibilities for  $T(\mathcal{C})$  are probably quite limited. Some possibilities are: a) the tangent and cotangent bundles of  $\mathcal{C}$  and tensor products of these vector bundles; b) the projective and Grassmann bundles associated to the examples in a); c) various jet bundles. It would

be interesting to actually determine all (reasonable) possibilities.

A vector field  $X$  on  $T(\mathcal{C})$  determines a trajectory  $c(t) \in T(\mathcal{C})^{\mathbb{R}^1}$  for every point  $c \in T(\mathcal{C})$ . Composing with the map  $\pi:T(\mathcal{C}) \rightarrow \mathcal{C}$  yields  $\pi \circ c(t) \in \mathcal{C}^{\mathbb{R}^1}$ , and hence we obtain a map from the full static configuration space  $T(\mathcal{C})$  to the external dynamic configuration space  $\mathcal{C}^T$ . In Hamiltonian dynamics one also assumes that group actions on  $T(\mathcal{C})$  are generated by Hamiltonian vector fields which are themselves determined by smooth observables. In particular, one assumes that  $T(\mathcal{C})$  comes equipped with a natural symplectic structure and that all groups act via symplectomorphisms (i.e., contact transformations). The simplest example here is the cotangent bundle  $T^*\mathcal{C}$  of  $\mathcal{C}$ . What are the other possibilities?  $T^*\mathcal{C}$  is the natural object to take if one thinks of translations on  $\mathbb{R}^n$ . What extended structures correspond to the full Euclidean group? (See §3)

Taking  $T^*\mathcal{C}$  for one's deep static configuration space and choosing a Hamiltonian  $H:T^*\mathcal{C} \rightarrow \mathbb{R}$  to determine the dynamics yields most of the special structure characteristic of classical mechanics. In particular,  $H$  is usually chosen to be the sum  $T+V$  of two parts:  $T$ , the kinetic energy, which is determined by a Riemannian metric on  $\mathcal{C}$ ; and  $V$ , the potential energy, which is determined by a scalar function on  $\mathcal{C}$ . Isometries of  $(\mathcal{C}, T)$  which leave  $V$  invariant lift to  $T^*\mathcal{C}$  in such a way as to be invariant under the dynamics determined by  $H$ ; their generators are thus invariants of the motion, i.e., we obtain the classical conservation laws (e.g., those of conservation of energy, momentum, and angular momentum). A similar Hamiltonian theory for systems of infinitely many degrees of freedom is still

under development.

In the theories above we have a static total configuration space  $\mathcal{C}'$  and a map  $\psi': \mathcal{C}' \rightarrow \mathcal{C}^{R'}$  which induces a map  $\psi'_*: \text{Prob}(\mathcal{C}') \rightarrow \text{Prob}(\mathcal{C}^{R'})$ . The dynamics is completely deterministic and any recourse to the use of probability is due to starting with a probability measure with dispersion on  $\mathcal{C}'$ . This may be necessary due to our ignorance of the "actual" initial state in  $\mathcal{C}'$ . In Wiener's theory of Brownian motion, we are also given a map  $x_*: \text{Prob}(\mathcal{C}') \rightarrow \text{Prob}(\mathcal{C}^{R^1})$ , but  $x_*$  is no longer induced by a map  $x: \mathcal{C}' \rightarrow \mathcal{C}^{R^1}$ , i.e., the dynamics is intrinsically stochastic. In this case, unlike the two above,  $\mathcal{C}'$  enters only as a computational device which enables us to compute the appropriate probability measures on  $\mathcal{C}^{R^1}$ . Generally speaking, physicists have taken the phase space  $T^*\mathcal{C}$  much more seriously perhaps too seriously, (but see the next section). Hamiltonian mechanics is a powerful and beautiful theory, but it has led to an overemphasis on  $T^*\mathcal{C}$  in contrast to  $\mathcal{C}^{R^1}$ . The essence of the notion of a particle is that of a lasting particular, and  $\mathcal{C}^{R^1}$  is the appropriate configuration space for describing a lasting particular.

### §2.3. Classical Statistical Mechanics.

There are two basic approaches to the study of systems such as a container of gas. In thermodynamics one takes a phenomenological approach and treats the gas as a basic entity and develops macroscopic observables - such as temperature, pressure, volume, density - and one attempts to discover laws governing their behavior - such as Boyle's Law and Charles' Law. On the other hand, one might attempt to model the gas as an enormous collection of very small molecules. The molecules could, in first approximation, be viewed as point particles or as hard spheres obeying Newton's Laws or, better yet, Hamiltonian mechanics. One would thus describe the state of the gas by a point in  $R^{6N}$ , where N is the number of molecules in the gas. Observables would correspond to functions on  $R^{6N}$ , and the dynamics of the gas would be governed by a Hamiltonian. The goal of statistical mechanics is to show how the phenomenological observables and laws of thermodynamics can be "reduced" to corresponding observables and laws of the molecular model. So far, all such reductionistic attempts have failed. The so-called paradoxes of reversal and recurrence show that, at best, one can only obtain from the molecular Hamiltonian model that the thermodynamical description is very probable, but not necessary. "Fluctuations" away from equilibrium will occur and will be important. This, by itself, is already an important accomplishment in that it focuses our attention on macroscopic phenomena that we may have overlooked as "just noise". But even this "at best" requires that the Hamiltonian flow be ergodic.

on the energy hypersurface - a result which was only comparatively recently proved by Sinai in the special case of hard spheres in a box. Backing off a bit, one can follow Gibbs' approach based upon ensembles of similarly prepared systems and try to show that the thermodynamical laws, while not strictly true of individual systems, are true of equilibrium ensembles. An ensemble will be described by a probability measure on the phase space  $R^{6N}$ . Each system in the ensemble is to be thought of in terms of the above molecular model. Hence, appeal to probability language here is strictly pragmatic. The individual system will usually not be in equilibrium; but if the Hamiltonian flow is ergodic, then almost any ensemble will approach equilibrium in the sense that its probability measure on  $R^{6N}$  will "approach" the unique invariant probability measure.

A very different viewpoint is suggested by our development in §1. Instead of viewing a molecular model in terms of lots of little billiard balls, we can instead view it as implying a mode of perception (i.e. the (at least theoretical) existence of an appropriate microscope) capable of looking at a gas and seeing a large collection of "molecules"; the result of such a look is to be recorded as a point in  $R^{6N}$ . From this point of view, one may very well describe a single sample of gas by a probability measure on  $R^{6N}$ , while disclaiming any lack of knowledge. The process of looking may be as intrinsically stochastic as the process of tossing a coin. One may now question whether the dynamics of the system should be described by a Hamiltonian flow. Possibly a Markov process or some other more general stochastic process may be appropriate for generating the time semigroup on

$\text{Prob}(R^{6N})$ . Maybe, even, the probability measures satisfy the Boltzmann equation! Now it is possible for a single sample of gas to be in equilibrium while at the same time repeated observations of it yield fluctuations (think of repeated tossings of a hundred fair coins)! Thus, it is possible to have a "molecular model" of a gas with a classical depth logic and to obtain from it both the laws of thermodynamics and also fluctuation laws without having to appeal to either ergodic theorems or to Gibbs' ensembles! Of course, one has here given up the classical atomic ontology. Our molecules do not have independent existence; they are more like the "splashes" of §§2.1 and 2.2. Had Boltzmann suggested such a move, he probably would have been laughed at. But such a move is now, in any case, necessitated by quantum mechanics. So, if it also helps when the deep logic is Boolean, why not make it?

§2.4. Field Theory .

The theory of continuous models was developed in the eighteenth and early nineteenth centuries by L. Euler et. al in order to describe the behavior of such diverse systems as vibrating strings, membranes and other elastic structures; heat conduction; sound waves; diffusion processes; fluid flow. All these systems were actually viewed as being fundamentally discrete (i.e. made up of atoms), and the continuous models were viewed as merely phenomenological theories (i.e. approximations not having fundamental significance). It wasn't until Faraday's conception of electric and magnetic lines of force that a field theory was viewed as fundamental (i.e., as a basic building block of the universe, having the same ontological status as atoms). But Faraday's ideas weren't well understood by his contemporaries. His mode of expression was essentially that of a twentieth century geometric topologist and not that of a nineteenth century mathematical physicist. Maxwell translated Faraday's intuitions into the more acceptable language of partial differential equations and founded the theory of the electro-magnetic field. (The relationship between the work of Faraday and Maxwell is essentially identical with the relationship between algebraic and differential topology in the d'Rham theory.) But even for Maxwell, the electro-magnetic field was viewed as essentially a phenomenological device similar to temperature. He sought for underlying mechanistic atomic models. It wasn't until Hertz that the true fundamental significance of the electro-magnetic field was fully realized. With the advent of relativity theory ontology had come full circle since in

relativity theory it is very natural to only consider fields as having basic ontological existence. There is no way to end this dialectic since one can probably always reduce discrete models to continuous ones and vice versa. In quantum mechanics one settles the dialectic with the slogan that the two approaches are "complementary". (In the classical case one also has, of course, that the appropriate phase spaces are isomorphic as Borel spaces.)

In the middle of the twentieth century, due mainly to the work of N. Wiener and A. Kolmogorov, one had the development of a statistical theory of random fields. The goal of this theory is to describe such complex phenomena as noise and turbulence by direct statistical assumptions on the properties of the fields. For example, one might assume that a fluid flow is to be described by a vector field on  $R^2$ , i.e. by a function from  $R^2$  to  $R^2$ . On the space of such vector fields  $\mathcal{F}$  one assumes the existence of a probability measure  $\mu$  which determines the statistical properties of the fluid under consideration, i.e. if we measure the fluid flow, then the resulting vector field will lie in a Borel set  $A \subset \mathcal{F}$  with probability  $\mu(A)$ . One might naturally investigate first the statistically homogeneous flows, i.e. those whose probability measures are invariant under the induced action of the Euclidean group  $E_2$  on  $\mathcal{F}$  from its natural action on  $R^2$ . Since  $\mu$  is a fairly complicated object, most work is done by simply making assumptions concerning the moments of the velocity field, e.g. assumptions on  $E(v_i(x))$ ,  $E(v_i(x)v_j(y))$ , etc. Recently, attempts have been made to reduce this intrinsically stochastic theory back to the deterministic

theory of the Navier-Stokes equations. This work has been buoyed up by the expectation that the flow described by these equations contains a "strange attractor". A dynamical system with such an attractor would have a very complex time evolution. But turbulence is also very complicated at a single time. Recalling the situation in classical statistical mechanics, one would expect that the deterministic model could, at best, obtain the same results as the statistical theory only "most of the time". It seems most unlikely that the statistical assumptions made by the statistical theory (e.g. homogeneous chaos) could be derived from the deterministic theory. If they could, one would, of course, feel that one had gained a greater understanding of such phenomena as turbulence. We guess that one can't, in general, obtain such a derivation and that, furthermore, experiment will support the statistical approach. In other words, our attitude towards turbulence is similar to our attitude towards a gas (and also a coin toss); they are both intrinsically stochastic phenomena. From a pragmatic point of view, this has certainly been the approach of applied mathematicians attempting to model such complicated phenomena. But we are going further and suggesting that these stochastic models may be fundamental, and not just useful because of our computational weakness.

Such stochastic theories as homogeneous chaos have been lately arising quite frequently. For example, besides its occurrence in the statistical theory of turbulence, it is also used in the statistical mechanics of infinite systems and in the Euclidean approach to quantum

field theory (see §3). Recently, Mandelbrot published a fascinating book using a general theory of random sets to describe such diverse phenomena as cloud patterns, shapes of island boundaries and rivers, the pattern of arteries and veins in the lungs. In these models, the sample set is monstrous with probability one. The prototype is, of course, Brownian motion with its nowhere differentiable sample paths. Mandelbrot's samples usually have Hausdorff dimension greater than their covering dimension; hence, the samples are spaces that one usually only thinks of as occurring in courses in general topology. But Mandelbrot makes a strong case for the claim that such monsters are all about us if we only care to look. He even suggests that physicists have independently discovered the notion of Hausdorff dimension in trying to describe some of the phenomena he considers.

We haven't even touched upon the biological and social sciences where almost all models are (or should be) stochastic. Most novel in some of these models is the use of game theory to describe conflict situations. But, to our knowledge at least, all of the models used so far in the biological and social sciences can be formulated in terms of classical (Boolean) probability theory. In the next section we discuss non-classical theories.

### §3. Quantum Theories.

In §2 we have considered various classical systems all of which have a depth logic which is either Boolean or a disjoint collection of Boolean logics. Thus, for the most part, classical systems are characterized by having all of their (theoretical) observables compatible, i.e., there is a single (theoretical) perspective from which one can see all there is to be seen. In order to obtain a truly non-classical theory, one must assume that the deep logic is non-Boolean and that its Boolean subalgebras interlock in a complex fashion. That quantum mechanics contains incompatible observables, such as the position and momentum observables,  $Q$  and  $P$ , was first realized by Dirac and Heisenberg in late 1926. This observation was not forced upon them by experimental failures to measure simultaneously the exact position and momentum of a particle; in fact, Planck's constant  $h$  is so small and our technology so weak, that even if an electron were a small classical billiard ball, no existing experiment could test Heisenberg's conclusion in his  $\gamma$ -ray microscope gedenken experiment. Instead, they came to their conclusion from the realization that the highly successful quantum mechanical formalism that had recently been established did not enable one to predict a joint probability distribution for  $Q$  and  $P$ . Thus, either the quantum mechanical algorithm, while very successful for some purposes, was statistically incomplete in that it did not enable one to predict the joint probability distributions for  $Q$  and  $P$ , or else the quantum mechanical

algorithm was statistically complete and hence  $Q$  and  $P$  didn't have a joint probability distribution, i.e., they were both empirically and theoretically incompatible. Thus, if one takes the quantum mechanical formalism very seriously and assumes that it is at least statistically complete, then one is forced to the conclusion (or discovery) that quantum mechanics contains incompatible observables. This approach can be summarized by Einstein's remark to Heisenberg: "It is the theory which decides what we can observe." Heisenberg's  $\gamma$ -ray microscope gedanken experiment was an attempt to retroactively rationalize and justify this viewpoint. He took Einstein's contribution to relativity theory as his paradigm. In Einstein's case also the formalism had already been established; and Poincaré at least was clearly aware that it called for a radical break with classical (Galilean) mechanics. But Einstein's approach via gedanken experiments involving rods and clocks was much more vivid and convincing. Thus, Einstein's approach was a way to rationalize the disappearance of an absolute notion of simultaneity, a disappearance which was already implied by taking the Lorentz group seriously, and taking the Lorentz group seriously was already implicit in the success and structure of Maxwell's theory. Of course, one also had the Michelson-Morley experiment. In his gedanken experiment, Heisenberg essentially shows that if  $Q_\epsilon$  is a position experiment with accuracy  $\epsilon$ , and  $P_\delta$  is a momentum experiment with accuracy  $\delta$ , then a necessary condition that  $Q_\epsilon$  and  $P_\delta$  be compatible is that  $\epsilon \cdot \delta \geq h$ . Note that Heisenberg's analysis here involves only the "old" quantum mechanics--in particular, wave-particle duality

and the Einstein-de Broglie relations. This inequality is totally unrelated to the similar-appearing inequality concerning statistical scatter relations which is easily derivable using the formalism of the "new" quantum mechanics. Heisenberg's gedanken experiment was as successful as Einstein's in providing a lucid and convincing rationalization of the established formalism. (Unfortunately for Heisenberg, Einstein wasn't particularly appreciative of the compliment of having his (former) approach followed. He is reputed to have remarked that "a good joke shouldn't be repeated too often".)

### §3.1 The Logic of Quantum Mechanics.

Upon leaving the classical (Boolean) tradition, one enters a chaos of possible deep logics. First, one investigates the simpler possibilities. Probably the simplest (and definitely the most important) non-Boolean deep logics are those obtained by assuming that  $\mathcal{L}^d$  is isomorphic to the diagram of Boolean sub-algebras of the orthocomplemented lattice of orthogonal projections on a Hilbert space. We will shortly see that this assumption "works", i.e., is capable of yielding the standard results of traditional quantum mechanics. But one can go even further and show how certain reasonable assumptions on a certain class of experimental situations actually yields deep logics which can be shown to be embeddable in  $\text{Proj}(H)$ . This has been done, for example, by Sneed [67]. Sneed main goal was to show that classical probability theory is completely adequate for describing certain quantum mechanical phenomena. He considers "finite filter experiments" which are an idealization of standard polarization experiments. He then shows how classical probability theory is capable of formulating all the questions one might have about such a system. Using certain "reasonable" axioms concerning the filters, he is able to show how one is led to introduce a certain orthomodular orthocomplemented lattice  $L$  and that the probabilities that one is interested in determine probability-like measures on  $L$ . Furthermore,  $L$  is embeddable in  $\text{Proj}(H)$ , and every probability measure on  $L$  is induced by one on  $\text{Proj}(H)$ . Thus, the standard quantum mechanics based upon  $\text{Proj}(H)$  is computationally adequate for describing any finite filter experiment. A similar

viewpoint is implicit in The Feynman Lectures on Physics and in work of Ludwig. On the other hand, we don't want to provide too strong a justification for the axiom that  $\mathcal{L}^d \cong \text{Proj}(H)$ , since it is known that other assumptions are appropriate in quantum statistical mechanics and in quantum field theory. In these subjects it is often reasonable to assume that  $\mathcal{L}^d \cong \text{Proj}(\mathcal{A})$ , where  $\mathcal{A}$  is a von Neumann algebra (or more generally a  $\Sigma^*$ -algebra (see [19])). Alternatively, if one works with the partial algebra  $\mathcal{A}_c$  of continuous real-valued observables, then one is led to assume that  $\mathcal{A}_c \cong \text{Herm}(\mathcal{A})$ , where  $\mathcal{A}$  is a  $C^*$ -algebra and  $\text{Herm}(\mathcal{A})$  is its partial algebra of Hermitian elements. (A natural question that arises here is whether  $\mathcal{A}$  is determined by  $\text{Proj}(\mathcal{A})$  or  $\text{Herm}(\mathcal{A})$ , i.e., if  $\text{Herm}(\mathcal{A}_1) \cong \text{Herm}(\mathcal{A}_2)$  (as partial algebras), then is  $\mathcal{A}_1 \cong \mathcal{A}_2$ ? Note that  $\text{Herm}(\mathcal{A}_1) \cong \text{Herm}(\mathcal{A}_2)$  does not imply that  $\mathcal{A}_1 \cong \mathcal{A}_2$  if  $\mathcal{A}_1$  and  $\mathcal{A}_2$  are only assumed to be real  $C^*$ -algebras: consider the complex numbers  $C$  and the quaternions  $H$  as real  $C^*$ -algebras, then  $\text{Herm}(C) \cong \text{Herm}(H) \cong R!$ ) Furthermore, it is possible that some of the difficulties occurring in quantum field theory in four dimensional space-time concerning the definability of commutators will require for their resolution the assumption that such commutators are not definable, i.e., that  $\mathcal{A}_c$  is an "exceptional" partial algebra not isomorphic to  $\text{Herm}(\mathcal{A})$  for any  $C^*$ -algebra  $\mathcal{A}$ !

Suppose now that  $\mathcal{L}^d \cong \text{Proj}(H)$ . Since a real-valued observable on  $\mathcal{L}^d$  is defined to be a  $\sigma$ -homomorphism from the Borel algebra of  $R$  to  $\mathcal{L}^d$ , it follows from von Neumann's spectral theorem that such observables are in 1-1 correspondence with the hypermaximal Hermitian operators on  $H$ . Furthermore, the partial algebra  $\mathcal{A}_m$  of bounded measurable

observables associated to  $\mathcal{L}^d$  is isomorphic to the partial algebra  $\text{Herm}(B(H))$  of bounded Hermitian operators on  $H$ . Similar results clearly hold for  $\mathcal{L}^d \cong \text{Proj}(\mathcal{A})$ , where  $\mathcal{A}$  is a von Neumann algebra.

The next problem is to describe the probability measures on  $\mathcal{L}^d$ .

Clearly, every unit vector  $\psi \in H$  determines a probability measure  $\mu_\psi$  on  $\mathcal{L}^d$  by the formula  $\mu_\psi(P) = \langle P\psi, \psi \rangle$ .  $\mu_\psi$  is called a vector state on  $\mathcal{L}^d$ . One can also form countable convex sums of vector states to obtain more general states  $\mu = \sum_i \alpha_i \mu_{\psi_i}$ . Such a  $\mu$  can alternatively be defined by the formula  $\mu_W(P) = \text{Tr}(PW)$ , where  $W$  is a von Neumann density matrix (i.e.,  $W$  is a positive self-adjoint operator of trace 1). One naturally asks if these are all the probability measures on  $\mathcal{L}^d$ . For  $H = \mathbb{C}^2$  the answer is a definitive no! One can embed the logic  $\text{Proj}(\mathbb{C}^2)$  into a Boolean algebra and thereby obtain lots of dispersion free probability measures on  $\text{Proj}(\mathbb{C}^2)$ . Thus, for  $H = \mathbb{C}^2$  one can contradict the Heisenberg uncertainty relations and also construct classical "hidden variables" models which are strict reductions (see [2] and [41]). In 1957 Gleason showed that for  $\dim H \geq 3$  all probability measures on  $\text{Proj}(H)$  can be represented by von Neumann density matrices.

That  $\text{Proj}(\mathbb{C}^2)$  is not "really" a quantum mechanical logic should come as no surprise since its diagram of Boolean subalgebras is an (uncountable) disjoint collection of copies of the Boolean algebra  $(0 \rightarrow P \rightarrow 1)$  with all the subalgebras  $(0 \rightarrow 1)$  identified. One needs more "coherence" in  $\mathcal{L}^d$  before one obtains truly non-classical phenomena. If  $\mathcal{L}^d \cong \text{Proj}(\mathcal{A})$ , where  $\mathcal{A}$  is a von Neumann algebra with discrete center, then  $\mathcal{A}$  decomposes into a discrete direct sum of factors,  $\mathcal{A} = \bigoplus \mathcal{A}_n$ , and

any pure state on  $\mathcal{Q}$  is determined by a pure state on one of the  $\mathcal{Q}_n$ . If  $\mathcal{Q}_n \cong B(H_n)$ , then every pure state on  $\mathcal{Q}$  is determined by a vector in  $H = \bigoplus H_n$  which lies in one of the  $H_n$ 's. Under these circumstances, physicists say that there exist super selection rules. If the center of  $\mathcal{Q}$  is continuous, then there are no pure states on  $\text{Proj}(\mathcal{Q})$ . This is a very embarrassing circumstance from the traditional point of view which considers only pure states as representing individuals and maximum knowledge. It provides no embarrassment - and even a strong argument in favor of - our point of view where any probability measure on  $\mathcal{L}^d$  might represent total possible knowledge of individuals.

We haven't even mentioned yet what was for Dirac the essence of quantum mechanics: namely, the principle of superposition of states. Since, by Gleason's theorem, the pure states on  $\text{Proj}(H)$  are determined by elements of  $H$ , if we are given a family  $\{\mu_{\Psi_i}\}$  of pure states on  $\text{Proj}(H)$ , then we can form new pure state  $\mu_\Psi$  with  $\Psi = \sum c_i \Psi_i \in H$ .  $\mu_\Psi$  is called a quantum mechanical superposition of the  $\mu_{\Psi_i}$ . This process is different from the process of forming classical superpositions such as  $\mu = \sum \alpha_i \mu_{\Psi_i}$ . In general, if  $\mathcal{B} = \{\mu_\alpha\}$  is a collection of states on  $\mathcal{L}^d$ , then we will (following Varadarajan [81]) say that  $\mu$  is a superposition of the states in  $\mathcal{B}$  if  $\mu_\alpha(P) = 0 \forall \mu_\alpha \in \mathcal{B} \rightarrow \mu(P) = 0$ . In the classical situation  $\mathcal{L}^d \simeq B(X)$ , the pure states correspond to probability measures  $\delta_a$  concentrated at points  $a \in X$  (assuming  $X$  is "reasonable"), all probability measures are superpositions of the pure ones  $\{\delta_a\}$ , but  $\delta_a$  is not a superposition of  $\{\delta_{a'}, |a' \neq a\}$ . In the quantum mechanical case, if  $\Psi = \sum c_i \Psi_i$ , then  $\mu_\Psi$  is a superposition of  $\{\mu_{\Psi_i}\}$ . So, quantum mechanics is distinguished from classical mechanics in the fact that even pure states can be superpositions of each other.

Our next task is to determine the automorphism group of  $\mathcal{L}^d$ . We first assume that  $\mathcal{L}^d \simeq \text{Proj}(H)$ , where  $H$  is a complex separable Hilbert space of dimension  $\geq 3$  (similar results hold over the reals and quaternions). Then it can be shown that every automorphism of  $\text{Proj}(H)$  is induced by a unitary or anti-unitary transformation of  $H$ . Automorphisms of  $\mathcal{L}^d$  induce convexity-preserving automorphisms of  $\text{Prob}(\mathcal{L}^d)$ . Conversely, every convexity-preserving automorphism of  $\text{Prob}(\mathcal{L}^d)$  is induced by an automorphism of  $\mathcal{L}^d$ . But we remind the reader that from our

interpretation of probability, there is no physical argument in favor of assuming that the maps of  $\text{Prob}(\mathcal{L}^d)$  that we are interested in are convexity-preserving unless we can give physical motivation for their being induced by maps of  $\mathcal{A}_m$ .

We expect the following general phenomena to hold for the cases  $\mathcal{L}^d \cong \text{Proj}(\hat{\mathcal{A}})$ , where  $\hat{\mathcal{A}}$  is a von Neumann algebra, and  $\mathcal{A}_c \cong \text{Herm}(\mathcal{A})$ , where  $\mathcal{A}$  is a  $C^*$ -algebra. First

from the partial states on  $\mathcal{A}_c$ ,  $S(\mathcal{A}_c)$ , to the probability measures on  $\text{Proj}(\hat{\mathcal{A}})$ , where  $\hat{\mathcal{A}}$  is the universal enveloping von Neumann algebra of  $\mathcal{A}$ . If  $\hat{\mathcal{A}}$  doesn't contain  $I_2$  as a direct summand, then this would show that every partial state on  $\mathcal{A}_c$  is actually a state, i.e., linear even across non-commuting elements, if every probability measure on such a von Neumann algebra determined a state. Second,

$$\text{Prob}(\text{Proj}(\hat{\mathcal{A}})) \rightarrow \text{Prob}(\text{Proj}(\hat{B}))$$

is induced by a positive partial linear map  $f: \hat{B} \rightarrow \hat{\mathcal{A}}$ , and, furthermore, if  $f^*$  is an isomorphism, then  $f$  is a partial isomorphism, and hence induces an isomorphism

$$f: \text{Proj}(B) \rightarrow \text{Proj}(\hat{\mathcal{A}}).$$

If  $\hat{\mathcal{A}}$  is a factor (not  $I_2$ ), then every partial homomorphism of  $\hat{\mathcal{A}}$  is actually either a homomorphism or an anti-homomorphism. It would be interesting to know to what extent these hold for exceptional partial algebras.

### §3.2. Dynamics.

The dynamics of a system is usually described by assuming that we have an action of  $R$  on  $\text{Prob}(\mathcal{L})$  by convexity-preserving automorphisms. As was seen in §3.1, this is equivalent to having a one-parameter group of automorphisms of  $\mathcal{L}^d$ . For  $\mathcal{L}^d \simeq \text{Proj}(H)$ , one can show that every such group of automorphisms can be induced by a (strongly continuous) one-parameter group of unitary transformations of  $H$ ,  $U_t = e^{-iHt}$ , where  $H$  is an Hermitian operator called, by analogy with classical mechanics, the Hamiltonian of the system, and assumed (defined) to be the observable representing the energy of the system.  $H$  is determined up to an additive constant by the action of  $R$  on  $\text{Proj}(H)$ .

Differentiating yields the abstract form of the Schrodinger equation  $\frac{\partial \Psi}{\partial t} = -iH\Psi$ . In the more general situations of  $\mathcal{L}^d \simeq \text{Proj}(\hat{A})$  or  $\mathcal{Q}_c \simeq \text{Herm}(\mathcal{Q})$ , Kadison has shown that one-parameter groups  $\alpha_t$  of convexity-preserving automorphisms of  $\text{Prob}(\mathcal{L}^d)$  or  $\mathcal{S}(\mathcal{Q}_c)$  are still induced by one-parameter groups  $\alpha_t$  of convexity-preserving automorphisms of  $\mathcal{L}^d$  or  $\mathcal{Q}_c$  (Kadison also assumes ~~that all partial states are states~~ that all partial states are states). But in these more general cases, a one-parameter group of automorphisms of  $\hat{A}$  may not be induced by a one-parameter group of unitary operators in  $\hat{A}$ , i.e., there may be no observable "energy". On the other hand, if each automorphism  $\alpha_t$  of the one-parameter group is inner (i.e., induced by a unitary operator in  $\hat{A}$ ), then under physically reasonable hypotheses-the one-parameter group is generated by a one-parameter unitary group  $U_t = e^{-iHt}$ , and there is an observable energy  $H$  in  $\hat{A}$  (possibly unbounded).

Up to now our model of the experimental process is that of a box with a timer and buttons: the timer is set to any time  $t_1 \in R$  and a button  $A_1$  is pressed yielding a result  $r_{A_1} \in R_{A_1}$  and then the box blows up (see §1). But suppose instead that the box has many timers and that we could set them for various times  $t_i \in R$  and choose various buttons  $A_i$  such that we obtain a sequence of values  $r_{A_i} \in R_{A_i}$  before the box blows up. What are we to do? The logically most consistent approach is to begin afresh and consider the sequence  $\{(t_i, A_i)\}$  as a single new experiment. We would thus be led to introduce entirely new surface and depth logics. But there is also a more conservative approach which often-but not always-suffices. Suppose that the state of the system is initially  $\mu \in \text{Prob}(\mathcal{L}^d)$ . Then at time  $t_1$  the state will be  $\alpha_{t_1}(\mu)$ . After  $A_1$  is pressed and  $r_{A_1}$  is recorded, it may be reasonable to assume that the system has been thrown into a new state which depends upon the old state  $\alpha_{t_1}(\mu)$  and the result  $r_{A_1}$ , the dependence being either deterministic, stochastic, or just restrictive.

After the experiment is completed, the new state evolves as if nothing had happened, i.e. via  $\alpha_t$ . Let us try to clarify the situation with some classical examples. A coin can usually be tossed repeatedly, but one doesn't usually introduce a complex depth logic which would incorporate the potential interdependencies. Instead, one usually assumes that the process of tossing the coin has a negligible effect on its intrinsic bias. Thus, one usually assumes that the new state of the coin is equal to its old state. Suppose, instead, that after being tossed a coin is totally biased

in favor of the result obtained, and that its bias evolves in time according to

$$\frac{d\mu}{dt} = \begin{cases} (\frac{1}{2} - \mu) & \text{if } \mu \leq \frac{1}{2} \\ - & \\ (1 - 2\mu) & \text{if } \mu \geq \frac{1}{2} \end{cases},$$

where  $\mu$  is the probability of a heads. Such a coin, if left to itself, seeks to become unbiased. Note, also, that the time semi-group is not acting in a convexity-preserving way on the states. If a box filled with such coins is shaken after a long rest, then one would expect approximately equal numbers of heads and tails to result. But if by chance there is a preponderance of heads, then upon an immediate retrial one would expect the preponderance to be sustained. One could generalize this example by supposing that the bias has only a probability of changing after a toss.

Now let us return to quantum mechanics. Let  $A$  be an Hermitian operator with discrete non-degenerate eigenvalues  $\{\alpha_i\}$  and associated eigen-vectors  $\{\Psi_i\}$ . Then a standard assumption in quantum mechanics is that no matter what the original state  $\mu$  is, the state after  $A$  is measured and  $\alpha_i$  is obtained will be  $\mu_{\Psi_i}$ . This assumption of "reduction of the wave packet" makes the quantum mechanics system behave in a manner analogous to our coins which become totally biased after having been forced into making a decision. If  $A$  has degenerate or continuous eigenvalues, then one only obtains the restriction that if the measured value of  $A$  lies in  $E \subset R$ , then the new state is  $\mu_\Psi$  for some  $\Psi \in P_E^H$ , where  $P_B$  is the spectral resolution of  $A$ . Many paradoxes

of quantum mechanics, such as those of Schrodinger's cat, Wigner's friend, and the E.P.R., can be traced to the hypothesis of reduction of the wave packet.

### §3.3. The Quantum Theory of Particles.

We will now consider the quantum mechanical notion of a particle. Suppose we have a system which has a position observable  $Q$  which yields a result in  $R^3$  and that the system is invariant with respect to an action of the Euclidean group  $E_3$  (= the group of transformations of  $R^3$  generated by translations and rotations), i.e., we are given actions of  $E_3$  on  $\text{Lab}$ ,  $R^3$ ,  $\mathcal{L}^d$ , such that the maps  $\mathcal{B}(R^3) \rightarrow \mathcal{L}^d$  and  $\text{Lab} \rightarrow \text{Prob}(\mathcal{L}^d)$  are equivariant, where the action on  $\text{Prob}(\mathcal{L}^d)$  is the natural one induced by the action of  $E_3$  on  $\mathcal{L}^d$ . In the classical case one usually takes  $\mathcal{L}^d \simeq \mathcal{B}(T^* R^3)$  with the action of  $E_3$  on  $T^* R^3$  induced by its natural action on  $R^3$  and the inclusion  $\mathcal{B}(R^3) \rightarrow \mathcal{B}(T^* R^3)$  induced by the natural projection map  $T^* R^3 \rightarrow R^3$ . In the case of the standard quantum logic  $\mathcal{L}^d \simeq \text{Proj}(H)$ , one can show that there is a natural 1-1 correspondence between isomorphism classes of equivariant maps  $(\mathcal{B}(R^3), E_3) \rightarrow (\text{Proj}(H), E_3)$  (called systems of imprimitivity for  $(\mathcal{B}(R^3), E_3)$ ) and isomorphism classes of (projective) representations of  $\text{SO}(3)$ . If  $V$  is a (projective) representation of  $\text{SO}(3)$ , then the system of imprimitivity corresponding to it has a natural representation in  $L^2(R^3, V) \simeq L^2(R^3) \otimes V$ . Usually one pays special attention to the "irreducible" systems of imprimitivity, i.e., those for which  $B(H)$  is the smallest von Neumann algebra containing the projections  $P_E$  for  $E \in \mathcal{B}(R^3)$  (i.e., the projections occurring in the spectral resolutions of the position observables) and also the projections occurring in the spectral resolutions of the momentum observables (i.e., the infinitesimal generators of the unitary

one-parameter groups inducing the one-parameter groups of automorphisms of  $\text{Proj}(H)$  which are determined by the one-parameter sub-groups of  $E_3$ ). In other words, if  $\zeta: (\mathcal{B}(R^3), E_3) \rightarrow (\text{Proj}(H), E_3)$  is an irreducible system of imprimitivity, then every observable in  $\text{Herm}(B(H))$  is a "function" of the position and momentum observables. (Note that this notion of function is purely mathematical and not physical; it depends upon the full algebra (and  $W^*$ ) structure of  $B(H)$  and not only its (more physical) partial algebra structure.) Given  $\zeta_i: (\mathcal{B}(R^3), E_3) \rightarrow (\text{Proj}(H_i), E_3)$  ( $i = 1, 2$ ), then we can form

$$\zeta = \zeta_1 \oplus \zeta_2: (\mathcal{B}(R^3), E_3) \rightarrow (\text{Proj}(H_1 \oplus H_2), E_3)$$

as follows. For  $g \in E_3$  let  $U_i(g)$  be a unitary operator on  $H_i$  inducing the action of  $g$  on  $\text{Proj}(H_i)$ . (Since  $E_3$  is a connected Lie group, every element of it is a square, and, hence, when  $E_3$  acts as an automorphism group of  $\text{Proj}(H)$ , each such automorphism is induced by a unitary operator on  $H$ .) Let  $U(g) = U_1(g) \oplus U_2(g)$  be the direct sum unitary operator on  $H = H_1 \oplus H_2$ , and let  $E_3$  act on  $\text{Proj}(H)$  via the induced action determined by  $U(g)$ . For  $E \in \mathcal{B}(R^3)$ , let  $\zeta(E) = \zeta_1(E) \oplus \zeta_2(E)$ . Then  $\zeta$  is equivariant. Given an arbitrary family  $\{\zeta_i\}$ , one can similarly define  $\oplus \zeta_i$ . Note that this construction would not work in general for the unrestricted Euclidean group which contains the reflections and is thus not connected because if  $r$  is the reflection about the  $x$ -axis and if  $\zeta_1(r)$  is induced by a unitary operator  $U_1(r)$  and  $\zeta_2(r)$  is induced by an anti-unitary operator  $U_2(r)$ , then  $U_1(r) \oplus U_2(r)$  is neither unitary nor anti-unitary, and hence does not determine an automorphism of  $\text{Proj}(H)$ ! Moreover, even for  $E_1 = R^1$  the construction is not well defined

on isomorphism classes of representations since  $\zeta_1(t) = e^{-iHt}$  and  $\zeta_2(t) = e^{iHt}$  are anti-unitarily equivalent actions of  $R^t$  on  $H_i$ , and hence are isomorphic actions on  $\text{Proj}(H_i)$ , but  $\zeta_1 \oplus \zeta_2$  is not isomorphic to  $\zeta_1 \oplus \zeta_1$ ! The importance of the direct sum construction lies in the fact that every system of imprimitivity for  $(\mathcal{B}(R^3), E_3)$  is isomorphic to a direct sum of irreducible such systems. The irreducible systems of imprimitivity for  $(\mathcal{B}(R^3), E_3)$  correspond to the irreducible (projective) representations of  $SO(3)$  and these are standardly denoted by  $D_j$ , where  $j$  is a half integer. The dimension of  $D_j$  is  $2j+1$  and the representation is an ordinary one if and only if  $j$  is an integer. In any case, associated to any system of imprimitivity for  $(\mathcal{B}(R^3), E_3)$  (in a complex Hilbert space  $H$ ) we have the special observables of position  $Q_i$ , momentum  $P_i$ , and angular momentum  $J_i$  (these observables are associated to (unbounded) Hermitian operators on  $H$ ). One can now derive the Heisenberg commutation relations  $[Q_i, P_i] = -\sqrt{-1}$ , and hence the Heisenberg uncertainty relations  $\Delta Q_i \cdot \Delta P_i \geq 1/2$ . Note that these uncertainty relations say nothing about one's ability to make simultaneous measurements of  $Q_i$  and  $P_i$ ; in fact, the non-vanishing of  $[Q_i, P_i]$  implies that  $Q_i$  and  $P_i$  are incompatible. These uncertainty relations are purely statistical scatter relations governing the statistics of independent measurements of  $Q_i$  and  $P_i$ . The main theorem required for the above derivation is Gleason's theorem asserting that the probability measures on  $\text{Proj}(H)$  can all be represented by density matrices. Given this, one can easily show (using the Schwartz inequality) that  $\Delta A \Delta B \geq 1/2 | \langle [A, B] \rangle |$ . To obtain the commutation relations  $[Q_i, P_i] = -\sqrt{-1}$ ,

one considers the special representation of  $(\mathcal{B}(\mathbb{R}^3), E_3) \rightarrow (\text{Proj}(H), E_3)$  in  $L^2(\mathbb{R}^3) \otimes V$ . In this representation,  $E \in \mathcal{B}(\mathbb{R}^3)$  goes to the projection operator corresponding to multiplication by its characteristic function  $x_E$ . Hence,  $Q_i$  is represented by multiplication by  $x_i$ .  $P_i$ , the infinitesimal generator of the one-parameter group of translations in the  $x_i$  direction, is represented by  $\frac{1}{\sqrt{-1}} \frac{\partial}{\partial x_i}$ . One thus obtains the commutation relations. The infinitesimal generator of the one-parameter group of rotations about the  $x_3$ -axis has the form

$$J_3 = L_3 + S_3 = \frac{1}{\sqrt{-1}} (x_2 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_2}) + S_3,$$

where  $S_3$  is an Hermitian operator on  $V$ .  $L_3$  is called the orbital and  $S_3$  the intrinsic spin angular momentum of the particle. A natural question to ask at this point is whether or not there is a classical analog of the intrinsic spin of a particle. That the answer is yes is quite recent! If one thinks of a particle as a small billiard ball, then it is natural to assume that the particle can rotate. This leads one to replace the configuration space  $\mathbb{R}^3$  by the configuration space  $\mathbb{R}^3 \times SO(3)$ .

Quantizing yields a natural action of  $E_3$  in  $L^2(\mathbb{R}^3 \times SO(3)) \cong L^2(\mathbb{R}^3) \otimes L^2(SO(3))$ . Here one has, besides the basic observables  $Q_i$ ,  $P_i$ ,  $L_i$ , and  $S_i$ , also new observables detecting the orientation of the particle. But not every system of imprimitivity for  $(\mathcal{B}(\mathbb{R}^3), E_3)$  is of this form. In particular, the system corresponding to the (projective) representation  $D_j$  of  $SO(3)$  is not of this form. In  $L^2(\mathbb{R}^3) \otimes D_j$  we have the extra angular momentum observables  $S_j$ , but we do not have any extra orientation observables. Classically this corresponds to replacing the phase space  $T^* \mathbb{R}^3$  by  $T^* \mathbb{R}^3 \times S^2$ , where  $E_3$  acts on  $S^2$  via the natural (symplectic) action of  $SO(3)$ . The closed two-form  $\omega$  on  $S^2$  defining the symplectic

structure also defines an area element. Let  $s = \int_{S^2} \omega$ . If  $s$  is an integer, then one can "quantize"  $T^* R^3 \times S^2$  and obtain a system of imprimitivity for  $(\mathcal{B}(R^3), E_3)$  which corresponds to  $D_{s/2}$ . Thus, the classical analog of the quantum mechanical system corresponding to  $D_j$  is not an extended billiard ball type of particle, but, instead, a point particle having extra pure angular momentum degrees of freedom in  $S^2$ .  $T^* R^3$  and  $T^* R^3 \times S^2$  are the only Euclidean invariant symplectic manifolds over  $R^3$  whose points are separated by the functions corresponding to the basic observables  $Q_i, P_i, J_i$ . One thus refers to these systems as "elementary" (symplectic) systems of imprimitivity for  $(\mathcal{B}(R^3), E_3)$ .  $S^2$  is an elementary system for  $SO(3)$ , and one can also obtain elementary systems for  $E_3$ . By and large, we feel that too much emphasis has been placed upon the "elementary" systems and not enough upon the composite or reducible systems (such as  $T^*(R^3 \times SO(3)) \rightarrow R^3$ ).

There has been much recent work in extending the above analysis to cases where  $(R_3, E_3)$  is replaced by some other pair  $(M, G)$ , and even further to cases where one is only given the classical phase space  $P$  (assumed to be a symplectic manifold). In a different direction, if one assumes that  $\mathcal{L}^d \simeq \text{Proj}(\mathcal{B})$ , where  $\mathcal{B}$  is a von Neumann algebra, and that  $E_3$  acts via inner automorphisms, then von Neumann's theorem concerning the essential uniqueness of the representations of the Heisenberg commutation relations implies that  $\mathcal{B}$  contains a direct sum of  $I_\infty$ 's and that on each summand we have a standard Schrodinger representation. We don't know whether or not there exist other systems of imprimitivity for  $(\mathcal{B}(R^3), E_3)$  in  $\text{Proj}(\mathcal{B})$ , with  $E_3$  not acting by

inner automorphisms.

The main point, so far, is that the physicist's standard quantum kinematical description of a particle - including the Heisenberg commutation relations and uncertain principle, the Schrodinger representation in terms of wave functions on  $\mathbb{R}^3$ , the notion of the intrinsic spin of a particle, and the superposition principle - follows from the single assumption that  $\mathcal{L}^d \cong \text{Proj}(H)$  together with standard symmetry assumptions! We shall now go further and assume that the dynamics of our system is compatible with its kinematics, i.e., that the action of the one-parameter time group  $T$  on  $\mathcal{L}^d$  commutes with the action of  $E_3$  on  $\mathcal{L}^d$ . If our system of imprimitivity is the one corresponding to  $D_0$  and represented in  $\mathcal{L}^d(\mathbb{R}^3)$ , then, by using Fourier transforms, one can show that the Hamiltonian  $H = g(P_1^2 + P_2^2 + P_3^2) = g(-\nabla^2)$ , where  $g$  is some Borel function on  $\mathbb{R}^+$ . If one assumes that  $g$  is smooth and approximates it by its first order Taylor expansion, then one has  $H \sim g(0) - g'(0)\nabla^2 = v_0 - \frac{1}{2m} \left( \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \right)$ . Hence, to first order, one obtains the usual Schrodinger equation

$$\frac{\partial \psi}{\partial t} = -iH\psi = \frac{i}{2m} \nabla^2 \psi - iv_0 \psi .$$

For the case of higher spin  $D_j$  one can show - using the theory of group representations - that  $H$  is determined by  $2j+1$  real-valued functions on  $\mathbb{R}^+$ .

What we have, in effect, assumed is that the action of  $E_3$  on  $\mathcal{L}^d$  extends to an action of the Newtonian space-time group  $E_3 \times T$  on  $\mathcal{L}^d$ . One can go further and assume that the action of  $E_3$  extends to an

action of the Galilean or Lorentz group. In the case of the Galilean group, one can show that  $H = -\frac{1}{2m} \nabla^2$ . For the Lorentz group, one has various possibilities, but, in any case, one always has the relation  $H^2 = m_0^2 c^4 - c^2 \nabla^2$ . More generally, one considers (projective) representations of a space-time group  $G$ , whether or not they extend a system of imprimitivity for  $(\mathcal{B}(R^3), E_3)$ . Irreducible such representations are often said to describe "elementary particles". But the term elementary systems is often more appropriate since these systems often lack many of the properties that we usually associate with the word particle. For example, the representations of the Lorentz group which are thought to describe photons, neutrinos, and gravitons cannot be localized, i.e., they do not extend a system of imprimitivity for  $(\mathcal{B}(R^3), E_3)$  (nor even one for  $(\mathcal{B}(R^4), G)$ ). The corresponding classical phase space for these systems is  $R^4 \times S^2$ ; we thus have energy, momentum, and angular momentum observables for these systems, but no location observables. Other representations of the Lorentz group are often said not to be of physical significance since they imply that the energy may be negative, or always zero, or that the mass parameter may be imaginary. We feel that these are inadequate grounds for eliminating the consideration of these representations; in fact, we shall shortly see that many of these representations occur in the decomposition of certain natural occurring reducible representations of  $G$ .

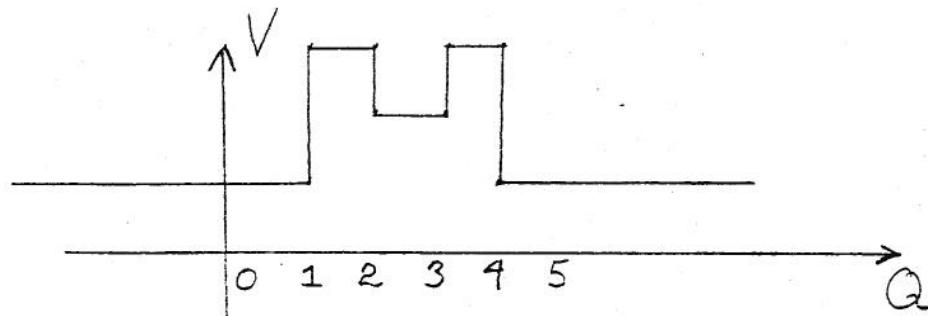
So far we have considered physical systems which are described using systems of imprimitivity for  $(\mathcal{B}(\mathbb{R}^3), E_3)$  or representations of the (inhomogeneous) Lorentz group. One can also consider systems whose description requires a system of imprimitivity for the Lorentz group. Suppose we have an experimental setup which yields an "event" in  $\mathbb{R}^4$  and admits the Lorentz group  $L$  as a symmetry group. Such a system will be described by a system of imprimitivity for  $(\mathcal{B}(\mathbb{R}^4), L)$ . Systems of imprimitivity for  $(\mathcal{B}(\mathbb{R}^4), L)$  are classified by the (projective) representations of the homogeneous Lorentz group  $L$ . One will now have a time observable  $T$  along with the position observables  $Q_i$ . Thus, one will also have the commutation relation  $[T, H] = i$ , and hence the (much used) uncertainty principle  $\Delta T \cdot \Delta H \geq 1/2$ . The resulting representation of  $L$  will not be irreducible. Furthermore, representations usually ignored in particle physics will occur in its decomposition as a direct integral of irreducibles. It is quite possible that phenomena such as resonances might be better thought of as "events" rather than "particles". Even the electron is not completely particle-like. At any time we choose, we will find "it" somewhere; but there are no paths. One can construct a quantum mechanical theory of particles with paths. Let  $C$  be the six dimensional homogeneous space of time-like lines in  $\mathbb{R}^4$ .  $L$  acts transitively on  $C$ . Systems of imprimitivity for  $(\mathcal{B}(C), L)$  have the most right to be said to be describing "particles". Such systems are classified by (projective representations of  $SO(3) \times \mathbb{R}$ ). Of course, the associated representation of  $L$  is not irreducible. Other possibilities readily come to mind -- in particular, theories based upon time-like, light-like or space-like

lines or hyperplanes. These theories should be systematically investigated. In any case, they show that one can have a quantum mechanical theory of particles with continuous paths.

### §3.4. Examples.

In this section we shall show how our formalism works on some standard examples.

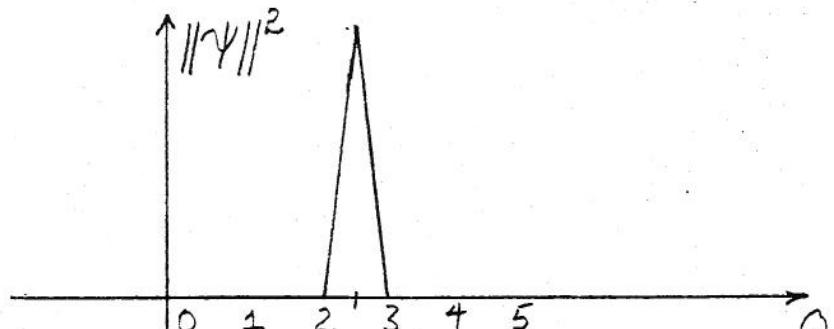
Example 1. Consider first a particle moving in a potential well described by the graph (3.4.1). What this means



(3.4.1)

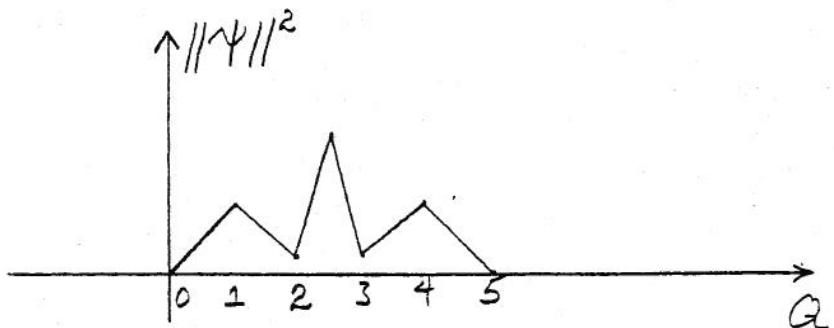
is that we have a basic location observable  $Q$  represented by the unbounded Hermitian multiplication operator  $M_Q$  acting on  $\mathcal{H} \cong L^2(\mathbb{R}^1, W)$ . Translations on  $\mathbb{R}^1$  generate a one-parameter unitary group on  $\mathcal{H}$  whose infinitesimal generator is  $\frac{1}{i} \frac{\partial}{\partial x} = P$ . One now assumes that the dynamics is generated by the Hamiltonian  $H = \frac{P^2}{2m} + M_V$ . Of course, one has to check that this sum defines an Hermitian operator on  $\mathcal{H}$ . Furthermore, one should note that since  $V$  is not translation invariant, the space and time translations will not commute; hence momentum may not be conserved in time. Suppose one has prepared the system so that if a position measurement is done at time  $T = 0$ , then the result recorded is always approximately 2.5. One usually describes this situation by saying that at time  $T = 0$  the particle was in the potential well. One might describe the particle's state by using a wave function as

graphed below.



(3.4.2)

Suppose that at time  $T = 1$  one sometimes finds the particle outside the well. Then its wave function might look like the following graph. (We aren't actually solving the wave equation - the points we are trying to make are qualitative and not quantitative).



(3.4.3)

One usually describes the above situation by saying that the particle which at time  $T = 0$  was inside the potential well has a non-zero probability of "tunneling" out. There are several problems with this (colorful) language from our point of view. First of all, if we do not make a position measurement at time  $T = 0$ , then  $Q$  has no value at time  $T = 0$ , even if the probability was 1 that had we made the measurement the value would have been 2.5! Secondly, if we do measure  $Q$  at  $T = 0$ , then we may destroy the system and not be able to make any further

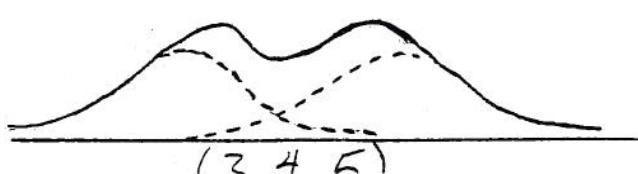
measurements. Thirdly, the language suggests that we are talking about a system whose basic observable is a continuous path, whereas the theoretical formalism being used doesn't contain any path questions. (This is also true of the usual Hamiltonian formulation of classical mechanics.) Strangely enough, the standard language of the physicist is often more compatible with some of the hidden variables models (such as those of Bohm and Nelson) than it is with the Copenhagen interpretation of quantum mechanics. (In spite of this, we have found physicists react very negatively when such hidden variables theories are mentioned.)

Example 2. We will now consider the classic two slit experiment. Suppose that at time  $T = 0$  we find the particle in region A with probability  $1/2$  or in region B with probability  $1/2$  (see diagram (3.4.4)); and that at time  $T = 1$



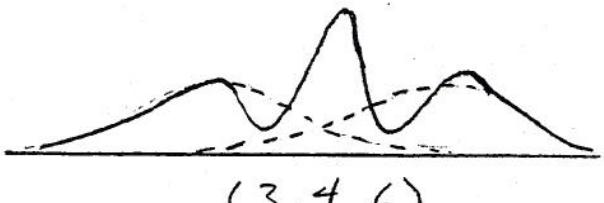
(3.4.4)

we find the particle in region C with probability 1. If we do a large ensemble of experiments at time  $T = 1$ , then what pattern will emerge? If our particles were billiard balls, then one would expect to see the solid line curve below emerge.



(3.4.5)

A similar experiment involving waves would yield an interference pattern such as figure (3.4.6).



(3.4.6)

If the experiment is done with electrons, then pattern (3.4.6) emerges even though at each repetition of the experiment a single location is observed. A classical model that would reproduce these results can be obtained by considering bubbles on water, the bubbles lasting only for a short time and their density being proportional to the wave intensity. Bohm and Nelson have other semi-classical models. If we assume that we are looking at a quantum mechanical particle, then at time  $T = 0$  the wave function  $\psi \in \mathcal{L}^2(\mathbb{R}^3, W)$  would have support in  $A \cup B$ . Let

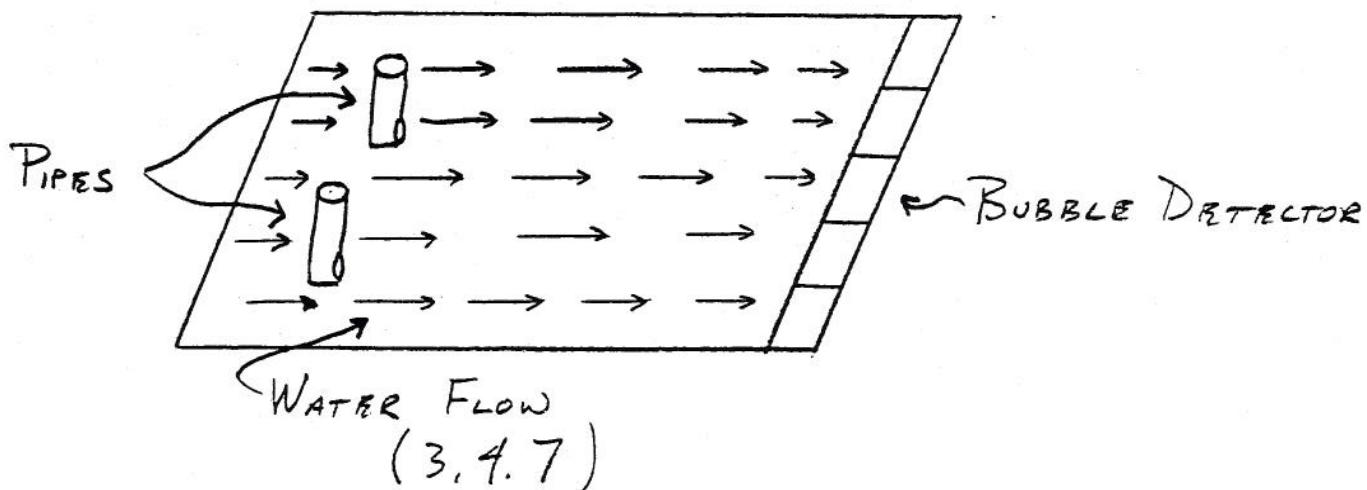
$$\psi_A(x) = \begin{cases} \psi(x) & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

and

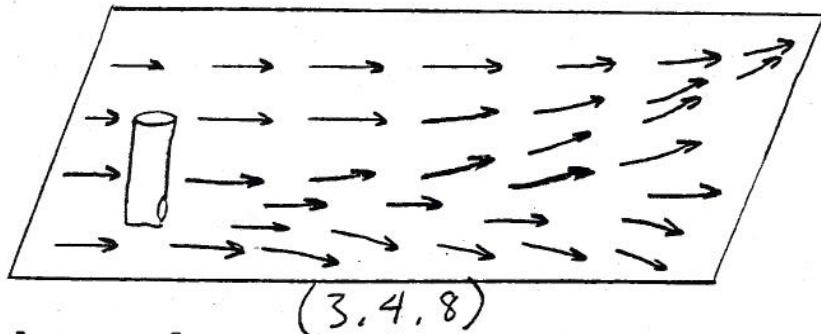
$$\psi_B(x) = \begin{cases} \psi(x) & \text{if } x \in B \\ 0 & \text{if } x \notin B \end{cases}$$

Then  $\psi = \psi_A + \psi_B$ , since  $A \cap B = \emptyset$ . This sum is a purely mathematical decomposition:  $\psi$  is a normalized state, but  $\psi_A$  and  $\psi_B$  are not normalized states. Of course,  $\mu_\psi$  is a "quantum mechanical" superposition of  $\mu_{\psi_A}$  and  $\mu_{\psi_B}$  in the sense described previously. But here we are focusing upon  $\psi_A$  and  $\psi_B$  and not upon  $\mu_{\psi_A}$  and  $\mu_{\psi_B}$ . By Wigner's theorem, the time

evolution is represented by a one-parameter unitary group  $U_t$  acting upon  $\mathcal{L}^2(R^3, W)$ . Hence, if  $\mu_\psi$  is the state at time  $T = 0$ , then at time  $T = 1$  the new state will be  $\mu_{U_1\psi} = \mu_{U_1\psi_A}$ . Thus,  $|U_1\psi|^2 = |U_1\psi_A + U_1\psi_B|^2 = |U_1\psi_A|^2 + |U_1\psi_B|^2 + 2\text{Re}(U_1\psi_A, U_1\psi_B)$ . The last term is called an interference term because it is responsible for the resulting pattern of observations looking like figure (3.4.6) instead of (3.4.5). This derivation of the interference pattern rests upon the following: the assumption that  $\mathcal{L}^d \simeq \text{Proj}(\mathcal{V})$ ; Gleason's description of the probability measures on  $\text{Proj}(\mathcal{V})$ ; the Schrodinger representation of  $\mathcal{V}$ ; and Wigner's theorem that time acts linearly on  $\mathcal{V}$ . Furthermore, we represented the original state by the "pure" probability measure  $\mu_\psi$ . If, instead, we started with a mixed probability measure, say  $\frac{1}{2}\mu_{\psi_A} + \frac{1}{2}\mu_{\psi_B}$ , then pattern (3.4.5) would emerge. Physicists usually only consider pure states as describing individuals, but we have argued otherwise. It is often very enlightening to view these experiments from the point of view of the hidden variable theorists. One can visualize Nelson's model in the following way. Consider water flowing smoothly in the plane to the right and two metal pipes above the plane (see diagram below (3.4.7) below). The pipes can be lowered into

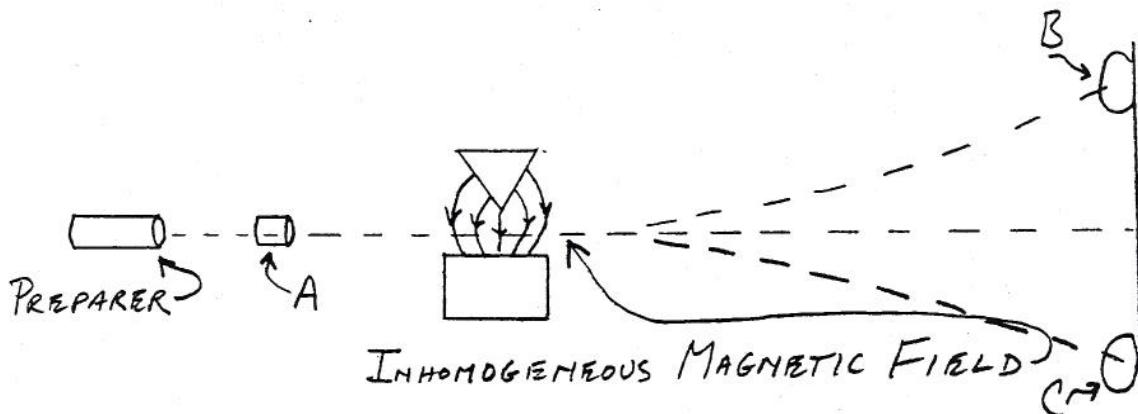


the water and plastic bubbles can be released through their holes. If one pipe is lowered, then one obtains the situation pictured in figure (3.4.8).



The mixture  $\frac{1}{2} \mu_{\psi_A} + \frac{1}{2} \mu_{\psi_B}$  is now physically described as the result obtained when the above system is prepared as follows: flip a coin to determine which pipe to lower, then lower that pipe and release one bubble from the pipe and then record its final location using the bubble detector (the bubbles are made of plastic to insure that they live long enough to reach the detector). The superposition  $\mu_{\psi} = \mu_{\psi_A} + \mu_{\psi_B}$  is obtained by first lowering both pipes into the water, and then flipping a coin to determine from which pipe the bubble should be released. In this model, as also in Bohm's model, the particles have continuous trajectories.

Example 3. We will now consider a model of the Stern-Gerlach experiment. A beam of silver atoms is sent through an inhomogeneous magnetic field and splits into two beams. Let  $Q$  denote the basic position observable. Then, at time  $T = 0$   $Q$  is in



(3.4.9)

region  $A$  with probability 1, and at time  $T = 2$   $Q$  is in either region  $B$  or  $C$ . Since the beam splits in two, it is reasonable to suppose that the state of the system is described by a two-component wave function

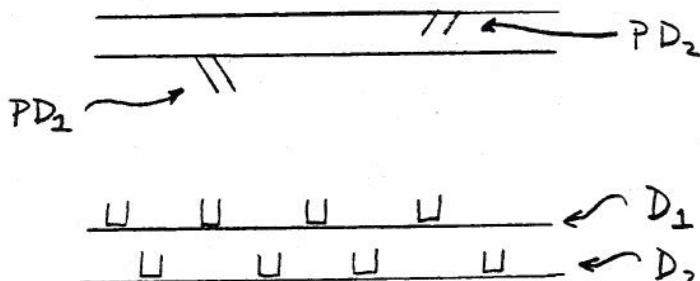
$\psi = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} \in \mathcal{L}^2(\mathbb{R}^3; D_{1/2})$ . Furthermore, at time  $T = 0$  one should obtain a reasonable approximation to  $\psi$  by considering it to be constant inside  $A$  and zero outside.

Suppose the system has been set up so that if  $\psi = \begin{pmatrix} \psi_+ \\ 0 \end{pmatrix}$  at  $T = 0$ , then  $Q$  will be in region  $B$  at  $T = 2$ , and if  $\psi = \begin{pmatrix} 0 \\ \psi_- \end{pmatrix}$  at  $T = 0$ , then  $Q$  will be in region  $C$  at  $T = 2$ . What happens if the preparer is rotated? The projective representation of  $SO(3)$  on  $D_{1/2}$  can be induced by a unitary representation of  $Spin(3) = SU(2)$ . Let  $R \in SO(3)$ , and let  $\tilde{R} \in Spin(3)$  such that  $\pi(\tilde{R}) = R$ , where  $\pi: Spin(3) \rightarrow SO(3)$  is the double covering homomorphism. Then, if the old wave function had value  $\psi$  in  $A$ , the new wave function will have value  $\tilde{R}\psi$  in  $A$ . Since  $\tilde{R}\psi = b\psi_+ + c\psi_-$  is a

superposition of  $\psi_+$  and  $\psi_-$ , and the evolution of the system is governed by a unitary one-parameter group, at  $T = 2Q$  will be in region B with probability  $|b|^2$ , and in region C with probability  $|c|^2$ . This prediction has been tested and is confirmed. One can similarly deduce the consequences of rotating the magnet. Note that nowhere in this discussion is it necessary (or possible) to picture a silver atom as a tiny spinning top.

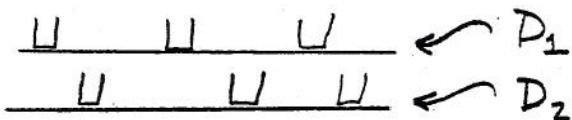
§3.5. The Quantum Theory of Interacting Particles.

In the previous sections we've considered the quantum theory of "individual isolated particles." But the world - as seen by the physicist - is composed of many particles undergoing complex interactions. Suppose, to begin with, that we consider the case of two "particles". What should the basic observables be? There are two natural choices. Clearly, the locations of the particles will yield a pair of points in  $\mathbb{R}^3$ . If our detecting devices can distinguish the two particles (for example, the particles might have different colors, or tastes, or charges), then our basic observable will yield a result in  $\mathbb{R}^6$ . If we cannot distinguish the particles, then our basic observable will yield a result in the space of (unordered) pairs of points in  $\mathbb{R}^3$ . In either case, it is natural to assume Euclidean invariance. In the first case one might go further and assume invariance under  $\mathfrak{S}_3 \times \mathfrak{S}_3$ . This larger symmetry group would be appropriate if we had two separate preparing devices and two separate arrays of detectors, one for each particle, all of which could be independently acted upon by  $\mathfrak{S}_3$  (see diagram below),



and we discovered that moving  $PD_1$  by  $g_1 \in \mathfrak{S}_3$  and  $PD_2$  by  $g_2 \in \mathfrak{S}_3$  had the same effect as moving  $D_1$  by  $g_1^{-1}$  and  $D_2$  by  $g_2^{-1}$ . This would certainly be the case for "non-interacting" particles. (But, of course, we wouldn't

call particles non-interacting unless this was the case!) In this case, one can show that the irreducible systems of imprimitivity for  $(R^3 \times R^3, \mathfrak{e}_3 \times \mathfrak{e}_3)$  all naturally decompose into a tensor product of two one-particle such systems. A similar result holds for a system of  $N$  particles based upon  $(R^3 \times R^3 \times \dots \times R^3, \mathfrak{e}_3 \times \mathfrak{e}_3 \times \dots \times \mathfrak{e}_3)$ . If we have only one preparing device (see below),



then one might discover that moving PD by  $g$  has the same effect as moving both detectors  $D_1$  and  $D_2$  by  $g^{-1}$ . In this case we are looking for systems of imprimitivity for  $(R^3 \times R^3, \mathfrak{e}_3)$ , where  $\mathfrak{e}_3$  acts on  $R^3 \times R^3$  via the diagonal map  $\mathfrak{e}_3 \rightarrow \mathfrak{e}_3 \times \mathfrak{e}_3$ . Here,  $\mathfrak{e}_3$  is not acting transitively on  $R^3 \times R^3$ . What do all the systems of imprimitivity for  $(R^3 \times R^3, \mathfrak{e}_3)$  look like? Those which come from a system of imprimitivity for  $(R^3 \times R^3, \mathfrak{e}_3 \times \mathfrak{e}_3)$  by restricting to the diagonal  $\mathfrak{e}_3 \rightarrow \mathfrak{e}_3 \times \mathfrak{e}_3$  can be shown to be determined by the Clebsch-Gordan series

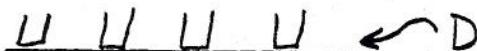
$$D_j \otimes D_k = D_{|j-k|} \oplus D_{|j-k|+1} \oplus \dots \oplus D_{j+k}.$$

There are other systems of imprimitivity for  $(R^3 \times R^3, \mathfrak{e}_3)$ , though they are not usually encountered in the physical literature! They describe systems with a basic configuration observable in  $R^6$ , but with only the momentum and angular momentum observables obtainable from the Lie algebra of  $\mathfrak{e}_3$ . One can decompose such systems of imprimitivity into "direct

"integrals" of irreducible such systems. A typical orbit of  $R^3 \times R^3$  under  $\delta_3$  is homeomorphic to  $R^3 \times S^2$ . A system of imprimitivity for  $(R^3 \times S^2, \delta_3)$  can be thought of as describing an "oriented rigid bar."

Possibly some binary molecules could be best described in this way.

We turn now to the case of two "identical particles", i.e., we have the experimental arrangement (below)



where we know that exactly two of the detectors in D will record a positive result. Assuming Euclidean invariance leads us to look for system of imprimitivity for  $(R^3 \times R^3/Z_2, \delta_3)$ . Two families of such systems are obtained by starting with a system of imprimitivity for  $(R^3, \delta_3)$  and forming either the symmetric or anti-symmetric tensor product of it with itself. These systems can be realized as spaces of symmetric or anti-symmetric  $L^2$ -vector-valued functions on  $R^3 \times R^3$ . There are other natural systems of imprimitivity for  $(R^3 \times R^3/Z_2, \delta_3)$ . A typical orbit of  $\delta_3$  acting on  $R^3 \times R^3/Z_2$  is homeomorphic to  $R^3 \times RP^2$ . A system of imprimitivity for  $(R^3 \times RP^2, \delta_3)$  can be thought of as describing an "unoriented rigid bar". Such a system may be appropriate for describing molecules composed of two identical atoms, e.g. the hydrogen molecule  $H_2$ .

The above considerations easily extend to the case of  $N(< \infty)$  particles. What about the case of infinitely many particles? (Even if one doesn't believe that any "real" system has infinitely many particles one could still be interested in exploring such models. For example,

they may provide insights into the behavior of very large systems ( $N \gg 0$ ) while being computationally simpler (recall the relationship between the binomial and normal distributions.) Suppose that we have a preparing device and a detector and that we expect to record positions in a region which occur with a density proportional to the region's volume. Such situations occur frequently, and a theory of random sets has been developed (see [53]). Let  $C(\mathbb{R}^3)$  denote the configuration space of infinite, locally finite subsets of  $\mathbb{R}^3$ .  $\delta_3$  acts naturally on  $C(\mathbb{R}^3)$ . What are all the possible systems of imprimitivity for  $(C(\mathbb{R}^3), \delta_3)$ ? (Work of Mackey and Glimm shows that there is little hope of actually being able to "survey" all possible such systems; the best we can expect to do is to explicitly construct large families of such systems!) There are various possible topologies for  $C(\mathbb{R}^3)$ ; besides the one usually used in the theory of random sets [53], one also has an interesting topology which arises in geometric measure theory [20]. In this second topology, a small neighborhood of the "vacuum" contains configurations which are very complicated; the vacuum in this theory is thus more like a Jackson Pollack painting than an empty canvass! If we use the topology described in [53], then  $\mu = \beta dx$ , where  $dx$  is Lebesgue measure on  $\mathbb{R}^3$ , induces an associated "Poisson" measure  $\mu_\beta$  on  $C(\mathbb{R}^3)$ .  $\delta_3$  acts ergodically on  $(C(\mathbb{R}^3), \mu_\beta)$ , and hence we can obtain a system of imprimitivity for  $(C(\mathbb{R}^3), \delta_3)$  in  $L^2(C(\mathbb{R}^3), \mu_\beta)$ . As  $\beta$  varies in  $\mathbb{R}^+$ , the induced Poisson measures  $\mu_\beta$  are mutually singular. We thus obtain an uncountable family of distinct systems of imprimitivity for  $(C(\mathbb{R}^3), \delta_3)$ . There are many more such systems!

So far we've only described the kinematics of interacting particles. The possible dynamics of interacting particles can also be obtained

from the group theoretic point of view. Of course, at this point, it becomes more natural to replace the Newtonian space-time group  $\mathcal{E}_3 \times T$  by the Galilean or Lorentz group and to also consider theories of interacting events and trajectories. Such analyses are capable of deriving the standard Schrodinger and Dirac equations for N-interacting particles. Thus, after a long arduous journey requiring some of the deepest results of modern mathematics, we arrive at the place where the physicist usually begins! But this trip is well worth taking, because in making it, we gain a much deeper understanding of the physicist's usual starting point. The trip also suggests other possible starting points.

One is now in a position to reap a fantastic harvest. By studying the Schrodinger equation with a Coulomb interaction potential, one can obtain the spectrum of atoms, molecules, positronium, munonium, etc. One can obtain a theoretical understanding of valence and of the periodic table. The structure of small molecules can be directly predicted, and at least qualitative insights can be obtained for the structure of large molecules. These successes have been repeated in the study of the structure of nuclei and more recently in the study of hadrons. With the recent proliferation of quarks having various "flavors, colors, and tastes", one can anticipate a similar theory of quark chemistry based upon an even more fundamental sub-domain. In all these theories, symmetry principles and group theory play a fundamental role. The importance of group theory in analysing these phenomena was first clearly seen by Weyl [84] and Wigner [90] (see the surveys by Michel and O'Raifeartaigh

in [98]). All that is further needed for a fundamental comprehensive understanding of the physical world is quantum statistical mechanics and quantum field theory.

### §3.6. Quantum Statistical Mechanics,

The goals of quantum statistical mechanics are identical to those of classical statistical mechanics: to derive the bulk properties of matter starting from an underlying atomic model. For example, one might model a container of gas by first assuming the existence of a microscope capable of yielding "the locations of the atoms in the gas", i.e., a microscope which outputs a finite subset of the region enclosed by the container. If one now assumes that one is looking at a "composite system of interacting quantum mechanical" particles, then one is led to the standard description of such a gas. The main problem with this program is that the Gibb's state describing a gas in equilibrium turns out not to be a pure vector state, but, instead, an impure density matrix state. The problem of justifying on the basis of first principles the use of the various canonical ensembles is the same for classical and quantum statistical mechanics (see [100] for a thorough discussion of the foundations of both classical and quantum statistical mechanics). We feel the "cure" is also the same: to consider all probability measures on our logics as capable of describing individual systems and to allow the dynamics to be governed by a stochastic process. In this way the many successes of quantum statistical mechanics (e.g., its theories of the specific heats of gases; of black body radiation; of plasmas; of low temperature phenomena; etc.) can be obtained from first principles without having to have recourse to (probably non-existent) ergodic theorems.

Over the past decade there has been a great deal of activity in the study of the limiting properties of atomic models as the number of atoms  $N$  goes to infinity, (i.e., in the thermodynamical limit see [104]). In fact, interest has focused upon the explicit construction of models

containing an infinite number of particles. In these models one can give very precise mathematical definitions of phenomena such as phase transitions. Wilson's pioneering work on the renormalization group appears to be leading to a satisfactory understanding of many such critical phenomena (see [94]). In the study of these models one is led away from the standard assumption that  $\mathcal{L}^d \cong \text{Proj}(H)$ . Instead, one typically assumes that one has a diagram  $\{\mathcal{A}_{c_\Lambda}\}$ , where  $\Lambda$  varies over the bounded subsets of  $R^3$  (or  $Z^3$ , or  $R^4$ , etc.) and where  $\mathcal{A}_{c_\Lambda}$  is the partial algebra of continuous real-valued observables in  $\Lambda$ . One also assumes that the diagram  $\{\mathcal{A}_{c_\Lambda}\}$  is isomorphic to the subdiagram of Hermitian elements in a diagram of  $C^*$ -algebras  $\{A_\Lambda\}$ , i.e., that  $\{\mathcal{A}_{c_\Lambda}\} \cong \{\text{Herm}(A_\Lambda)\}$ . Instead of working directly with  $\{\text{Herm}(A_\Lambda)\}$ , one usually works with  $\text{Herm}(A_\infty)$ , where  $A_\infty$  is the direct limit  $C^*$ -algebra,  $\lim_{\leftarrow} \{A_\Lambda\}$ . One of the usual reasons for passing to the direct limit  $A_\infty$  is that one can assume that the natural action of the Euclidean group  $\delta_3$  on  $R^3$  induces an action on  $A_\infty$  (and, as we've seen, symmetry principles are crucial for obtaining results). But  $\delta_3$  also acts naturally on the diagram  $\{A_\Lambda\}$  if it is thought of as an object in the category  $\text{ginj-}(C^*\text{-algebras})$ . One could thus create a satisfactory theory based upon the diagram  $\{A_\Lambda\}$ , while avoiding having to introduce certain extra ad hoc assumptions to eliminate consideration of "unphysical" states on  $A_\infty$  (see, for instance, the discussion by B. Simon in [69]).

### §3.7. Quantum Field Theory.

The original goal of quantum field theory was to describe such phenomena as the interaction of electrons with the electro-magnetic field. From the very beginning, all non-trivial models were plagued with infinities. This provoked some to attempt to abandon the continuum and to work with discrete space-times or to make other radical departures from the traditional formalism. But the collection of ad hoc tricks to remove the infinities, called renormalization theory, yielded such accurate predictions that it was clear that there was at least some "truth" in the traditional formalism. The traditional formalism assumed the existence of observable fields  $\varphi(t, x)$  and  $\pi(t, x)$  obeying the canonical commutation relations

$$(3.7.1) \quad [\varphi(t, x), \varphi(t, y)] = 0 = [\pi(t, x), \pi(t, y)]$$

$$[\varphi(t, x), \pi(t, y)] = i\hbar\delta(x-y) ,$$

and of a Hamiltonian given (for example) by

$$(3.7.2) \quad H = \int_{R^3} [\pi^2 + (\nabla\varphi)^2 + m^2\varphi^2 + \lambda\varphi^4] dx .$$

For systems having finitely many degrees of freedom, von Neumann showed that the corresponding Heisenberg commutation relations (in Weyl form) had a unique representation; but von Neumann's result does not extend to systems having infinitely many degrees of freedom. In fact, work of Mackey and Glimm shows that there are so many representations of (3.7.1) that they cannot even be parametrized by countably many real-valued parameters. Thus, one cannot "survey" all the possibilities. The traditional favorite was the Fock representation. But, during the

the 1950's it was realized that the Fock representation could almost never be used to describe Euclidean invariant interacting systems (e.g., one in which sense could be made of (3.7.2)). Over the past ten years one has learned how to make (3.7.1) and (3.7.2) precise, at least in two-dimensional space-time (see [66]). In four-dimensional space-time there are arguments which suggest that the equal time commutators in (3.7.1) will not exist in interesting models (even after being smeared in the space variables). Thus, one will have to begin by making more basic assumptions. (Note that in our approach the commutation relations for systems of finitely many degrees of freedom was not a basic assumption, but, instead, a consequence of looking for systems of imprimitivity for  $(R^3, E_3)$ .)

In general, a (smeared, scalar) field is defined to be an assignment  $f \rightarrow \varphi(f)$  of a real-valued observable  $\varphi(f)$  to every test function  $f$  on space-time from an appropriate space  $\mathcal{S}$  of such test functions (e.g., the rapidly decreasing  $C^\infty$ -functions). The assignment should be linear and continuous. One also usually assumes that the field is relativistically invariant (i.e., one has an action of the inhomogeneous Lorentz group  $L$  on the partial algebra of real-valued observables such that  $\varphi(g(f)) = g\varphi(f)$ ) and local (i.e., if  $\{f_i \in \mathcal{S}\}$  have space-like separated supports, then the family of observables  $\{\varphi(f_i)\}$  is compatible). In classical field theory one usually models a field by a continuous function  $\varphi(x)$  on space-time. Associated with such a  $\varphi(x)$  one has the smeared field  $f \rightarrow \int_{R^4} f(x)\varphi(x)dx$ . Smeared fields can be very singular locally. In quantum field theory one usually assumes that one's fields

satisfy the Wightman axioms. Namely, that there is a complex separable Hilbert space of states  $H$  on which we are given a (projective) representation of  $L$  which satisfies the spectral condition that the spectral resolution of the energy-momentum vector observable is supported on the positive light cone. Furthermore, there is a unique invariant vector  $\psi_0 \in H$  (the "vacuum"), and a dense invariant domain  $D \subset H$  such that each  $\varphi(f)$  is represented by a Hermitian operator on  $D$ . Major successes of the Wightman axioms include the PCT Theorem, the Spin-Statistics Theorem, and the Haag-Ruelle scattering theory. The Spin-Statistics Theorem provides a fundamental justification for a rule that was introduced ad hoc into particle physics and is crucial for the theoretical explanation of the Mendeleev table and many other important phenomena. The Haag-Ruelle scattering theory is important because in most scattering experiments one has direct information only about the incoming and outgoing states and physicists tend to prefer to describe the incoming and outgoing states using a particle as opposed to field language. A major weakness of the Wightman axioms is that there are still no known non-trivial models in four-dimensional space-time. Non-trivial models have been constructed in two-and three-dimensional space-times and they exhibit all the phenomena one expected of them from perturbation theory arguments. It is not clear at this point whether more hard analysis will yield the appropriate four-dimensional models or, instead, one will need a major departure from the Wightman axioms to obtain the desired models. One alternative approach is to start with the diagram of local partial algebras of observables  $\{\mathcal{O}_\Lambda\}$ , where  $\Lambda$  is a bounded subset of space-time, and make

direct assumptions on its structure. Using such an approach, Araki [102] was able to derive an analog of the Haag-Ruelle scattering theory and Haag and Kastler [103] were able to give a theoretical explanation of super-selection sectors.

### §3.8. Final Remarks.

We have presented a coherent approach to quantum mechanics which yields all its positive results while simultaneously short-circuiting all its paradoxes. For example, we have no "measurement problem". The measurement problem is avoided by not introducing the notion of the "system itself". Our fundamental axioms relate directly to the structure of the deep logic which itself represents interlocking modes of perceiving. E-P-R type paradoxes are avoided in a similar fashion. In particular, even if the probability is one of yielding the value  $a$  upon measurement of the observable  $A$ , we say nothing about "its value" if a measurement is not made. On the other hand, we have no ontology, not even when the deep logic is Boolean. We thus have an adequate formulation of how the world appears, but not of what it actually is.

### References

1. Aronson, J. L. and Edwards, D. A., "Quantum Logics: I-IV"(mimeographed notes).
2. Bell, J. S., "On the Problem of Hidden Variables in Quantum Mechanics," Reviews of Modern Phys. 38, 447 (1966).
3. Birkhoff, G. and MacLane, S., A Survey of Modern Algebra, MacMillan, New York, 1964.
4. Birkhoff, G. and von Neumann, J., "The Logic of Quantum Mechanics," Annals of Mathematics 37, 823 (1936).
5. Bohm, D., "A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden Variables' I," Physical Review 85, 166 (1952).
6. Bohm, D., "A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden Variables' II," Physical Review 85, 180 (1952).
7. Bohm, D., "Quantum Theory as the Indication of a New Order in Physics," Foundations of Physics.
8. Bohr, N., "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" Physical Review 48, 696 (1935).
9. Bohr, N., Kramers, H. A., and Slater, J. C., "The Quantum Theory of Radiation," Philosophical Magazine 47, 785 (1924).
10. Bohr, N. and Rosenfeld, L., "Field and Charge Measurement in Quantum Electrodynamics," Physical Review 78, 794 (1950).
11. Bopp, F., "The Principles of the Statistical Equations of Motion in Quantum Theory," in S. Karner (ed.), Observation and Interpretation in the Philosophy of Physics, Dover Publication, New York, 1962.
12. Bub, J., "On the Completeness of Quantum Mechanics," in C. A. Hooker (ed.), Contemporary Research in the Foundations and Philosophy of Quantum Theory, Reidel, Dordrecht, 1973.
13. Bub, J., "The Interpretation of Quantum Mechanics," D. Reidel Pub. Co. Dordrecht, 1974.
14. Colodny, R. G., Paradigms and Paradoxes: R.G. (ed.), The Philosophical Challenge of the Quantum Domain," University of Pitts. Series in the Phil. of Science, vol. 5, (1972).
15. Dirac, P. A. M., The Principles of Quantum Mechanics, Oxford University Press, London, 1958 .

16. Domotor, Z., "The Probability Structure of Quantum Mechanical Systems," Synthese 29, 155 (1974).
17. Einstein, A., "Remarks to the Essays Appearing in This Collective Volume," in P.A. Schilpp (ed.), Albert Einstein: Philosopher-Scientist, Harper and Row, New York, 1959.
18. Einstein, A., Podolsky, B., and Rosen, N., "Can Quantum-Mechanical Description of Physical Reality be Considered Complete?" Physical Review 47, 777 (1935).
19. Emch, G. G., Algebraic Methods in Statistical Mechanics and Quantum Field Theory, Wiley - Interscience, New York, 1972 .
20. Federer, H. Geometric Measure Theory, Springer-Verlag, New York, (1969).
21. Feynman, R. P., "Concept of Probability in Quantum Mechanics," Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, pp. 533-541 1951 .
22. Feynman, R. P., "Space-time Approach to Non-Relativistic Quantum Mechanics," Reviews of Modern Physics 20, 367 (1948).
23. Feynman, R. P. and Leighton, R. B. and Sands, M., The Feynman Lectures on Physics, Addison-Wesley, 1963 .
24. Finkelstein, D., "The Physics of Logic", in Paradigms and Paradoxes: The Philosophical Challenge of the Quantum Domain, ed. R. G. Colodny. Univ. of Pittsburgh Press, 1972 .
25. Foulis, D. J. and Randall, C. H., "Empirical Logic and Quantum Mechanics," Synthese 29, 81 (1974).
26. Freistadt, G., "The Causal Formulation of Quantum Mechanics of Particles," Supplemento del Nuovo Cimento 5, 1 (1957).
27. Gleason, A. M., "Measures on the Closed Subspaces of a Hilbert Space," Journal of Mathematics and Mechanics 6, 567 (1957).
28. Greechie, R. J., "An Orthomodular PO set with a Full Set of States Not Embeddable in Hilbert Space," Caribbean Journal of Science and Mathematics 1, 15 (1969).
29. Greechie, R. J., "Some Results From the Combinatorial Approach to Quantum Logic," Synthese 2, 113 (1974).
30. Greechie, R. J. and Gudder, S., "Quantum Logics," in C.A. Hooker (ed.), Contemporary Research in the Foundations and Philosophy of Quantum Theory, Reidel, Dordrecht, 1973.
31. Grothendieck, A., "Technique De Decente Et Theorems D'Existence En Geometrie Algebrique I-IV", Seminar Bourbaki, Exp. 190, 195, 212, 221, 1959-60, 1960-61.

32. Gudder, S., "Hidden Variables in Quantum Mechanics Reconsidered," Review of Modern Physics 40, 229 (1958).
33. Halmos, P. R., Measure Theory, D. van Nostrand Company, Princeton, 1950 .
34. Heisenberg, W., Physics and Philosophy, New York: Harper and Row,
35. Hilbert, D., Foundations of Geometry, The Open Court Publishing Co., 1965 .
36. Jammer, M., The Conceptual Development of Quantum Mechanics, McGraw-Hill, New York, 1966.
37. Jammer, M., The Philosophy of Quantum Mechanics, Wiley-Interscience, New York, 1974.
38. Jauch, J., Foundations of Quantum Theory, New York: Addison-Wesley, New York, 1968.
39. Jordan, P., von Neumann, J., and Wigner, E. P., "On an Algebraic Generalization of the Quantum Mechanical Formalism," Annals of Mathematics 35, 29(1934).
40. Kadison, R., "Transformations of States in Operator Theory and Dynamics," Topology Vol. 3, Suppl. 2, 1965, pp. 177-198.
41. Kochen, S., and Specker, E., "The Problem of Hidden Variables in Quantum Mechanics," Journal of Mathematics and Mechanics 17, 59 (1967).
42. Kuhn, T., The Structure of Scientific Revolutions, Chicago: University of Chicago Press, Chicago, 1962.
43. Lawvere, F., "The Category of Categories as a Foundation for Mathematics," Proc. of the Conf. on Categorical Algebra, La Jolla, 1965, ed. S. Eilenberg, et al., Springer-Verlag, New York, 1966.
44. Ludwig, G., "The Measuring Process and an Axiomatic Foundation of Quantum Mechanics," In Foundations of Quantum Mechanics, 1971, II Corso, Academic Press Inc., New York, .
45. Mackey, G. W., The Mathematical Foundations of Quantum Mechanics, W. A. Benjamin, Inc., New York 1963 .
46. Mackey, W. G., "Infinite Dimensional Group Representations," Bulletin of the American Mathematical Society 69, 628 (1963).
47. Mackey, G. W., Induced Representations of Groups and Quantum Mechanics, New York: W. A. Benjamin, Inc., 1963 .

48. Mackey, G. W., "Ergodic Actions of the Euclidean Group and the Canonical Commutation Relations," In Proceeding of the 1967 International Conference on Particles and Fields, ed. Hagen, C., Garalnik, G. and Mathur, V. Interscience Pub., (1967).
49. Mackey, G. W., "Ergodic Theory and Its Significance for Statistical Mechanics and Probability Theory," Advances in Mathematics 12, pp. 178-268 (1974).
50. MacLane, S., "Categorical Algebra", Bull. Amer. Math. Soc. (1965).
51. MacLane, S. and Birkhoff, G., Algebra, MacMillan, New York, 1970.
52. Mandelbrot, B., Fractals: Form, Chance and Dimension, W. H. Freeman and Co., San Francisco, 1977.
53. Matheron, G., Random Sets and Integral Geometry, New York: John Wiley and Sons, Ltd., 1975 .
54. Messiah, A. Quantum Mechanics, Amsterdam: North Holland Publishing Co., 1964.
55. Murray, F. J. and von Neumann, J., "On Rings of Operators," Annals of Mathematics 37, 116 (1936).
56. Nelson, E., "Derivation of the Schrodinger Equation from Newtonian Mechanics," Physical Review 150B, 1079 (1966).
57. Newton, T. D. and Wigner, E. P., "Localized States for Elementary Systems," Reviews of Modern Physics 21, 400 (1949).
58. Putnam, H., "Is Logic Empirical?" in Proceedings of the Boston Colloquium in the Philosophy of Science, 1966-1968. ed. R. S. Cohen. New York: Humanities Press, 1969.
59. Reid, C., Hilbert, Springer-Verlag, 1972 .
60. Sachs, M., "A New Approach to the Theory of Fundamental Processes," British Journal for the Philosophy of Science 15, 213 (1964).
61. Schilpp, P. A., ed. Albert Einstein: Philosopher-Scientist. 2 volumes. New York: Harper and Bros., Harper Torchbooks, (1959). Now published by The Open Court Publishing Co., La Salle, Ill.
62. Schweber, S., An Introduction to Relativistic Quantum Field Theory, Row, Peterson and Company, Evanston and New York, 1961 .
63. Segal, I. E., "Postulates for General Quantum Mechanics," Annals of Mathematics 48, 930 (1947).

64. Segal, I. E. Mathematical Problems of Relativistic Physics,  
Lectures in Applied Mathematics, American Mathematical Society,  
Providence, I. R., 1963 .
65. Segal, I. E. "Algebraic Integration Theory", Bull. Amer. Math. Soc.  
May, (1965) .
66. Simon, B., The Euclidean (Quantum) Field Theory, Princeton: Princeton  
Univ. Press 1974 .
67. Sneed, J. D., "Quantum Mechanics and Classical Probability Theory",  
Synthese 21, 34 (1970) .
68. Stein, H. "On the Conceptual Structure of Quantum Mechanics,"  
In Paradigms and Paradoxes, ed. Colodny, Univ. of Pittsburgh  
Press, 1972 .
69. Streater, R. F., Mathematics of Contemporary Physics, Academic Press,  
New York, 1972 .
70. Streater, R. F. and Wightman, A. S., PCT, Spin and Statistics and All  
That, New York: W. A. Benjamin, Inc., 1964 .
80. Varadarajan, V. S., "Probability in Physics and a Theorem on Simultaneous  
Observability," Communications in Pure and Applied Mathematics  
15, 189 (1962) .
81. Varadarajan, V. S., Geometry of Quantum Mechanics, D. von Nostrand  
Co., Princeton, N. J., 1968 .
82. Velo, G. and Wightman, A., Constructive Quantum Field Theory,  
Springer Lectures Notes in Physics, Vol. 25, New York, 1973 .
83. von Neumann, J., Mathematical Foundations of Quantum Mechanics,  
Princeton: Princeton Univ. Press, 1955 .
84. Weyl, H., The Theory of Groups and Quantum Mechanics, Dover  
Publications, Inc., New York, 1931 .
85. Weyl, H., The Classical Groups, Princeton: Princeton Univ. Press,  
1946 .
86. Weyl, H., Space, Time, Matter, Dover Publications, Inc., 1950 .
87. Whittaker, E. A., A History of the Theories of Ether and Electricity,  
2 volumes. New York: Harper and Row, Harper Torchbooks. (1960).
88. Wick, G. C., Wightman, A. S. and Wigner, E. P., "The Intrinsic Parity  
of Elementary Particles," The Physical Review 88, 101 (1952) .
89. Wightman, A. S., "On the Localizability of Quantum Mechanical Systems,"  
Reviews of Modern Physics 34, 845 (1962) .

90. Wigner, E. P., "Unitary Representations of the Inhomogeneous Lorentz Group," Annals of Mathematics 40, 149 (1939).
91. Simms, D. and Woodhouse, N., Lectures on Geometric Quantization, Springer Lectures Notes in Physics #53, New York, 1976.
92. Bleuler, K. and Teetz, A. Differential Geometrical Methods in Mathematical Physics, Lecture Notes in Math. #570, New York, 1977.
93. Guillemin, V. and S. Sternberg, Geometric Asymptotics, Mathematical Surveys #14, American Math. Soc., Providence, R. I., 1977.
94. Brey, J. and Jones, B., Critical Phenomena, Springer Lecture Notes in Physics #54, New York, 1976.
95. Nelson, E., Dynamical Theories of Brownian Motion, Princeton University Press 1967.
96. Wightman, A. S., "Hilbert's sixth problem: mathematical treatment of axioms of physics", in Mathematical Developments Arising from Hilbert Problems, Proceedings of Symposia in Pure Mathematics, Vol. XXVIII, Part 1, American Mathematical Society, Providence, Rhode Island, 1976.

97. P. D. Finch, "On the Structure of Quantum Logic", Journal of Symbolic Logic, 1969, 34, 275-282.
98. V. Bargmann, Group Representations in Mathematics and Physics, Lecture Notes in Physics 6, Springer-Verlag, New York, 1970.
99. E. Wigner, Group Theory, Academic Press, New York, 1959 .
100. R. Jancel, Foundations of Classical and Quantum Statistical Mechanics, Pergamon Press, New York, 1969.
101. A. Volkel, Fields, Particles and Currents, Lecture Notes in Physics, 66, Springer, New York, 1977.
102. H. Araki, "Local Quantum Theory I", in Local Quantum Theory, Proceedings of the Int. School of Phys. Enrico Fermi, Course 45, Ed. R. Jost, Academic Press, New York, 1969, p. 65-96.
103. R. Haag and D. Kastler, "An Algebraic Approach to Quantum Field Theory", Jour. Math. Phys. 5, 1969, p. 848-61.
104. Ruelle, D., Statistical Mechanics Rigorous Results, W. A. Benjamin, New York, 1969.

