

FROM SCHRODINGER TO DIRAC : A STUDY OF RELATIONS AND STATISTICS

TERRANCE QUINN

(Received 26 November 1998; Revised ??)

Analysis of the quantum theories given by Schrodinger, Heisenberg and Dirac reveal a general structure in which 'primary relations' can be distinguished from 'secondary determinations'. This has a bearing on several issues, including simultaneous measurement, uncertainty relations, Schrodinger's equation, Dirac's equation and general state space transformations.

Introduction

The 20th century saw the emergence of quantum physics, a range of theories quite distinct from the classical physics of the 19th century. Rising on efforts to understand certain complex processes, the quantum physics was able to achieve some measure of success, where the earlier physics remained frustrated. There has been much debate, however, as to the nature of the new theories, and the discussion continues to this day. Moreover, several arguments assert paradoxes of one kind and another. Yet, despite problems and unanswered questions, predictions of the quantum physics continue to be experimentally verified.

This state of affairs can be viewed as a positive stage in the understanding of science. For in the first place, actual physical processes continue, with or without successful enquiry. In the second place, data, as given, is free of paradox, for paradox occurs in the logical order of terms, relations and propositions. (So a breakdown in logic does not mean a breakdown in actual physics but is evidence that further progress remains to be made, that new terms and relations are required.) In the third place, there is a contemporary quest for "unity" in physics. It follows that the alleged paradoxes and the conflicting opinions can be considered not as justification for an obscurantism, but as an invitation, both to find resolutions and to meet the needs of the developing science. Yet, part of the problem is where to begin, particularly when

1991 AMS Mathematics Subject Classification : 81, 82.

Key words and phrases : quantum physics, quantum dynamics, state, measurement, primary relations, secondary determinations.

there is such diversity in the literature. Admittedly, there have already been comprehensive texts organizing the various views on quantum theories. But that there are various views is partly what calls for further investigation.

Our paper then is an effort to review what has been going forward, both in the sense of analysing development and in the sense of isolating certain ambiguities and inconsistencies. The paper therefore is partly historical in its purpose. And it would seem that by investigating origins of current views, some new progress is possible, and that the results are relevant to contemporary issues. Indeed, the analysis would, at least in part, seem to expose certain oversights, which remain embedded in modern work such as quantum field theories. So it is a purpose of this paper to give at least some partial analysis (in hindsight) of the successful Quantum Physics, and so to hopefully contribute to the beginnings of some future account of the current successes of contemporary Gauge Theories and Quantum Dynamics.

The structure of the paper is as follows. The first three sections are devoted to a study of the work of Schrodinger, Heisenberg and Dirac. Section 1 is mainly concerned with the papers of Schrodinger, although some attention is given to the work of De Broglie, because of his influence on Schrodinger's wave mechanics. Section 2 is relatively brief and concerned exclusively with the uncertainty relations of Heisenberg. Section 3 is the longest, because Dirac was so thorough in his construction of an axiomatic structure. Since his results are cumulative, it seemed advisable to consider a range of topics, in sequence. In Section 4 some effort is made toward understanding not quantization in general, but quantization within the Hamiltonian context. And while the results given are incomplete, they do seem to lead to certain generally valid propositions. So it is hoped that Section 4 might be a beginning for later work on the general problem of the heuristic context of quantization dynamics.

1. Wave Mechanics

The theory of "wave-mechanics" is often thought to have begun with the work of Louis de Broglie. In fact, Schrodinger himself acknowledges de Broglie's papers for leading him to his own "initial deliberations" [11]. So this section begins by raising certain questions about de Broglie's "wave theory". The second part of the section then concentrates on Schrodinger's papers found in *Collected Papers on Wave Mechanics*, [11] particularly parts I-IV of 'Quantisation as a problem of proper values'.

To begin, recall that in de Broglie's time there were various experiments on light diffraction suggesting that light had a wave property; and there was the photoelectric effect, accounted for by Einstein with his theory of light quanta. As summarised in Jammer's text [7, p. 243], de Broglie tried to put these results together.

He "considered the motion of a particle of rest mass m_0 and velocity $v = \beta c$ with regard to a stationary observer and assumed that the particle is the seat of a periodic internal phenomenon whose frequency ν_0 is given by $m_0 c^2 / h$, where h is Planck's constant and $m_0 c^2$ the internal energy of the particle. Now, the stationary observer, de Broglie maintained, ascribes to the particle the energy $m_0 c^2 / (1 - \beta^2)^{1/2}$ and hence to the periodic phenomenon the frequency $\nu = \nu_0 / (1 - \beta^2)^{1/2}$. Looking on the internal phenomenon, however, the observer will see its frequency, owing to the relativistic time dilation, being lowered to $\nu_1 = \nu_0 (1 - \beta^2)^{1/2}$ and its vibrations varying as $\sin 2\pi \nu_1 t$." To account for the difference between ν and ν_1 de Broglie introduced a "fictitious wave associated with the motion of the mobile which spreads with velocity c/β and frequency ν as defined above. He then could prove that if at the beginning the internal phenomenon of the moving body is in phase with the wave, 'the harmony of phase will persist'. Since the body, which at time $t = 0$ coincided with the wave at the origin, is at time t , at $x = vt$, its internal phenomenon is proportional to $\sin (2\pi \nu_1 x / \nu)$, whereas the wave at that place is given by

$$\sin [2\pi \nu (t - \beta x/c)] = \sin [2\pi \nu x (1/\nu - \beta/c)].$$

But by definition $\nu_1 = \nu (1 - \beta^2)^{1/2}$, which proves the theorem." This result", continued de Broglie, "suggests, that any moving body may be accompanied by a wave and that" it is impossible to disjoin motion of body and propagation of wave".

In a second and third memoir de Broglie demonstrated that the velocity $v = \beta c$ is the group velocity of the associated wave, or phase wave. More generally to a particle with kinetic energy E and momentum (p_x, p_y, p_z) is associated with the wave-function

$$\psi = \exp \left\{ \frac{2\pi i}{h} (Et - p_x x - p_y y - p_z z) \right\} \quad (1.1)$$

so that the wavelength λ is h/p , $p = (p_x^2 + p_y^2 + p_z^2)^{1/2}$ and the group velocity V is given by $V = \lambda \nu = E/p$.

Now, in empirical science there is a basic requirement that hypotheses be verifiable in data. This is borne out, if not by literature on science, at least by scientific method itself. But this then excludes any problem concerning the picture of objects too small to be seen. For the seen image can be verified only by seeing. Therefore when the seeing cannot occur, all that can be verified are equations, and terms implicitly defined by equations. (See also [5] p. 10, par. 3) Since by hypothesis

the "fictitious wave" is "internal" to the particle, and there are no "looks" inside particles, it follows that the picture of such a wave is non-verifiable. It may be claimed however that, while waves have not been observed as such, there are well known experiments which verify de Broglie's relation $\lambda = h/p$ and so give the de Broglie "internal periodic motion" empirical significance.

To determine λ experimentally, there must be a diffraction pattern. Measurement provides the distance between the "crests" in the laboratory frame, and these distances give a number proportional to λ . A particle is associated with a given crest when a mark on the apparatus (e.g. photoplate) corresponds to a collision of that single particle. So a given crest is determined by the clustering of aggregates of collisions and the quantity λ is associated with the aggregate. The de Broglie relation is verified in as much as an incident particle of momentum p is diffracted to the crest with $\lambda = h/p$. But this does not mean that the "particle" is a "wave", but that data on an aggregate of collisions can cluster according to a prescribed pattern. [It is perhaps worth noting the similarity here to the fact that a single type of distribution (e.g. a normal distribution) may pertain to several differently defined events. But it would be contradictory to conclude that the events could then be identified.]

Now consider de Broglie's wave function ψ , which for a particle of momentum p and energy E is defined by equation (1.1). Mathematically, the map $(p_x, p_y, p_z, E) \rightarrow \psi$ is given by definition. But the group velocity of ψ equals the particle velocity, and in certain situations ψ implicitly defines a range of possible trajectories. But as in the above, this has only been verified in experiments involving large numbers of particles. Furthermore, exceptional single trajectories do not affect the conclusions. In other words, the wave functions would not seem to refer to single particles.

Proceeding now to Schrodinger's wave mechanics, one place to start is with his paper Part I on "proper values" [11] where Schrodinger treats the problem of the non-relativistic hydrogen atom. Part of his aim is to show that Sommerfeld's quantum conditions can be replaced by another more "natural" postulate. He starts by considering the Hamilton-Jacobi differential equation.

$$H(q, \frac{\partial S}{\partial q}) = E, \quad p_j = \frac{\partial S}{\partial q_j} \quad (1.2)$$

where

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad \frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}.$$

Here S is the classical generating function for a canonical transformation and p_i, q_i

are the canonically conjugate variables. Integration of the equation, with suitable boundary conditions, provides the trajectory.

$$(q_1(t), \dots, q_n(t), p_1(t), \dots, p_n(t))$$

in "phase-space". In his calculation Schrodinger introduces a function ψ and constant K by setting $S = K \log \psi$. The Hamilton-Jacobi equation becomes

$$H(q, \frac{K}{\psi} \frac{\partial \psi}{\partial q}) = E.$$

Choosing rectangular coordinates and using the Hamiltonian function for Keplerian motion, the equation then becomes

$$Q(x, y, z) = \nabla^2 \psi - \frac{2m}{K^2} (E + \frac{e^2}{r}) \psi^2 = 0 \quad (1.3)$$

where e = charge, m = mass of the electron and $r = \sqrt{x^2 + y^2 + z^2}$ is the distance of the orbiting electron to the centre of mass of the atom. He then introduces a new problem, namely, the variational problem

$$\delta \int \int \int dx dy dz Q(x, y, z) = 0$$

where integration is over all space. The variation yields

$$\nabla^2 \psi + \frac{2m}{K^2} (E + \frac{e^2}{r}) \psi = 0 \quad (1.4)$$

and

$$\int df \delta \psi \frac{\partial \psi}{\partial n} = 0$$

where df is a surface element.

A first remark is that electrons cannot be observed in orbit about the hydrogen nucleus. A second is that in equation (1.3), $K \log \psi$ is the generating function for a canonical transformation, and so up to boundary conditions, defines possible trajectories of an electron in the atom. But by the first remark this would seem to make the relation so-far non-verifiable. However, the success of equation (1.4) in spectral analysis is well known. The problem then is to determine what it is that has been verified. But Schrodinger's wave-equation (1.4) can be written

$$\nabla^2 \psi - \frac{e^2}{r} \psi = E\psi$$

an operator equation providing "proper-values" E for solutions $\psi = \psi_E$. Verification of the equation then has involved the association of wave-lengths to spectral lines, as in measurements described above. For each spectral line of wave-length λ say, the "frequency" ν is defined by $\nu = c/\lambda$. Then Planck's rule is invoked to obtain $E = h\nu$. The spectrum of energies $E = h\nu$ obtained experimentally is then compared to the proper-values of the operator equation. But each spectral line, l_E say, is determined by a clustered aggregate of photons, each of energy $E = h\nu_E$. So up to experimental error, the operator equation gives the range of possible energies of single photons involved in the process.

Schrodinger hypothesised that the solution ψ represent "some vibration process in the atom". [11] Indeed, in his second paper [11] he suggests an "innate connection between wave propagation and the movement of the representative point We can attempt to build up a wave group which will have relatively small dimensions in every direction. Such a wave group will then presumably obey the same laws of motion as a single image point of the mechanical system." Schrodinger therefore associates a wave-function to a particle as de Broglie did, but in such a way that the motion is governed by the classical (Hamilton-Jacobi) equation. But again, verification of the classical equation requires at least some data on single trajectories of sub-atomic electrons (not free electrons), and this remains unavailable. And if Schrodinger's proposition is verified for perhaps a free particle, then a result would be obtained about the motion of the particle, i.e. that its trajectory agree with that of a certain wave-group function. But it does not follow that the particle itself is a small wave-group.

Schrodinger comments on possible problems in his first interpretation of the wave functions, particularly when applied to several electrons. He suggests in fact that the space density of electricity, as a function of space coordinates and time, is given by $\psi\bar{\psi}$. So it must be asked what could be meant by the density function $\psi\bar{\psi}$ which is 'associated' with one of the particles of classical mechanics. It has been suggested that this is the density "internal" to the particle. But yet again this would seem to be non-verifiable. Furthermore, it was application of his theory to the Stark-Effect which vindicated Schrodinger's results. But the Stark-effect refers to the relative densities of aggregates of particles. So that even if an "internal" property such as "electrical density" were to become known, it would be a further issue to relate this to a property of aggregates. Finally, in his remarks in Part IV on the physical significance of the field scalar, similar problems are revealed. There " $\psi\bar{\psi}$ is

a kind of weight-function in the system's configuration space. The *wave-mechanical* configuration of the system is a *superposition* of many, strictly speaking of all, point-mechanical configurations kinematically possible. Thus each point-mechanical configuration contributes to the wave-mechanical configuration with a certain weight, which is given precisely by $\psi\psi^*$ [11]. Yet any single particle will only exhibit one possible, point-mechanical configuration. It would seem then that, much as with de Broglie's results, if Schrodinger's superposition principle can have any physical significance, it cannot be a property as such of individual particles.

2. The Uncertainty Relations

In particle physics measurement itself can interfere with a concrete situation. To deal with this problem Heisenberg proposed his "uncertainty relations".

Following the argument in Chapter II of his book [6] Heisenberg writes that

"the position of an electron is known to within a certain accuracy Δx at the time t (and) can be visualized by the picture of a wave packet in the proper position with an approximate extension Δx . By 'wave-packet' is meant a wavelike disturbance whose amplitude is appreciably different from zero only in a bounded region. ... The velocity of the electron corresponds to that of the wave packet, but this latter cannot be exactly defined, because of the diffusion which takes place. This indeterminateness is to be considered as an essential characteristic of the electron, not as evidence of the inapplicability of the wave picture. Defining momentum as $p_x = \mu v_x$ (where μ = mass of electron, v_x = x -component of velocity), this uncertainty in the velocity causes an uncertainty in p_x of amount Δp_x , and from the simplest laws of optics, together with the empirically established law $\lambda = h/p$, it can readily be shown that

$$\Delta x \Delta p_x \geq h. \quad (1)$$

Suppose the wave packet is made up by superposition of plane sinusoidal waves, all with wave-lengths near λ_0 . Then, roughly speaking, $n = \Delta x / \lambda_0$ crests or troughs fall within the boundary of the packet. Outside the boundary the component plane waves must cancel by interference; this is possible if, and only if, the set of component waves contains some for which at least $n + 1$ waves fall in the critical range. This gives

$$\Delta x / (\lambda_0 - \Delta \lambda) \geq n + 1$$

where $\Delta \lambda$ is the approximate range of wave-lengths necessary to represent the packet. Consequently

$$\Delta x \Delta \lambda / \lambda_0^2 \geq 1. \quad (2)."$$

On the other hand, "the group velocity v_g can be calculated from elementary optical principles to be

$$v_g = \frac{dE}{dp} = \frac{p}{\mu} = \frac{h}{\lambda\mu}.$$

Therefore, "the spreading of the wave packet is characterized by the range of velocities

$$\Delta v_g = \frac{h}{\mu\lambda_0^2} \Delta\lambda."$$

And since Heisenberg assumes that $\Delta p_x = \mu\Delta v_g$, equation (2) implies the uncertainty relations, that is

$$\Delta x \Delta p_x \geq h. \quad (3)$$

A first remark is that, as in the work of both de Broglie and Schrodinger, the electron "wave-picture" is so far not verified and probably non-verifiable. Secondly, in equations (2) and (3) [6] Δx refers to a single electron while in experiment λ is either the "wave-length" of a patterned aggregate of marks on an apparatus, or when associated to a single electron refers to a radial space (and time) coordinate of a collision. Thirdly, the equation $\lambda = h/p$ has a similar experimental basis, as discussed in Section 1 of this paper. Now if λ and $\Delta\lambda$ are taken as referring to coordinates of a single electron, then the wave argument has no basis. If λ and $\Delta\lambda$ are taken as referring to aggregates of collisions then since $\lambda\Delta x$ refers to a single electron, there would seem to be no pre-experimental grounds for comparing these terms.

Despite these particular inconsistencies, the expression $\Delta p_x \Delta p_x \geq h$ can be valid, if its terms are defined to be compatible. Either it refers to aggregates of electrons or it refers to single electrons. If it refers to single electrons verification would require determination of the error terms Δx and Δp_x , both at some time t . But this would seem to need a degree of non-interference in measurement which is contrary both to Heisenberg's assertions and to known experimental results. It would seem then that if the uncertainty relation is to be empirically significant, then it must refer to aggregates and be statistical in nature.

In fact, using statistical terms, Heisenberg goes on to give an alternate derivation of the uncertainty relations [6]. With $S(q')$ as a probability amplitude and $|S(q')|^2 dq'$ the probability of finding the numerical value of the coordinate of the electron between q' and $q' + dq'$, he defines $\bar{q}$ and Δq by

$$\bar{q} = \int q' |S'(q')|^2 dq'$$

and

$$\Delta q = 2 \int (q' - \bar{q})^2 |S'(q')|^2 dq'.$$

And similar definitions are given for $\bar{p}$ and Δp , where p is the conjugate momentum from Hamilton mechanics.

Again though questions arise. For Heisenberg equates Δq , as statistically defined, with experimental error. But as observed by Ballentine ([3], p. 168) in empirical investigation it is the band width of a histogram which corresponds to experimental error, while the statistically defined Δq is the standard deviation of the entire histogram; and certainly in general these are not the same.

The conclusion of this section is not that the uncertainty relations do not possess empirical validity. What has been shown is that if the relations are to be verified as pertaining to electrons, for example (which seem to involve randomness and suffer interference from measurement), then the meaning of the relations must be statistical. And as statistical relations, their verification must also be statistical, involving sufficiently large samples. In particular, the inequalities do not refer to any single electron.

A remark on nomenclature: the Heisenberg relations have been called both "uncertainty" relations and "indeterminacy" relations. The name "indeterminacy" however would seem inappropriate, since for the relations to have physical significance they must be verifiable. But to verify a statistical hypothesis requires that actual events be at least determinate enough to be counted. The name "uncertainty" could be retained, as long as it is understood as referring to, for example, standard deviation. But standard deviation is a statistical quantity of a frequency distribution and so does not pertain to any single event.

Finally, note that it will be in the next section that the problem of "simultaneous measurement" will be considered, particularly as it relates to Heisenberg's uncertainty relations.

3. An Axiomatic Quantum Mechanics

3(a). States and Measurement

This section (3(a)) begins by considering the term "state" as it appears in Dirac's formulation of quantum mechanics. The purpose is to determine its physical meaning, and a question for interpretation comes from the fact that Dirac seems to have used this term in two different ways. To see this, begin by recalling a few

definitions. By "dynamical variables are meant quantities such as the coordinates and the components of velocity, momentum and angular momentum of particles, and functions of these quantities - in fact the variables in terms of which classical mechanics is built up" ([5], p. 26). To each classical variable is associated a self-adjoint linear operator ξ which by definition acts on a vector space of bra and ket vectors; and the vector space is equipped with a scalar product. The operator ξ is an observable if its eigenvectors form a *complete set*, i.e., if they determine a basis. In "The General Physical Interpretation" ([5], p. 45) a "state" is given in terms of ket vectors $|x\rangle$. If "measurement of the observable ξ for the system in the state corresponding to $|x\rangle$ is made a large number of times, the average of all the results will be $\langle x|\xi|x\rangle$, provided $|x\rangle$ is normalized" ([5], p. 46). Earlier in the text, though a state of any atomic system is defined to be "any possible motion of the particles or bodies consistent with the laws of force".

This definition of state in terms of motion has been carried over from classical mechanics. But in many processes determination of such a classical state is beyond the reach of experiment. And it would be extra-scientific to assert that an atomic system were in a particular "state of motion", when the assertion could not be verified. Furthermore, even if data are available, "state" in the classical sense is not equivalent to state in the statistical sense. The average value of a measurement does not uniquely determine the value of any single measurement; and any single measurement would only be one component of the classical state. Moreover, if data under investigation were amenable to the classical definition, the statistical definition would seem to be of doubtful purpose. For example, it is not usual for statistics to play a vital role in the investigation of planetary motion. And in any case, efforts to determine the classical states of, for example, "systems" of electrons remain effectively frustrated. It seems then that for quantum mechanics the classical definition of state can have no adequate experimental basis and that therefore the statistical definition must be chosen for Dirac's axiomatic structure. (See also [5], p. 14, par. 1).

Part of Dirac's discussion includes the further problem of "collapsing state". When a photon meets a tourmaline crystal, "we are subjecting it to an observation. We are observing whether it is polarized parallel or perpendicular to the optic axis. The effect of making this observation is to force the photon entirely into the state of parallel or entirely into the state of perpendicular polarization. It has to make a sudden jump from being partly in each of these two states to being entirely in one or other of them. Which of the two states it will jump into cannot be predicted, but is governed only by probability laws" (See also [5], p. 34). But a statistical state does not refer to particular events. And a single photon is not in a statistical state - by definition a statistical state is determined by large numbers of measurements for single photons. So to assert the "superposition" or "collapse" of a state as referring to

single photons is, from the statistical point of view, to assert what has no link to statistical definition. From the classical point of view, "superposition" for single photons would seem to be non-verifiable, since either a particular motion is detected or it is not. And "collapse" of a classical state will also be non-verifiable because in general, as already mentioned, quantum mechanics concerns itself with processes with which classical method is frustrated, and where a classical state can simply not be determined. Therefore, at least, within Dirac's quantum mechanics, there seems to be no place for the "superposition" or "collapse" of single photons.

Both types of state are carried forward in Dirac's text, but are frequently not distinguished. Take, for instance, his discussion of eigenstates. "If the dynamical system is in an eigenstate of a real dynamical variable ξ , belonging to the eigenvalue ξ' , then a measurement of ξ will certainly give as a result the number ξ' " ([5], p. 35). But the statistical definition of state only requires that the average value of sufficiently large numbers of measurements approximate the eigenvalue. Hence, without contradiction, actual averages are permitted to diverge from the ideal average, particularly an actual average obtained from a single measurement.

Dirac goes on to relate eigenstates to the collapse of states. For he says that "a measurement always causes the system to jump into an eigenstate of the dynamical variable that is being measured, the eigenvalue this state belongs to being equal to the result of the measurement" ([5], p. 35). But it has already been argued that within quantum mechanics the definition of state must be the statistical one. Therefore, a single measurement is not sufficient to determine a state, "collapsed", or otherwise. So, as it stands, the assertion that the system jumps into an eigenstate is ambiguous. Perhaps though, certain propositions on the effect of measurement are possible. If single measurements are considered, one could inquire into the possibility that one measurement of an observable ξ following quickly on another would give approximately the same value. In this case a result on physical stability would be obtained. Note, however, that this has no necessary relation to a result on classical states, for by hypothesis, only one observable ξ is being measured. Alternatively, one could investigate possible statistical state transformations induced by measurement. For an example, suppose that a process is such that large numbers of electrons determine a state ψ_0 , and that measurements are made providing a sequence (E_n) of values from the energy spectrum. If there are enough terms of the sequence for it to be statistically representative, then an energy state η_1 is obtained. It might then be possible (e.g. with a filtering device) to provide a second sequence (F_n) of energies. If this second sequence also determines a state η_2 , one could consider possible relations M between η_1 and η_2 . The problem therefore is to obtain an operator M by performing a suitable statistical analysis. And this analysis would depend on a number of factors such as the particular observable being measured, the apparatus

being used, the kinds of particle in the process, the type of randomness involved, the events of interest, etc. Note that the example given is not meant to exclude other possible investigations. The main point is that if the statistical definition of state is being used, then the enquiry must be statistical, and if successful will yield a relation M between statistical state spaces.

The next part of the present discussion concerns the question of compatible observables and simultaneous measurement. In fact, it will be shown that simultaneous measurement and commutativity have no necessary relation within Dirac's structure. This is not to say that compatibility is without its possible physical significance, but that distinctions need be made. First recall Dirac's definitions from Sections 11 and 12 of his text ([5], pp. 41-48). There, if ξ is an observable with a in its (discrete) spectrum and if δ_{ξ_a} is the spectral projection for a , then for a state determined by a normalized ket $|x\rangle$, the probability of ξ having the value a is hypothesized to be the average of δ_{ξ_a} , that is,

$$P_{\xi}(a) = \langle x | \delta_{\xi_a} | x \rangle .$$

Suppose now that η is another observable, which commutes with ξ , so $\xi\eta = \eta\xi$. If b is in the (discrete) spectrum of η and as above δ_{η_b} is the spectral projection of b , then according to Dirac, for the state determined by $|x\rangle$ the probability of ξ and η having the values of a, b respectively, is given by

$$P_{ab} = \langle x | \delta_{\xi_a} \cdot \delta_{\eta_b} | x \rangle .$$

However, by a general rule of probability theory, the probability of the occurrence of independent events is the product of their respective probabilities. Hence, if ξ and η are independent in this statistical sense, the probability of ξ having the value of a and η having the value of b is

$$P_{\xi}(a) P_{\eta}(b) = \langle x | \delta_{\xi_a} | x \rangle \langle x | \delta_{\eta_b} | x \rangle .$$

And these values do not in general agree with P_{ab} given in the preceding equation. For an example, suppose

$$\delta_{\xi_a} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \delta_{\eta_b} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$|x\rangle = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Then

$$\langle x | \delta_{\xi_a} \delta_{\eta_b} | x \rangle = 0$$

but

$$\langle x | \delta_{\xi_a} | x \rangle \langle x | \delta_{\eta_b} | x \rangle = |x_1|^2 |x_2|^2.$$

Perhaps though, it will be suggested that the premise of "statistical independence" is at fault, in the sense that if $\xi\eta = \eta\xi$ then ξ and η must be dependent, and in precisely the way needed to make $P_\xi(a)P_\eta(b)$ equal to P_{ab} . But on what grounds could this be affirmed? Even if there exists a third observable, ρ say, and functions f and g such that $f(\rho) = \xi$ and $g(\rho) = \eta$, this would not seem to be enough, in general, to force statistical dependence unless, perhaps, f and g are invertible in some suitable space of functions.

Another argument might be that for a normalized ket $|x\rangle$, the formula for P_{ab} determines not the probability of ξ having the value a and η having the value b , but of ξ and η having the values a and b "simultaneously". In itself this may be a legitimate hypothesis. Although, verification of the equation may pose some problems. For to verify the probability of an event requires that the event be accessible to an observer over sequences of occasions. So in the laboratory one would at least need to be able to determine simultaneity in the experimental situation. Hence the results and definitions of special relativity would have to be drawn into the calculations. It would then be necessary to actually simultaneously determine the values for ξ and η .

There is the question however of how this relates to Dirac's axioms. His definitions associate a self-adjoint operator ξ to a classical variable u , in such a way that average values of the single variable u correspond to average values of the single operator ξ . But the above procedure would involve the simultaneous measurement of two classical variables, which is not accounted for in Dirac's definitions. Still, it might be proposed that there is some variable w with operator π , whose average values determine P_{ab} , i.e. so that

$$P_\pi(c) = \langle x | \delta_{\pi_c} | x \rangle = P_{ab}.$$

But what is that operator π ? How does the spectrum of π relate to the spectra of ξ and

η ? Is π to have some verifiable relation to ξ and η ? Some authors have suggested that π should be the product $\xi\eta$.

But again, if

$$\xi = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad \text{and} \quad \eta = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

then

$$\delta_{\xi_2} \cdot \delta_{\eta_2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

whereas

$$\xi\eta = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I.$$

So $\xi\eta$ has only one spectral projection and

$$\langle x | \delta_{\xi_2} \cdot \delta_{\eta_2} | x \rangle \neq \langle x | I | x \rangle.$$

Alternatively, it has been argued that if $\xi = f(\rho)$ and $\eta = g(\rho)$ then f and g need not be invertible, for it is measurement of ρ which allows for the simultaneous measurement of ξ and η . Now to see in what sense this could be true, consider the following. If measurement of ρ is possible then by definition ρ must be an observable for some variable w . So the spectrum of w is the operator spectrum of ρ and the average values agree, that is,

$$\langle w \rangle = \langle x | \rho | x \rangle$$

for any normalized ket vector $|x\rangle$. With this correspondence it can follow, under certain conditions on the operators involved, that $u = f(w)$ and $v = g(w)$ are the variables corresponding to ξ and η respectively, that is, the operator relation can be pulled back to the classical variables. Hence, if w is measured, then by definition one obtains values for u and v , through the functions f and g . Moreover the distribution of w then determines the distributions for u and v so that

$$\langle u \rangle = \langle f(w) \rangle = \langle x | f(\rho) | x \rangle$$

and

$$\langle v \rangle = \langle g(w) \rangle = \langle x | g(p) | x \rangle.$$

Notice however that a crucial premise is that p be an observable corresponding to the variable w . And the existence of the variable w does not seem to be guaranteed in general. If $\xi = f(p)$, for instance, one cannot define $w = f^{-1}(u)$ unless f maps the spectrum of w one to one and onto the spectrum of u . For otherwise, ambiguities arise, as is well known. As for simultaneity, it does not enter the calculations. It may be simultaneously true that $u = f(w)$ and $v = g(w)$, so that a single measurement of w provides single values for both u and v . But this simultaneity is mathematical definition, whereas the problem of physically simultaneously measuring u and v is quite distinct, as further discussed in paragraphs below.

In fact, it would seem that these difficulties are beside the point and can be avoided by adhering to Dirac's axioms. For in the first place, as mentioned above, the axioms refer to averages of single variables and single operators. And in the second place, if the problem of simultaneous measurement is of interest, then a sharp distinction must be drawn between physical simultaneity and mathematical simultaneity. Physical simultaneity refers to events accessible to observers. Mathematical simultaneity refers to intelligibly linked propositions. So a set G may simultaneously be a group, a topological space and a manifold. Or more to the point in question, in operator theory two commuting self-adjoint operators ξ and η with discrete spectra $\{a_n\}$ and $\{b_n\}$ respectively, may be "simultaneously" diagonalized, in the sense that there is one orthonormal basis $\{|x_n\rangle\}$ with respect to which both operators are diagonal. Dirac's physical axioms then require that for an eigenvector $|x_n\rangle$ with $\xi |x_n\rangle = a_n |x_n\rangle$ and $\eta |x_n\rangle = b_n |x_n\rangle$, probabilities are

$$P_\xi(a_n) = 1$$

and

$$P_\eta(b_n) = 1.$$

But the first probability may be verified by sufficiently many measurements of ξ ; the second probability may be verified in a similar way. But this need not involve physically simultaneous measurements of ξ and η ; and nor is it called for by the statistical axiom.

These remarks also have a bearing on Heisenberg's uncertainty relations. If the variable u has observable ξ , the variable v has observable η , and ξ and η do not commute, it does not seem to follow from the axioms that simultaneous measurement of ξ and η (or of u and v) is not possible. Granted, one may calculate quite generally that the product of their respective standard derivations, as defined within the

axiomatic structure, satisfy.

$$\Delta_{\xi} \Delta_{\eta} \geq \frac{1}{2} |\langle x | \xi \eta - \eta \xi | x \rangle|.$$

(See also Heisenberg's calculation referred to in Section 2.) But this provides a possible relation between statistical terms and make no references to any single event, let alone any physically simultaneous pair of events such as $u = u_0$ and $v = v_0$. In fact, as remarked on by Ballentine [3], for a given state vector $|x\rangle$, Δ_{ξ} is determined by its own sequences of measurements, as is Δ_{η} . The product of these quantities may then be compared with the deduced quantity $\frac{1}{2} |\langle x | \xi \eta - \eta \xi | x \rangle|$. So again, physical simultaneity is not part of the required statistical investigation.

Before concluding this subsection, a few remarks will be given on certain algebraic aspects of "quantization". One of the aims in "quantization" has been to map classical relations between variables u and v say, to operator relations between the corresponding observables ξ and η . That obstructions have arisen is already known. For example there is the paper of Arens and Babbitt [1] discussing the "algebraic difficulties of preserving dynamical relations" - of particular interest to them were the poisson bracket and the operator commutator. And in 1967 Kochen and Specker [8] gave their algebraic contradiction, now known as the "Kochen-Specker Paradox". But it can be shown that more elementary considerations are needed. For suppose that $w = u + v$, ξ corresponds to u , η corresponds to v and ρ corresponds to w . Then even if ξ and η commute, it does not seem to follow from the axioms that $\rho = \xi + \eta$. Note however that, as in the work of Arens and Babbitt, the equality $\rho = \xi + \eta$ is often just assumed.

The usual proof of additivity is that $\langle x | \xi + \eta | x \rangle = \langle x | \xi | x \rangle + \langle x | \eta | x \rangle$, which is a correct statement for any two operators ξ and η when $|x\rangle$ is in the domain of both ξ and η . But this proves only that there is additivity of expectation values for two given operators ξ and η , and says nothing about the operator ρ , which in the axiomatic structure corresponds to w . Recall too that commutativity has no necessary relation to simultaneous measurement. The axioms only require that samples of measurements of the single variable w give rise to the average value $\langle x | \rho | x \rangle$.

It may be argued that the average values should satisfy $\langle u \rangle + \langle v \rangle = \langle w \rangle$, and if this were the case, then the operator $\xi + \eta$ would satisfy the statistical requirements and so justify the correspondence with w . But on what grounds can such an assertion be made prior to experiment? It could be put forward as an hypothesis. But finally verification would require measurements of w , which might not even be possible. Or,

if such measurements are possible, they might require their own experiment and apparatus.

3(b). Unitary Transformations and the Equations of Motion

To obtain the equations of motion, Dirac introduces certain transformations, namely, displacement operators, referring to either distance or time. Beginning with "distance", Dirac argues that the "displacement of a state or observable is a perfectly definite process physically. Thus to displace a state or observable through a distance δx in the direction of the x -axis, we should merely have to displace all the apparatus used in preparing the state, or all apparatus required to measure the observable, through the distance δx in the direction of the x -axis, and the displaced apparatus would define the displaced state" ([5], p. 99). Dirac then makes various assumptions about such a transformation. A first is that superposition relationships between states remain invariant under the displacement. So, for example, if

$$|R\rangle = c_1 |A\rangle + c_2 |B\rangle, \quad c_1, c_2 \in C$$

and $|Rd\rangle, |Ad\rangle, |Bd\rangle$ represent the displaced kets, then

$$|Rd\rangle = c_1 |Ad\rangle + c_2 |Bd\rangle.$$

In other words, the transformation is given by a linear operator D , so that for the general ket $|P\rangle$,

$$|Pd\rangle = D |P\rangle. \quad (3.1)$$

Dirac secondly assumes that "we must have

$$\langle Qd | Pd \rangle = \langle Q | P \rangle" \text{ ([5], p. 101).} \quad (3.2)$$

Thirdly, if v (Dirac's notation) is a dynamical variable (operator), then he defines the displaced dynamical variable v_d (operator) by

$$v_d |Pd\rangle = D v |P\rangle. \quad (3.3)$$

Then using equations (3.1) and (3.2) ([5], p. 101) Dirac deduces that

$$v_d = D v D^{-1}. \quad (3.4)$$

It seems worthwhile to consider possible empirical grounds for equations (3.1) - (3.4). And while Dirac's work is already in this context, the importance of the equations would seem to warrant the effort of making some of these grounds explicit. First then is the question of linearity. In fact, initially the operation of displacement

need not be defined for more than single vectors corresponding to states determined by situations. But from the data of the double-split experiments, there can emerge the question of state "addition". And because states are represented by ket vectors, state "addition" could correspond to ket vector addition. Note however that concretely the new state is determined by a new situation. So, there is nothing in the axioms which requires that state "addition" be obtained simply by adding probabilities. On the other hand, if there are two slits, the four possible arrangements involve similarly defined processes, so one might expect relations between the possible states. Indeed, it is known that if $| \epsilon_1 S_1, \epsilon_2 S_2 >$ is the ket for having the slits open or closed, depending on whether $\epsilon_i = 1$ or 0 , $i = 1, 2$ then

$$| \epsilon_1 S_1, \epsilon_2 S_2 > = \begin{cases} (\epsilon_1 | S_1 > + \epsilon_2 | S_2 >) / (\epsilon_1^2 + \epsilon_2^2)^{1/2}, & \epsilon_1 \text{ not both zero,} \\ 0, & \epsilon_1 = \epsilon_2 = 0 \end{cases}$$

But data accumulated over investigations of many such processes, together with the Schrodinger functions of the form $e^{i\theta(r, t, R, E)}$ and their time derivatives have all partly contributed to establishing the need of an addition of states represented by vector addition, with all complex numbers as possible coefficients. It is then, once linearity is part of the structure, that the effect of displacement on linear relationships can be considered. And from empirical evidence it would seem that for $a_i \in C$, $i = 1, 2$, $D(a_1 | S_1 > + a_2 | S_2 >)$ can be equivalent to $a_1 D | S_1 > + a_2 D | S_2 >$, within the axiomatic structure. The linearity of D could then be hypothesized; for if the two vectors just given are equal, then they would also be empirically equivalent.

Next to consider is Dirac's equation (3.2). In fact, if D is surjective, this condition is equivalent to D being unitary and linear. So again, possible empirical grounds for the assertion are discussed. Now there is nothing in scientific method which can demand, prior to experiment, the invariance of either space-time components of measurement, or frequencies of events accessible to observers. On the other hand, for processes sufficiently isolated from extraneous interference, conservation of energy has proven to be a principle of limitation. For certain cases it is therefore plausible that the statistical distribution function for energy is preserved under sufficiently small spatial displacement — this would have to be verified. If true, and if for example $\{E^n\}$ is the discrete spectrum of possible non-zero energies with $\{|E^n >\}$ a basis of corresponding eigenkets, then

$$\langle E^n | H | E^n \rangle = \langle E^n | d | H_d | E^n \rangle \quad (3.5)$$

for all n . Since $E^n \neq 0$,

$$1 = \langle E^n | E^n \rangle = \langle E^n d | E^n d \rangle \quad (3.6)(a)$$

and for $n \neq m$

$$0 = \langle E^n | E^m \rangle = \langle E^n d | E^m d \rangle. \quad (3.6)(b).$$

Furthermore, if $|P\rangle$ is any ket, then the probability of the energy being E^n is given by

$$\text{Prob}(E^n) = |\langle P | E^n \rangle|^2 \quad (3.7)$$

which by hypothesis is invariant and so is equal to

$$\text{Prob}(E^n d) = |\langle Pd | E^n d \rangle|^2. \quad (3.8)$$

Evidently, if D is extended linearly from its values on the basis kets $|E^n\rangle$, then D is surjective and unitary, and satisfies equations (3.2) and (3.5) - (3.8).

Note, however, that even in this special case the linearity of D is not strictly necessary, for the relation of kets to states is one of representative to equivalence class. For instance, if $|P\rangle$ is a ket and θ is real then $e^{i\theta}|P\rangle$ would be empirically indistinguishable from $|P\rangle$. And this equivalence relation induces a partition on the transformations of the vector space. For an example consider the ket space C^2 and suppose that $|1\rangle, |2\rangle$, the standard basis vectors, are eigenkets for energies E^1, E^2 respectively. For a linear combination $a_1|1\rangle + a_2|2\rangle$, $a_i \in C$, define

$$T: (a_1|1\rangle + a_2|2\rangle) \mapsto \begin{cases} a_1|1\rangle + a_2|2\rangle, & a_2 \neq 0, \\ ia_1|1\rangle, & a_2 = 0. \end{cases}$$

Then the transformation T satisfies all of the conditions (3.5) - (3.8), preserves the modulus of the linear product, is "additive" with respect to the hypothesized state space structure, would be indistinguishable from the identity, and is neither linear nor anti-linear. Nevertheless, many known equivalence classes of state space transformations have linear representatives and so linearity is obtained by equivalence relation. Hence, in the given example T could be replaced by the identity; and with regard to displacement, D could be taken to be unitary linear.

Recall that one of the suggestions in the above remarks is that conservation of energy be applied to average values. This might be generalized to the hypothesis that average values of a complete set of commuting variables ([5], p. 57) be preserved.

But this raises questions about equations (3.3) and (3.4). For if equation (3.4) holds for all observables, then all spectra and all probabilities are preserved under displacement. But as mentioned above such invariance cannot be imposed prior to experiment. Furthermore, not only would it run counter to relativistic considerations, but one could ask the purpose of the investigation, for by hypothesis there would be no measurable differences. It seems then that equations (3.3) and (3.4) require certain restrictions. Let it be emphasized though, that this is not to say that the equations cannot be valid for some observables, but that they should not apply to all observables under all displacements.

Dirac defines operators by their average values. Hence to determine v_d in general it is necessary to consider the quantities $\langle Pd | v_d | Pd \rangle$. In some cases, simple relations might exist between the average values for the displaced kets and the original average values $\langle P | v | P \rangle$. For example suppose that the displacement operator D is unitary and that for a particular observable ξ , there exist a positive function $\lambda : d \rightarrow \lambda(d)$ so that

$$\langle Pd | \xi_d | Pd \rangle = \lambda(d) \langle P | \xi | P \rangle$$

for all kets $|P\rangle$. Then by the polarization identity it follows that

$$\langle Pd | \xi_d | Qd \rangle = \lambda(d) \langle P | \xi | Q \rangle$$

for all kets $|P\rangle$ and $|Q\rangle$. Therefore, if

$$\xi | P \rangle = | R \rangle$$

then

$$\langle Q | \xi | P \rangle = \langle Q | R \rangle$$

which implies that

$$\frac{1}{\lambda(d)} \langle Qd | \xi_d | Pd \rangle = \langle Q | R \rangle = \langle Qd | Rd \rangle.$$

Since D is surjective and $\langle Qd |$ is arbitrary, it follows that

$$\xi_d | Pd \rangle = \lambda(d) | Rd \rangle.$$

But then just as in Dirac's text [5],

$$\xi_d | Pd \rangle = \lambda(d) (D \xi D^{-1}) | Pd \rangle.$$

And so (again from the fact that D is surjective) it must be that

$$\xi_d = \lambda(d) D\xi D^{-1}.$$

This ends the remarks on possible empirical grounds for equations (3.1) - (3.4). Returning now to Dirac's discussion, his next step is to "pass to an infinitesimal displacement" ([5], p. 102). He proposes that because of physical continuity the mapping $|P\rangle \rightarrow |Pd\rangle$ should be differentiable with derivative d_x , the displacement operator for the x -direction. But some care should be taken; for trajectories in atomic processes are known to exhibit abrupt changes. But recall that states are defined statistically and that data can consist of aggregates determining, for example, diffraction patterns on a photographic plate. And under displacement the variations of the distributions are typically approximately continuous (a distribution as such being unaffected by random differences). Therefore, much as in the derivation of the classical wave equation, one may abstract from the continuum to obtain possible differentiable relations between representative kets $|P\rangle$ and $|Pd\rangle$. (See also the paragraph following Schrodinger's equation of motion, below.)

Continuing to Dirac's next chapter the topic is "The equations of motion" ([5], Ch. V). Dirac writes:

"To get a complete theory of dynamics we must consider also the connexion between different instants of time. When one makes an observation on the dynamical system, the state of the motion gets changed in an unpredictable way, but in between observations causality applies, in quantum mechanics as in classical mechanics, and the system is governed by equations of motion which make the state at one time determine the state at a later time ... They will apply so long as the dynamical system is left undisturbed by any observation or similar process. Their general form can be deduced from the principle of superposition of Chapter I." ([5], p. 108)

Notice that Dirac would again seem to be using the term state in two ways. For in reference to the unpredictable change he seems to be speaking of a classical state (positions, momenta, etc.) And yet the equations of motion he deduces are for kets, which are defined by operators and average values. But for the processes in question "we cannot observe a small system with the amount of detail which classical theory supposes. The limitation in the power of observation puts a limitation on the number of data that can be assigned..." [5] So by Dirac's own arguments the assertion of the time evolution of the classical state is non-verifiable. If, however, the term ket is adhered to in the statistical sense, then as Dirac proceeds to do, possible equations of motion for their time evolution can be deduced.

To obtain these equations, arguments much like those given for space displacement operators provide a unitary operator T , for time displacement, which is assumed to be differentiable in the sense that the following limit exists:

$$\lim_{t \rightarrow t_0} \frac{|Pt\rangle - |Pt_0\rangle}{t - t_0} = \frac{d}{dt_0} |Pt_0\rangle.$$

Then, since by definition

$$|Pt\rangle = T |Pt_0\rangle,$$

it follows that

$$\frac{d}{dt_0} |Pt_0\rangle = K(t_0) |Pt_0\rangle,$$

where the pure imaginary $K(t_0)$ is the strong limit

$$\lim_{t \rightarrow t_0} \left\{ \frac{T - I}{t - t_0} \right\}.$$

Dirac then identifies the Hamiltonian operator H with ihK and so obtains "Schrodinger's form of the equation of motion", ([5], p. 110).

$$ih \frac{d}{dt} |Pt\rangle = H(t) |Pt\rangle.$$

This equation has been remarkable in its success. And it is this success which invites further comment. The problem is to account for the identification of ihK with the Hamiltonian H . On this point, Dirac's rationale calls on "classical analogy" ([5], p. 110). It would seem however that some grounds for the relation $H = ihK$ can be given from within the bounds determined by the axiomatic structure. Indeed, it seems possible to transpose the argument found in Schrodinger's notebooks. [11] For diffraction experiments provide aggregates of data distributed as critical points of classical wave functions. So while the empirical significance is statistical, this suggests the possibility (requiring verification) of (ket) functions ψ which satisfy the classical wave equation. In particular one could hypothesize that a function

$$\psi = e^{2\pi i(kx - vt)}$$

corresponds (within the axiomatic statistical structure) to one line in the pattern, each of which, by the relations of deBroglie and Bohr determines a state verified to be of zero dispersion and fixed energy. For such an "eigenfunction", ψ_E say,

$$\frac{h^2}{2m} \nabla^2 \psi_E + V\psi_E = -\frac{h}{2\pi i} \frac{\partial \psi_E}{\partial t},$$

where V and E are the potential energy and total energy of a projectile corresponding to the given line of the pattern. Denote the energy operator $-\hbar^2/2m \nabla^2 + V$ by H . Then by linearity,

$$H\psi = i\hbar \frac{\partial \psi}{\partial t}$$

for any ket function ψ . But the general argument in Dirac's book gives

$$K\psi = \frac{\partial \psi}{\partial t}$$

Multiplying this last equation by $i\hbar$, the Schrodinger representation is obtained,

$$H = i\hbar K.$$

Suppose now that average values for energy are invariant under change of representation. In this case the energy operator can be taken to be similar to H , and so the relation between H and K remains invariant. Hence, one might hypothesize that $H = i\hbar K$ for a general class of processes.

To conclude this section, some remarks will be given on the time and space operators. For in Dirac's exposition there are in fact two kinds of transformation. Recall then that his characterization of space-displacement is in terms of apparatus ([5], p. 99). It follows that the operator D refers to a change of (spatial) frame, and that if the reasoning is to be extended to time t , the time-displacement should refer to a change of (time) frame as well. But Dirac's operators $T(t)$ refer to time evolution ([5], p. 108). Consider then a change of (time) frame $T_f(t)$, and for the process under investigation, assume that unitary operators $T_f(t)$ are verified to transform initial kets $|P\rangle$ to kets $|Pt\rangle = T_f(t)|P\rangle$. Consequently, an observable v transforms to v_t so that, by definition, the average value for v in the new frame is

$$\langle v \rangle_t = \langle Pt | v_t | Pt \rangle.$$

Now, for time-evolution, assume here too that the process is such that the kets transform (evolve) according to a unitary operator $T(t)$. Then for the observable v , there are the time-dependent average values $\langle v(t) \rangle$ given by $\langle v(t) \rangle = \langle P(t) | v(t) | P(t) \rangle$, where $|P(t)\rangle = T(t)|P\rangle$. (Note then that $v(t) = v_S(t)$, the usual operator from the "Schrodinger Representation".) It follows that in general there are two distinct time equations namely

$$\frac{d}{dt} |Pt\rangle = d_t |P\rangle$$

and

$$\frac{d}{dt} |P(t)\rangle = L |P\rangle$$

where the first refers to a change of frame and the second refers to temporal evolution within a frame. Note that the derivative of $T_f(t)$ (strong operator topology) is denoted by d_f , in analogy with Dirac's spatial displacement operators.

Perhaps the distinction between $T(t)$ and $T_f(t)$ will be considered unnecessary. But consider a process (e.g. one that is dissipative) where average energy values are evolving in time. Then, just as changes in distances may be invariant under certain changes of frame, so too might evolving changes in energies be invariant under certain changes of time frame. In that case $T_f(t)$ could preserve energy eigenvalues while $T(t)$ need not.

Finally, observe that the reasoning extends to the geometry of special relativity, where there is no position without time and no time without position. Indeed, for a permissible change of coordinates.

$$\underline{s} = (t, x) \rightarrow (t', x') = \underline{s}'$$

there would be the map

$$|P_{\underline{s}}\rangle \rightarrow |P_{\underline{s}'}\rangle \quad (3.9)$$

which would correspond to the change of frame. So average values as accessible to observers would transform according to

$$\langle v \rangle_{\underline{s}} \rightarrow \langle v \rangle_{\underline{s}'} \quad (3.10)$$

A second transformation would represent "space-time evolution" of the ket space, within the given frame. So here there would be the ket space transformation

$$|P(t_1, x_1)\rangle \rightarrow |P(t_2, x_2)\rangle \quad (3.11)$$

and the transformation of average values

$$\langle v \rangle(t_1, x_1) \rightarrow \langle v \rangle(t_2, x_2). \quad (3.12)$$

And as will be indicated in the next section, the space-time evolution can be related to Dirac's spinor equation for the electron.

4. Primary Relations and Secondary Determinations

The main purpose of this section is to distinguish two components in concrete relations, that is, a primary relativity and other secondary determinations. This distinction would seem to be relevant to certain fundamental issues. The structure of the section then will be as follows: Firstly, the problem is posed in the Hamiltonian context. Secondly, the possibility of primary relations and secondary determinations is discussed. Thirdly, the results are related to Dirac's equation. Fourthly, the form of Schrodinger's equation is re-considered. Finally, It is briefly indicated how operators should be a part of any quantum dynamical theory.

To present the problem, recall from analytical mechanics that a system may have a Lagrangian $L(q_j, \dot{q}_j, t)$ with generalized coordinates q_j and generalized momenta $p_j = \partial L / \partial \dot{q}_j$. In this classical theory, the Hamiltonian function is given by

$$H = \sum_j p_j \dot{q}_j - L$$

so that the canonical equations of motion are

$$\dot{q}_j = \partial H / \partial p_j$$

and

$$\dot{p}_j = -\partial H / \partial q_j.$$

These equations have been particularly useful in the investigation of planetary motion, and other classical systems. However, they have proved no less useful in dealing with processes typical to quantum mechanics. And while there are notable differences in the significance of the canonical equations in these respective domains, these differences also point to a similarity. For in both classical and quantum mechanics, the canonical equations have proven basic. Following on comments made in the Introduction, the problem here is to determine what role classical variables play in the theories of Schrodinger, Heisenberg and Dirac. With the question so posed there is no claim to generality. However, it would seem that the distinctions to be made could be relevant to the general problem of defining quantization.

One of the distinguishing features of small particle physics is the effectiveness with which the data continue to frustrate classical method. Laplace's goal is nowhere in sight and randomness would seem to have an objective meaning. So statistical method has found a place in modern physics, and quantum mechanics is noted for its indeterminacy relations and operator equations. Unlike classical physics, quantum mechanics can indicate spectra of possible values together with probabilities. And so,

while there is measurement in both classical and quantum mechanics, its function would seem to be quite different in each. Still, the fact that there is measurement in both theories is perhaps a clue to the link between them. If a measurement is made, it is a measurement of something. In other words, measurements are not merely sequences of numbers, but numbers correlated to terms, and terms which find their meaning in relations. So, for example, generalized momentum is not something somehow separate, unto itself, but is the term p_j defined by $p_j = \partial H / \partial q_j$. Therefore, it would seem that there is a complementarity in Dirac's quantum mechanics, a complementarity which is distinct from Bohr's meaning for the word. For in Dirac's theory, the quantum structure sets boundaries or norms from which concrete events cannot diverge non-randomly. On the other hand, the relational structure can provide definitions which are invariant under certain (canonical) transformations. Another way of saying this is that the operator structure of quantum mechanics provides probabilities of events which are defined by relations. So the energy operator H is to be clearly distinguished from the classically defined energy function. The operator has a spectrum of possible values for the function; and probabilities are introduced through the ket vectors.

The relations of Hamiltonian Mechanics therefore play a key role in quantum mechanics. And this leads to the question of relations in general. It would seem that the relations ultimately sought by physicists would be those which would survive all revisions in the scientific enterprise. In this regard both Newton's relations and Hamilton's original equations are problematic, for they occur in a provisional explanation and are not invariant under Lorentz transformations. Consider however the following. If it is true, for example, that A is twice the size of B , then A is proportional to B . The numerical ratio "twice" gives the proportion between A and B at given times under given conditions. But the proportion may change in accord with probabilities without altering the relation of proportionality (See [9], Sec. 16.2). And this would seem to be the general situation. In other words, in concrete relations there are two components to distinguish, namely primary relativity and other secondary determinations. The primary relativity is invariant while the secondary determinations include the numerical; and just what these are at any given time will be governed by probability distributions. So classical method reveals the primary relativity without the secondary determinations of concrete relations, that is, "provides an abstract relational field", say, for the positions and momenta of masses, but leaves to observation and, in the general case, to probabilities the determination of how many masses with what momenta at what position.

If the analysis is correct, then a corollary follows, namely, that the discovery of discrete spectra was not a failure as such in classical method. For classical relations say nothing of secondary determinations. Verified classical relations do not rule out the possibility of randomness in world process, for even in the investigation of highly

"regular" processes there is the proviso of "other things being equal". (That a meteor recently collided with Jupiter does not invalidate the laws of motion.) This is not to say that the discovery of discrete spectra was not significant. On the contrary, it was one of the turning points in Physics and was a catalyst for the discovery of new definitions and to the emergence of quantum mechanics. Still, classical method did not fail, but was confronted with processes which seemed to involve randomness. But if, for example, aggregates of electrons happen to involve randomness, it does not follow that velocity and momentum are not defined. What does follow is that their occurrence, in a sense which can be made rigorous, is *non-systematic* ([9], Sec. 2.3). This is in contrast to a *systematic* process like planetary motion which, among other things, allows for exact predictions. In both cases though there must be a primary relativity defining motion.

Now it is not the purpose of this section to focus on the meaning of *systematic* and *non-systematic*, but rather to further consider the significance, within quantum mechanics, of primary relations as distinct from secondary determinations. And perhaps one way of doing this is to review the problem of electron *spin* and the discovery of Dirac's spinor equation. Recall that when spectral lines of hydrogen were investigated with instruments of greater resolution, a discrepancy was found. Where single energy levels were predicted, pairs of levels were discovered (except for the case where the azimuthal quantum number $l = 0$). From this new statistical data the existence of a new variable was suggested, and it was Uhlenbeck and Goudsmit who proposed (1925) that an electron had a magnetic moment due to a motion about its own axis. Assuming the angular momentum to be $h/4\pi$ the corresponding magnetic moment would be $\mu = (he)/(4\pi mc)$. Therefore the energy in a uniform magnetic field $\mathcal{H}$, in the absence of all other forces, should be $\pm\mu\mathcal{H}$. And this correction term provided agreement with observed energy levels.

At this stage Pauli incorporated the experimental results into quantum mechanics by introducing his *spin matrices* σ_x , σ_y and σ_z , each of whose eigenvalues was to be the intrinsic angular momentum $\pm h/4\pi$. Since the probabilities (relative intensities) for each value seemed equal, a new energy operator H was given by

$$H = H_0 \otimes I_2 + I \otimes H_s$$

where H_0 was the original Schrodinger operator and H_s the spin energy operator. The new operator H acts on the tensor product of the original ket space with C^2 . For then if ψ_0 and $\psi_{\pm}$ are eigenvectors for H_0 and H_s respectively, with eigenvalues E_0 and $\pm\mu\mathcal{H}$, then

$$H(\psi_0 \otimes \psi_{\pm}) = (E_0 \pm \mu\mathcal{H})(\psi_0 \otimes \psi_{\pm})$$

as required. In a technical sense however, Pauli's results are incomplete. For in the first place the "spin variable" s is not a classical variable, but indexes the spectrum of possible spin energies. In other words, the quantity s would provide a range of secondary determinations, but is not defined in Pauli's theory. In the second place, the image of a spinning (microscopic) electron is probably non-verifiable. In the third place, Pauli's equations are not relativistically invariant. Still his inspired solution had the advantage of being formulated in accord with the basic tenets of quantum mechanics, and in such a way that proved useful for Dirac.

One of Dirac's questions was whether the "duplexity" phenomenon ([5], p. 254) could be accounted for by equations which were relativistically invariant. But Einstein's formula for energy is relativistically invariant. And if the equation is allowed its full relativity, two energies are obtained, E and $-E$. But Pauli's theory only considered positive energies and required a 2-fold ket space structure, one dimension for each spin $\pm h/4\pi$. Owing to the symmetry in E^2 of the relativistic equation, Dirac therefore hypothesized a 4-fold structure, two dimensions for each of $\pm E$.

It is well known that certain calculations allow Dirac's equation to be compared with Schrodinger's equation. And that by neglecting terms involving $1/c$, two terms remain from Dirac's equation which distinguish it from Schrodinger's, namely,

$$\frac{he}{4} \underline{\sigma} \cdot \underline{H} + \frac{he}{4\pi mc} \underline{\alpha} \cdot \underline{E}$$

where $\underline{H}$ is the magnetic field intensity and $\underline{E}$ is the electric field intensity. But again recalling known results, if an electron had a magnetic moment $\underline{m}$ and an electric moment $\underline{E}$, its additional energy in an electromagnetic field would be

$$\underline{m} \cdot \underline{H} + \underline{E} \cdot \underline{E}.$$

Because of this it is often said that Dirac's theory attributes the electron with a spin energy - without using a 19th century kind of assumption like a "spinning electron". Note however that spin itself requires classical definition. The operators $\underline{\sigma}$ and $\underline{\alpha}$ to be sure provide corrections to the range of possible energies. But just as velocity requires definition, so too does spin require definition, at least in order to survive as a defined term.

In addition to the correction terms for energy values, it is significant that as differential equations, the Schrodinger and Dirac equations are also different in type. This is not part of the main discussion, but since the topic does come up, a few comments will now be given on this issue. So where the Schrodinger time-dependent equation is of degree 2 in its space coordinates and of degree 1 in its time coordinate,

Dirac's equation (as originally required by Dirac) is of uniform degree 1 in all of the space and time coordinates. The usual justification for Dirac's form of the equation first assumes that the time-dependence must be of the first degree. Then by arguing that, in relativity, the space coordinates should hold no priority over the time coordinate, it is deduced that the energy equation must also be of the first degree in the space coordinates.

It would seem however that the uniform degree of Dirac's equation might also be obtained in another way. As mentioned at the end of Section 3 of this paper, space-time evolution can be related to Dirac's spinor equation. For the equation implicitly defines a transformation of the ket space as it evolves in space-time. But in the last paragraph of Section 3, the general form of such a transformation was given in equation (3.11). It could be possible therefore to introduce assumptions appropriate to the Hamiltonian context, so that Dirac's equation could be deduced from the general equation.

The attempt to give an exact solution to this problem must be left for a later paper. However, the following preliminary calculations may prove relevant. First suppose a 4-fold ket vector, as suggested by the relativistic formula for energy together with Pauli's results. Suppose also that as functions of space-time the ket vectors are at least once differentiable (this might be true, for example, if the space-time dependence of the energy average values is sufficiently smooth). Then by considering the first degree approximation to the general equation (3.11), one obtains a differential equation for the ket vectors, of uniform degree 1, and necessarily involving both 4×4 complex matrices and the operators $\partial/\partial x^\mu \otimes I_4$, $\mu = 0, 1, 2, 3$.

But it is known that when the electromagnetic potential is $A^\mu \equiv (\phi, \underline{A})$, the covariant form of the Dirac equation is

$$[\gamma^\mu (i\partial/\partial x^\mu - eA_\mu)] \Psi = m\Psi$$

where γ^μ , $\mu = 0, 1, 2, 3$ are certain 4×4 complex matrices. So it seems plausible that with certain further hypotheses in place, the first order approximation to equation 3.11 could reproduce Dirac's spinor equation.

Returning to the main discussion, our final remarks concern operators and the role they play in quantum dynamics. Consider then that every classical variable is implicitly defined by a correlation, or a unified set of correlations. On the other hand, places and times differ as a matter of fact. Places and times may be united by reference frames; and different geometries may offer different orderings for places and times. But unless there is to be a different physics for each place and time, there would seem to be no way of explaining why one place and time is different from another. If this is correct, it follows that position and time as such cannot be classical

variables. Moreover, this is consistent with the postulates of relativity. However, various velocities can be defined and their changes correlated and so have been candidates for classical variables. Therefore, any such defining (and so invariant) differential equation involving the space and time coordinates would be at least of the second degree. (This could relate then to the geodesic equations as given by Einstein, and to the much earlier equation $F = ma$ given by Newton). So while the primary (defining) differential relations would (with respect to these quantities) be of degree 2 or higher, the distribution relations for the secondary determinations (e.g. possibly Dirac's equation) could be of degree 1. Note that this difference would involve no contradiction, for the physical significance of the two types of relation are distinct, the first being definition and the second being statistical.

If a process being investigated involves randomness then it necessarily requires statistical method. But to a set of classes of classically defined events, statistical analysis associates respective probabilities, that is, statistical method determines states ([9], p. 81). Verification involves measurement which occurs at particular places and particular times, and so the states typically depend on space-time coordinates. Therefore, any quantum dynamics involving states which vary in space-time according to some law will necessarily include state equations defined by space-time operators. These equations within a frame would be determined by the particular type of process.

A change of frame however would have a different physical significance. A change of frame can (and often does) affect the distributions as accessible to observes, and so the states and state equations in a frame need not all remain invariant under all permissible transformations of a frame. For instance, the Dirac equation need not be invariant under all continuous transformations, but may be invariant only under the permissible transformations of Special Relativity. And this would not contradict the principles of General Relativity, for the invariance of General Relativity does not pertain to particular values, to secondary determinations, but to general defining relations.

To conclude : it seems that Quantum Dynamics would normatively determine a State Equation Theory. These equations would provide the space-time dependence of the statistical states, as verifiable within a frame. Without contradiction to General Relativity, these state equations and their state spaces need not be invariant under all permissible changes of frame, but could be invariant under a restricted group of permissible transformations (perhaps the Lorentz group) linked to the invariance properties of measurements accessible to observes. All permissible transformations however must by definition leave the defining primary relations invariant. In particular certain symmetry groups of the state evolution space-time dependent equations could be subgroups of symmetry groups of the defining primary relation

equations. This finally seems to indicate a possible line of enquiry whereby Quantum Dynamics and General Relativity might be correlated. There cannot be a unification in the sense of having a single system of equations and terms, for the theories each have their distinct physical significance. There could however be a unification corresponding to the unification in concrete situations of Primary Relations and Secondary Determinations. A network of verifiable classical relations would invariantly and so unambiguously define the physical terms, while the operator state equations and statistical structures would provide ideal norms accessible to observers from which actual relative frequencies of the classically defined terms could randomly, but not systematically diverge.

Acknowledgment

It is with pleasure that I take this opportunity to thank Ms. Noreen Dullea and Ms. Ann McCarthy (both of University College Cork) for taking my handwritten pages and producing this manuscript.

REFERENCES

1. R. ARENS AND D. BABBIT, Algebraic difficulties of preserving dynamical relations when forming quantum-mechanical operators, *Journal of Mathematical Physics*, **6** (1963) 1071-1075.
2. See J. BAGGOTT, *The Meaning of Quantum Theory*, Oxford University Press, Oxford, 1993.
3. L.E. BALLENTINE, *Quantum Mechanics*, Prentice Hall, New Jersey, 1990.
4. P.A.M. DIRAC, Quantum theory of the electron, *Roy. Soc. Proc. A* **117** (1928) 610-624.
5. P.A.M. DIRAC, *The Principles of Quantum Mechanics*, Fourth Edition, Oxford University Press, Oxford, 1962.
6. W. HEISENBERG, *The Physical Principles of the Quantum Theory*, University of Chicago Press, Chicago, 1930.
7. M. JAMMER, *The Conceptual Development of Quantum Mechanics*, McGraw-Hill Company, New York, 1966, p. 243.
8. S. KOCHEN AND E. SPECKER, The problem of hidden variables in quantum mechanics, *Journal of Mathematics and Mechanics*, **17** (1967) 29-87.
9. B. LONERGAN, *Collected Works*, vol. 3, University of Toronto Press, 1992.
10. A. MESSIAH, *Quantum Mechanics*, volume II, Fifth Printing, North Holland Pub. Co., Amsterdam, Theorem III.
11. E. SCHRODINGER, *Collected Papers on Wave Mechanics*, Blackie and Son Ltd., London and Glasgow, 1928.

President	: D.P. Gupta, Allahabad
Vice-Presidents	: M.S. Rangachari, Madras
	: S.L. Singh, Rishikesh
Secretary	: Mona Khare, Allahabad
Treasurer	: Shalini Srivastava, Allahabad
Editor (IJM)	: S.P. Singh, St. John's, NF, Canada
Librarian	: S.K. Srivastava, Varanasi

Other Members of the Council : P. Srivastava, Allahabad; T. Pati, Allahabad; U.K. Saxena, Allahabad; R.S. Sengar, Allahabad; Y.K. Srivastava, Varanasi; S.P. Khare, Allahabad.

Editorial Board

R.S. Sengar, Allahabad (Editor, Bulletin)
 Z.U. Ahmed, Aligarh
 R.K. Jain, Kanpur
 S.L. Singh, Haridwar
 V. Sitarahmaiah, Hyderabad
 T. Thrivikraman, Cochin

Committee for Publications : Pramila Srivastava (Chair); Bindhyachal Rai; Mona Khare; Y.K. Srivastava; Shalini Srivastava (Secretary for Publications)

The Society is registered under the Societies Registration Act XXI of 1860 and has reciprocity arrangement with various Societies including the American Mathematical Society, the Canadian Mathematical Society and the London Mathematical Society.