SLAC-PUB-2312
April 1979
(T/E)

AN OPERATIONAL ANALYSIS OF THE DOUBLE SLIT EXPERIMENT*

H. Pierre Noyes
Stanford Linear Accelerator Center
Stanford University, Stanford, California 94305

ABSTRACT

An attempt is made to reduce the discussion of the quantum mechanical interference patterns predicted to occur when a beam of particles of mass M and momentum p_0 is incident on two slits to macroscopic dimensional measurements of the sizes of slits, detectors, and collimators, and the macroscopic time measurements of whether or not detectors fire during macroscopic time intervals. By introducing detectors in the slits, the same apparatus yields statistical information about both single and double slit interference patterns, whose intensities add without interference, which can check the quantum mechanical prediction to arbitrarily high (statistical) precision. But discussion of the detectors themselves reveals that this scale invariant prediction (which depends only on $\hbar$ and M) can be carried through only if the detecting systems in the detectors have masses $m << M$. The existence of a smallest mass (empirically, the electron mass) or the limiting velocity c and the mass-energy relation (via the Wick-Yukaw mechanism) break the scale invariance of the theory. We conclude, as did Bohr and Rosenfeld in their analysis of the measurability of the electromagnetic field, that the existence of a smallest mass prevents an operational definition of the meaning of space time intervals of order h/m_ec or less.

(To be published by the Philosophy of Science Association in a volume on quantum mechanics edited by P. Suppes.)

* Work supported by the Department of Energy under contract number EY-76-C-03-0515.

I. INTRODUCTION

This paper illustrates the contention that quantum mechanical effects can be reduced to statistical distributions of yes-no events corresponding to counts in detectors. We claim that these distributions correspond to objectively distinguishable classes of events, and that their probabilities add without interference. In our view, noncommuting 'observables' and complementarity are confined to the model space employed by the quantum mechanical formalism in the calculation of these probability distributions; our 'objective reality' resides in the individual events, the macroscopic measurements which specify their approximate spatial and temporal relations, and the finite distributions which result from any particular test of the theory. All other aspects of the problem are discussed within the model space, and are posed under the restriction that the model space allows us to carry out the implied operations.

We hold that classical physics also employs a model space. This space differs subtly from the model space of quantum mechanics in that it allows limiting points of space and velocity (or time) to be defined independently in terms of abstract 'particles'. In contrast, we will discover in our specific context that the model space of quantum mechanics, also using abstract 'particles', easily allows the velocity limit to be taken (given enough space) but can define short distances only indirectly. Hence the space-time points in the model space do not have the same operational significance as those in the model space of classical physics, even though both are treated formally as mathematical continua.

We believe that one major source of confusion in discussions about the foundations of quantum mechanics is the often made but usually unstated assumption that the model space of classical physics is identical to 'real physical space'--whatever that means. This confusion becomes worse once it is also assumed that 'particles' are 'real physical objects' whose positions and velocities can be defined in that space. For us this last assumption is patently a metaphysical assertion rather than a scientific statement. Once that assertion is made, or even worse assumed without explicit recognition of the fact that this has been done, the fact that quantum mechanics employs a model space with different properties becomes distorted into a belief that quantum mechanics requires the 'real world' to have paradoxical, alogical, or even antilogical properties. All these problems can be avoided by not falling into that trap in the first place.

It has been forcefully pointed out to us that by confining the essential discussion to the model spaces of classical or quantum physics we disengage just where the philosophical problems become most interesting. Our excuse is that most of the problems we avoid are common to classical and quantum physics, at least from our own point of view. We are not willing to grant that physicists exist in the abstract, let alone the "Great Mathematician" of Jeans' cosmology. For us, either classical or quantum physics has the burden of showing that it is capable of a scientific retrodiction of the past which leads to physicists, using "historical" data compatible with the model in question; a brief effort of ours along these lines has been attempted (Noyes, Note 1). Most of this discussion does not involve explicitly the points at issue between classical and quantum physics. So far as we can see, any evolutionary or dialectical scheme which leads from physical cosmology through terrestrial biological development to the cultural evolution of physicists must, in current scientific context, use both conserved elements (matter in the broadest sense) and statistical elements on which selection can act in a unidirectional manner leading to increased complexity. We have argued elsewhere (Noyes, Note 2) that the absolute conservation laws for discrete quantum numbers coupled with the fixed-past uncertain-future interpretation of quantum mechanics (Noyes, 1975) form a more satisfactory scientific basis for such schemes than does classical continuum physics, but obviously cannot pursue this further here. In what follows we assume that, to requisite accuracy, physicists can measure "real" macroscopic space and time intervals, and agree whether counters fire or not whatever their basic beliefs about the models they correlate with these measurements.

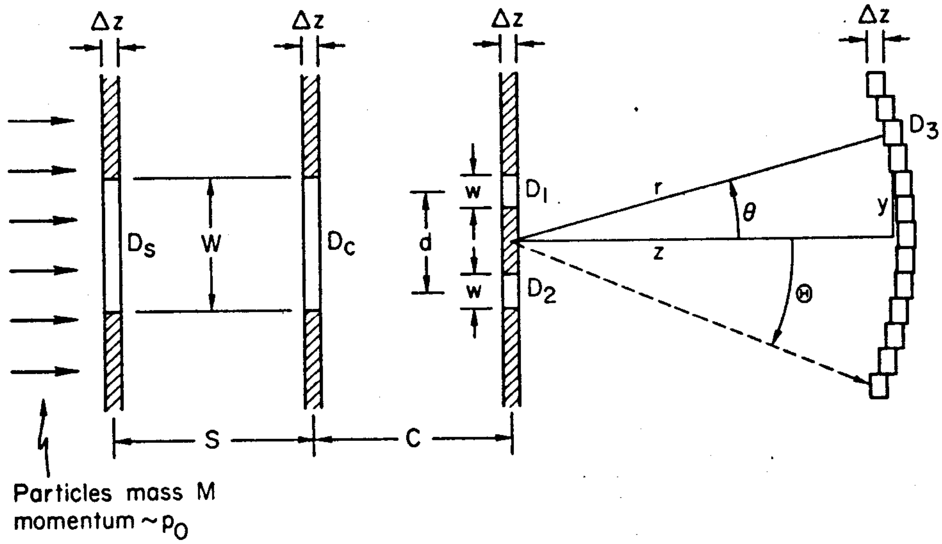
The paradigm we choose to illustrate our point of view is the double slit experiment. In order to bring out the objective nature of the predictions for this experiment, we employ particles of finite mass rather than 'light quanta' and assume that the two slits enclose particle detectors which sometimes detect the passage of a particle and sometimes do not. We then predict that three different statistical distributions will be detected in an array of particle detectors placed a sufficient distance behind the two slits. The usual double slit interference pattern corresponds to the case when neither of the slit detectors fires, while the two other patterns arise when one or the other of the slit detectors fires in addition to one of the detectors in the final array. These three distributions are predicted to add without interference, and to be objectively distinguishable by making use of appropriate time gates on the detectors and (delayed) coincidences between the detector signals.

Corrections to this initial prediction arise from nonuniform illumination of the slits and can be calculated from a detailed discussion of the way in which the beam of particles striking the slits is prepared. Additional corrections arise from the detection events in the detectors themselves and can be calculated from a model of the elementary process that activates the detector. We show that these corrections can, in principle, be made arbitrarily small, thus justifying the simpler treatment with which we start. We also discover that we do not actually have to perform coincidence measurements to distinguish the three distributions, because these distributions change in predictable ways as we change the density of the material in the detectors. This allows us to extrapolate our macroscopic thought experiment down to atomic dimensions and relate it to actual measurements of the interference effects observed between de Broglie waves. Some known limitations on this extrapolation are discussed in our concluding section.

II. FIRST APPROXIMATION TO THE PROBABILITY DISTRIBUTIONS

The model of the double slit experiment we consider is presented in Figure 1. We assume that particles of mass M and momentum close to some value p_0 (velocity $V_0 = p_0/M$) traveling in the z direction are produced by an accelerator or some other source. For simplicity we assume uniformity in the dimension perpendicular to the figure until we come to discuss corrections. To select particles with a known uncertainty in momentum we employ a collimator consisting of two slits (also containing particle detectors) of width W a distance S apart; the exit detector D_C is in a slit a distance C from the double slit arrangement we are studying. If we pick our zero of time as the time when a particle from the source is supposed (classically) to pass through one of the two slits, the momentum of this particle is ensured to be close to p_0 by gating the detector D_S closest to the source to be active at time $t \sim -(S + C)M/p_0$ and the detector D_C to be active at time $t \sim -CM/p_0$. We collect data only for cases in which D_S and D_C both fire while these time gates are open. Our third requirement is that the final detector array D_3 fire at time $t \sim +rM/p_0$, where $r = (y^2 + z^2)^{1/2}$ is the distance from the center of the double slit arrangement to the coordinates y,z describing the center of the detector in the array that fires. In addition to these three requirements (i.e., D_S, D_C, and D_3 must fire during the time intervals Δt set by the time gates centered around the two times given above) we also record whether or not D_1 or D_2 fires during a time interval Δt_0 centered on $t = 0$. We assume that the intensity is low enough so that corrections due to pileup (e.g., D_1 and D_2 both firing during this time interval) are not severe.

The prediction of the quantum mechanical model is that we can distinguish three different distributions in the detector array D_3. These are defined as: (I_0),



$$D_S \text{ fires at } t = -(S+C)\,M/p_0 \pm \frac{\Delta t}{2}$$

$$D_C \text{ fires at } t = -cM/p_0 \pm \frac{\Delta t}{2}$$

$$(I_0): D_3 \text{ fires at } t = rM/p_0 \pm \frac{\Delta t_3}{2}$$

$$(I_1): D_1 \text{ fires at } t = 0 \pm \frac{\Delta t_0}{2} \quad D_3 \text{ fires at } t = rM/p_0 \pm \frac{\Delta t_3}{2}$$

$$(I_2): D_2 \text{ fires at } t = 0 \pm \frac{\Delta t_0}{2} \quad D_3 \text{ fires at } t = rM/p_0 \pm \frac{\Delta t_3}{2}$$

4-77 3155A1

FIGURE 1. Center section of the geometrical arrangement for the experiment. The minimal dimension of the detectors is Δz, so w, d, and W are larger. Dimensions of the detectors in the x direction (perpendicular to the plane of the figure) are ℓ for the detectors in the slits and L for the detectors in the counters. The minimal time resolution Δt exceeds $M\Delta z/p_0$ (see text). If Δt_0 and Δt_3 are less than $\Delta t_{min} = M\Delta z/p_0$, events which would be allowed in the data set as valid are excluded. Optimal values for Δt_0 and Δt_3 depend on the intensity (i.e., number per unit time of valid events), since high intensity leads to an increase in the number of valid events, but also to a more rapid increase in the number of ambiguous events. The maximum value of Δt_3 which will ensure the maximum value of valid events allowed by the collimator is given in the text, but the optimization of this conjointly with the intensity is not performed.

those events in which neither D_1 nor D_2 fires when $t \sim 0$; (I_1), those events in which D_1 fires when $t \sim 0$; and (I_2), those events in which D_2 fires when $t \sim 0$. We assume that both the slit widths w and the distance between the two slits d are small compared to the distance to the detector array r. We further assume that the detector array D_3 subtends a finite angular aperture $\pm\Theta$, that the angles $\theta = \tan^{-1} y/z$ within this aperture are small compared to d/r, and that $r \gg \hbar/p_0$. We define three functions

$$f_S^{\pm}(\theta;w,k/p,d/r,\Theta) = \sin^2[pw(\theta \mp d/2r)/\hbar]/N_S(\theta \mp d/2r)^2 \tag{1}$$

$$f_d(\theta;w,d,\hbar/p,\Theta) = \sin^2[pw\theta/2\hbar]\cos^2[pd\theta/2\hbar]/N_d\theta^2 \tag{2}$$

with

$$N_S = \int_{-\Theta}^{\Theta} d\theta f_S(\theta) \; ; \qquad N_d = \int_{-\Theta}^{\Theta} d\theta f_d(\theta) \; . \tag{3}$$

Then the prediction for these three distributions, assuming a total of N data, is

$$I_0 = N(1 - e_1 - e_2) f_d(\theta;w,\hbar/p_0,d/r,\Theta) \tag{4}$$

$$I_1 = Ne_1 f_S^{+}(\theta;w,d,\hbar/p_0,\Theta) \tag{5}$$

$$I_2 = Ne_2 f_S^{-}(\theta;w,d,\hbar/p_0,\Theta) \; . \tag{6}$$

These are, of course, the usual double and single slit diffraction patterns for wavelength $\hbar/p_0$ in the small angle approximation. The parameters e_1 and e_2 are simply the fraction of the total number of data for which D_1 or D_2 fires. The prediction is statistical; for large N it will be approached in the sense of the law of large numbers. The fractions e_1 and e_2 are predicted to be proportional to the detector efficiencies, as will be discussed in Section IV.

We note that this single experimental arrangement suffices to provide data exhibiting both single slit and double slit interference patterns. This appears to conflict with loose statements about complementarity, such as "an experiment which allows us to determine which slit the particle passes through is complementary to an experiment which yields the double slit pattern." We see that if by experiment we mean a single event, it is correct to say that the three possibilities belong to disjoint and objectively distinguishable classes of events and hence in that restricted sense are complementary. However, if the term 'experiment' means, as it usually does in physics, a system capable of producing numbers which can be compared with the theoretical prediction, the statement is palpably false.

In order to test the prediction, we must accumulate enough data to compare the prediction, at some confidence level, with the large number limit. If it happens that e_1 and e_2 are very small, we will get a reasonable check on the prediction of I_0 for a smaller value of N than we need to check the predictions for I_1 and I_2; but if $e_1 + e_2$ is close to unity, the reverse will be true. Since e_1 and e_2 are adjustable by changing the density of the detecting systems in the detectors, we can obviously choose conditions such that our confidence in the test of all three predictions reaches roughly the same level for the same value of N. Further, for the same reason, if we collect data for various values of e_1 and e_2 we can verify the prediction without even making use of the time tagging of the three distributions. Hence, allowed this freedom, the experiment can be 'overdetermined'; this is important from an experimental point of view, since it allows us a check on some types of systematic error.

We now derive the approximate results quoted above. Concentrating first on states of a single momentum p traveling in the +z direction, if there were no boundaries, the wave function would be $\psi(z,t) = \exp i(pz/\hbar - p^2t/2M\hbar)$. The corresponding flux is $j_s = (\hbar/2Mi)(\psi^*\partial\psi/\partial z - \psi\partial\psi^*/\partial z) = p/M = v$ particles per unit 'area' per unit time; in our two-dimensional problem, 'area' is unit length perpendicular to the z direction, while in three dimensions it is area in the xy plane. We assume that our source and collimator are such that, to a first approximation, our two slits are illuminated with a plane wave of this type and that the detector does not significantly alter the momentum of the individual particles in this beam. We also assume that the randomness in phase allowed by the first two time gates is uniform enough so that only the time-average value of ψ matters. All of these approximations will be justified subsequently. They allow us to look for a solution of the time-independent Schrödinger equation $(\partial^2/\partial z^2 + \partial^2/\partial y^2 + p^2/\hbar^2)\psi(y,z) = 0$ in the half space $0 \leq z < \infty$, $-\infty < y < \infty$ which vanishes on the line $z = 0$ except for two uniform distributions of equal magnitude for y lying in the two slit intervals $(d + w)/2 > y > (d - w)/2$ and $-(d - w)/2 > y > -(d + w)/2$.

Since we also need the single slit pattern, we first consider a single slit of width w centered at $y = 0$, and then show how the desired result can be derived from that. Our boundary condition for this simpler problem is that $\psi_w(y,0) = \frac{1}{w}[\theta(y + w/2) - \theta(y - w/2)]$. The solution must be of the form

$$\psi_w(y,z) = \int_{-\pi/2}^{\pi/2} d\phi f_w(\phi) e^{ip\cos\phi z/\hbar} e^{ip\sin\phi y/\hbar} \quad . \tag{7}$$

Hence, by taking the Fourier transform of the boundary condition we find that

$$f_w(\phi) = (1/\pi w)\sin(pw\sin\phi/2\hbar)\cot\phi \tag{8}$$

or, letting $\xi = \sin\phi$, that the single slit wave function is

$$\psi_w(y,z) = \frac{1}{\pi w}\int_{-1}^{1} \frac{d\xi}{\xi}\sin(pw\xi/2\hbar) e^{ipz\sqrt{1-\xi^2}/\hbar} e^{ipy\xi/\hbar} \tag{9}$$

or, in polar coordinates,

$$\psi_w(r \sin \theta, r \cos \theta) = \frac{1}{\pi w}\, d\, \frac{\sin(pw\xi/2\hbar)}{\xi} \cdot e^{i(pr/\hbar)(\sqrt{1-\xi^2} \cos \theta + \xi \sin \theta)}. \tag{10}$$

If we had used a rectangular aperture of finite length ℓ in the x direction as well as finite width w, we could clearly have obtained a similar result using two angles for the resolution of $\underline{p}$ in the x, y, and z directions, and ended up with a somewhat more complicated double integral representing $\psi_{w\ell}$.

In order to obtain the desired result we now make use of our assumption that the detector array D_3 is so far away from the slits that $pr/\hbar >> 1$. We can then evaluate the integral by the method of stationary phase, as explained, for example, by Jeffries and Jeffries (1950), obtaining

$$\psi_w \rightarrow \left[\frac{2\pi\hbar(1 - \xi_0)^2}{pr}\right]^{1/2} \frac{\sin(pw\xi_0/2\hbar)}{w\xi_0}\, e^{i(pr/\hbar-\pi/4)} + O(\hbar/pr) \tag{11}$$

where

$$\xi_0 = \sin \theta . \tag{12}$$

If detector 1 fires, we know that the particle passed through the upper slit. We can therefore obtain the wave function for this case by using the method just developed with our boundary condition displaced upward by d/2, rather than being centered at y = 0. The result is again Equation 11 with $\xi_0 \rightarrow \xi_1$, defined by

$$(\xi_1)^2 = \frac{(\sin \theta - d/2r)^2}{\cos^2\theta + (\sin \theta - d/2r)^2} . \tag{13}$$

If detector 2 fires the result is obviously given in the same way with $\xi_0 \rightarrow \xi_2$ and

$$(\xi_2)^2 = \frac{(\sin \theta + d/2r)^2}{\cos^2\theta + (\sin \theta + d/2r)^2} . \tag{14}$$

If neither 1 nor 2 fires, our boundary condition spreads the same intensity over the two slits, and the two contributions add coherently, yielding the wave function

$$\psi_{2w,d}(r\cos\theta, r\sin\theta) = \frac{1}{\pi}\int_{-1}^{1} d\,\frac{\sin(pw\xi/2\hbar)}{w\xi}\cos(pd\xi/2\hbar) \cdot e^{i(pr/\hbar)(\sqrt{1-\xi^2}\cos\theta + \xi\sin\theta)}\ . \tag{15}$$

For this wave function the stationary phase is given by $\xi = \sin\theta$. Using these three wave functions to calculate the flux through D_3 and neglecting θ^2 compared to θ, we finally obtain the three results given in Equations 1-3, as was to be proved.

III. CORRECTIONS DUE TO THE STRUCTURE OF THE COLLIMATOR AND THE SOURCE

This section and the next deal with what is customarily called 'preparation of the initial state' and 'measurement'. Elementary treatments simply assume, as we have done in the last section, that it is possible to start with a beam of particles with precisely defined momentum $\underline{p}_0$, and that the measurement corresponds to the detection events in D_3 assumed proportional to $\psi^*\psi$ integrated over the space-time volume defined by that detector and its time gate. I made these assumptions myself in my own derivation of the scattering formalism starting from quantum mechanical particles and particle detectors (Noyes, 1976). Yet this definition of measurement as proportional to $\psi^*\psi$ is hard to relate to the abstract formulation of quantum mechanics as given by von Neumann, as will be readily appreciated after reading the discussion of this problem given by Nancy Cartwright in another article prepared for this volume. The assumption that the initial momentum is well defined is known to be suspect because of the uncertainty principle. Qualitatively we can argue that with all past time available to prepare the beam, both its energy and direction can be defined as sharply as we wish. More precise treatment, for example, that by Goldberger and Watson (1964), estimates the size of the wave packet allowed by the uncertainty principle, and shows that this introduces no problems for the particle beams in actual use, but implicitly assumes that some function $f(p - p_0)$ used to define the wave packet centered on p_0 can itself be known to arbitrarily high precision. We have gone to considerable trouble to show that this assumption can indeed be justified using only particle detectors to measure momentum distributions.

Because of the complexity of the argument, I have found it useful to break my analysis into three parts. First, I assume that the geometrical configuration of the two detectors in the collimators and the time gates applied to them define the momenta which emerge and, under the assumption that the detection events are uniformly distributed over these space-time volumes, calculate the probability distribution of the momenta of the particles thus selected. Then I ask under what restrictions the quantum mechanical model allows us to treat this calculation as a first approximation, with bounded corrections that can be made arbitrarily small, still assuming that the detection events themselves within these geometrical and temporal limits can be treated as punctiform. Finally, I discuss assumptions which suffice to allow the quantum mechanical detection events to be approximated in this manner.

Returning to Figure 1 and the statement of the conditions for data collection, we see that the firing of the two detectors D_S and D_C guarantees (except for the accidental coincidences that will occur if the flux of particles is too intense) that the probability of the particle being in the two space-time volumes

$$-\frac{L}{2} < x_S < \frac{L}{2}, \qquad -\frac{W}{2} < y_S < \frac{W}{2},$$

$$-S - C - \frac{\Delta z}{2} < z_S < -S - C + \frac{\Delta z}{2}$$

$$-\frac{M(S + C)}{p_0} - \frac{\Delta t}{2} < t_S < -\frac{M(S + C)}{p_0} + \frac{\Delta t}{2} \qquad (16)$$

and

$$-\frac{L}{2} < x_C < \frac{L}{2}, \qquad -\frac{W}{2} < y_C < \frac{W}{2},$$

$$-C - \frac{\Delta z}{2} < z_C < -C + \frac{\Delta z}{2}$$

$$-\frac{MC}{p_0} - \frac{\Delta t}{2} < t_C < -\frac{MC}{p_0} + \frac{\Delta t}{2} \qquad (17)$$

is unity; this implies that the baffles around the slits and the interior walls of the volumes containing the detectors are 'perfectly absorbing', or in other words that any particle that fails to meet these boundary conditions does not reach the two slits under study or return to either detector in the collimator. Corrections due to the failure of this approximation constantly harass the experimental physicist but do not concern us in this thought experiment.

For simplicity we will ignore the lateral (x,y) dimensions and calculate the probability distribution in $p_z = M(z_C - z_S)/(t_C - t_S)$ implied by these boundary conditions. We further assume, to begin with, that the probability of finding a particle within the specified ranges of time and distance is uniform. The quantity of physical interest, the probability distribution in the momentum p_z, depends only on the intervals $z_C - z_S$, and $t_C - t_S$. This fact allows us to define two new variables $z = z_C - z_S - S$ and $t = t_C - t_S - MS/p_0$, and integrate over the uniform distributions in t_S and z_S to obtain

$$dP(z,t) = \Theta\left(\frac{z}{\Delta z}\right)\Theta\left(\frac{t}{\Delta t}\right)\left(1 - \frac{|z|}{\Delta z}\right)\left(1 - \frac{|t|}{\Delta t}\right)\frac{dz\,dt}{\Delta z \Delta t} \qquad (18)$$

where

$$\Theta\left(\frac{x}{\Delta x}\right) \equiv \theta\left(\frac{x}{\Delta x} + 1\right) - \theta\left(\frac{x}{\Delta x} - 1\right). \qquad (19)$$

The final step is then to change variables to p_z and some second variable which can be integrated out. In order to bring out more clearly the physical magnitudes involved, we introduce the dimensionless parameters and variables

$$\varepsilon = p_0 \Delta t / MS, \qquad f = \Delta z / \varepsilon S,$$
$$q = (p_z - p_0)/\varepsilon p_0, \quad \xi = \Delta t / t \ . \tag{20}$$

The time a particle of velocity p_0/M takes to cross a length Δz is $M\Delta z/p_0$. If the time resolution Δt were less than this value, we could use the timing device to reduce the spatial resolution Δz. We therefore assume that $\Delta t \geq M\Delta z/p_0$, which implies that f is less than or equal to unity. The momentum p_0 and the mass M are specified in advance. As we will see in the next section, the thickness of the detector Δz depends on what we are willing to assume about its construction. However, the length of the collimator S is still at our disposal, which allows us to make ε as small as we like if we are willing to go to a large enough apparatus for preparing the beam. For similar obvious geometrical reasons, the lateral spread in momentum is proportional to W/S and L/S, and these two dimensions (if bigger than the limiting dimension Δz) are also at our disposal. We can therefore guarantee that the momentum p is well defined in direction as well as in magnitude, if there are no additional wave mechanical restrictions.

Making use of these definitions, we can change variables from z,t to q,ξ and using

$$\frac{z}{\Delta z} = \frac{1}{f}\,[q(1 + \varepsilon\xi) - \xi] \tag{21}$$

find that the probability distribution in q is

$$\begin{aligned} dP(q) &= \frac{dq}{f}\int_{-1}^{1} d\xi (1 + \varepsilon\xi)(1 - |\xi|)(\xi)(1 - |\,\frac{1}{f}\,[q(1 + \varepsilon\xi) - \xi|) \\ &\quad \cdot \left(\frac{1}{f}\,[q(1 + \varepsilon\xi) - \xi]\right) \qquad (22) \\ &= [\theta(q + \frac{1 + f}{1 + \varepsilon}) - \theta(q - \frac{1 + f}{1 - \varepsilon})]F(q)dq \ . \end{aligned}$$

For $f \approx 1$ (optimum time resolution) and small ε

$$F(q) \approx \frac{2}{3} - \frac{3}{2}|q| + 2q^2 - |q|^3 \qquad |q| < 1$$
$$\approx \frac{1}{6}(2 - |q|)^3 \qquad 1 < |q| < 2 . \tag{23}$$

The net effect of the calculation is to show that the collimator does indeed produce a momentum distribution peaked about p_0 with the limits

$$p_0 - \varepsilon p_0 \frac{1+f}{1+\varepsilon} < p_z < p_0 + \varepsilon p_0 \frac{1+f}{1-\varepsilon} . \tag{24}$$

If the source illuminating the collimator does not provide an absolutely uniform distribution of momenta p_z in this range, the assumption that the spatial and temporal distributions within the counters are uniform is not justified due to the correlations between them introduced by this fact. However, the classical calculation is easy to correct if we know the source distribution $P_S(p_z)$, the obvious result being $P_S(p_z)F((p_z - p_0)/\varepsilon p_0)dp_z/\bar{P}$, where $\bar{P}$ is the integral of $P_S F$ over the interval in Equation 24. Since we can construct collimators of arbitrarily high quality by making them long enough, we can measure the probability distribution P_S to any requisite accuracy. This is part of what any experimental physicist would do to calibrate the equipment used in the experiment and so is usually assumed without question. The difficulty from the point of view of this paper is that the calibration so far described uses classical physics and hence cannot be invoked by us if we aim at a self-consistent operational analysis. Even careful treatments such as that of Goldberger and Watson (1964) content themselves with the assumption that the source provides a wave packet $(P_S(p_z))^{1/2} = f_S(p_z)$ whose size can be estimated from the size of the source and with showing that quantum mechanical corrections due to 'wave packet spreading' are, under current conditions, unimportant until the apparatus extends from the earth to the moon or farther. We try to provide below a somewhat more careful analysis. However, to do this we find that it is necessary to invoke some background material which, though familiar to physicists, cannot be expected to form part of the background philosophers and mathematicians invariably bring to the study of the foundations of quantum mechanics.

In classical wave motion, the maximum amplitude A of a monochromatic wave $A \sin [(z - vt)/\lambda\!\!\!^{-} + \delta]$ and the phase δ are physically measurable quantities. For instance, for a vibrating string, A is the maximum displacement of the string from its equilibrium configuration at position z along the string, and δ can be determined, if the motion is slow enough, by measuring the displacement as a function of time relative to some clock. However, for sound waves in which the amplitude is the average displacement of the air molecules from their equilibrium position, or for light waves in which it is the maximum strength of the electromagnetic field vector, neither A nor δ is experimentally accessible from a practical point of view. Usually for any pitch, the ear, and for all colors, the eye, respond not to the displacement (whose average value is obviously zero) but to the <u>intensity</u> I, which is the square of the displacement averaged over one cycle; clearly, $I = 1/2\ A^2$. Therefore, it is convenient to represent the wave by a complex function $I^{1/2} \exp i[(z - vt)/\lambda\!\!\!^{-} + \delta] = \psi$, whose absolute square $\psi^*\psi = I$ performs the averaging automatically. Clearly, if we measure only the intensity of a <u>single</u> wave, the phase δ remains unobservable.

At this point, one might well ask how the wave character of either sound or light can be demonstrated. For instance, the source of light might be a candle flame containing $\sim 10^{18}$ individually excited molecules emitting wave trains randomly phased with respect to each other. If we bring beams from two different candles together, there is no interference; the intensities simply add. But if we use a single candle to illuminate a single slit (through a collimator or lens, or far enough away so that the candle can be geometrically approximated by a point source) and look on a far screen, the path length through different points across the width of the slit to the point on the screen can differ by a wavelength; we then can observe the single slit diffraction pattern due to the interference between elements of a single wave train coming from a single atomic source. If the sources are effectively at a point, the intensities on the screen add to give the visible pattern predicted for each separate wave train. The calculation is equivalent to the derivation we gave in Section II. If we use two slits illuminated by a single source, we can observe the double slit pattern. We see that the quantum mechanical calculation for a single particle is like the classical calculation for a single wave train of unknown phase. All that our detectors D_S, D_C, and D_0 do in the particulate case is to ensure, by our time tagging, that we are dealing with a single wave train with the further simplification, if $\Delta z/S$ is small, of a well-defined wavelength. However, if either D_1 or D_2 fires in addition, we destroy any possibility of phase coherence between the two slits, and obtain only the appropriate single slit pattern. The quantum mechanical character of the calculation is thus restricted to the discrete (yes-no) character of the individual particulate detection events, as claimed in our introduction.

The point I hope to have brought out by this digression is that for quantum mechanics the absolute phase of the wave is in principle unknowable, in contrast to the classical situation where the 'random phase approximation' is simply a practical method for dealing with physical situations in which a well-defined quantity happens to be beyond the reach of direct experimental measurement. In my previous discussion of the scattering process (Noyes, 1976), I made use of this principle to derive the conventional scattering formalism by noting that if the absolute coordinates of the positions where scattering takes place entered the theory, they would have to enter the theory as 'quantum mechanical hidden variables' unless the sum over all such processes is performed in a particular way. This enabled me to separate the kinematic (i.e., descriptive) aspect of scattering theory from the dynamical aspect (i.e., the calculation of scattering amplitudes, including relative but not absolute phases). But in that treatment I still assumed that an incident beam of precisely defined momentum could be used as a starting point, an assumption which the discussion presented here is intended to justify.

As we will discuss in more detail in the next section, the detector can be so designed (under restrictions which turn out to be conceptually important) that the detection event does not in itself significantly perturb the incident wave's momentum. Thus if we know $f_S(p_z)$, but not its phase, we can assume that the scattering event that initiates the detection in D_S results in a wave function, for $t > t_S$, which can be written as

$$\psi_C(z,t) = \int_{-\infty}^{\infty} dp e^{i\delta(p)} f_S(p) e^{i[(z-z_S)p/\hbar-(t-t_S)p^2/2M\hbar]} . \tag{25}$$

Under our assumption that this process is incoherent with respect to the source, the phase δ can have any value, and, as just discussed, the parameters z_S and t_S can be known only to the extent that they lie within the ranges given by Equation 16. To justify this assumption a fortiori, we must be able to show that by making this posit the collimator can be used to measure the probability distribution $P_S(p) = f_S^2(p)$ independent of any knowledge of $\delta(p)$; clearly we are assuming both f_S and δ are real and that f_S is appropriately bounded. As in the classical calculation, we assume that all values of z_S and t_S are equally probable within the specified ranges.

If we are correct in assuming that this is the wave function that results from the first detection event, we can then use it to calculate the probability of the second detector (D_C) firing during the gated time interval as proportional to $\psi_C{}^*\psi_C$ when z and t are in the ranges specified by Equation 17. Hence,

$$dP(x_S,t_S,x_C,t_C) \alpha dx_S dt_S dx_C dz_C$$
$$\cdot \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dk f_S(p) f_S(k) e^{i(\delta(p)-\delta(k))}$$
$$\cdot e^{i[(z_C-z_S)(p-k)/\hbar-(t_C-t_S)(p^2-k^2)^2/2M\hbar]} . \tag{26}$$

As in the classical calculation, we can immediately integrate over t_S and z_S to obtain

$$dp(z,t) = \frac{dzdt}{\Delta z \Delta t}\left(1 - \frac{|z|}{\Delta z}\right)\left(1 - \frac{|t|}{\Delta t}\right)\Theta\left(\frac{z}{\Delta z}\right)\Theta\left(\frac{t}{\Delta t}\right) \cdot \int_{-\infty}^{\infty}\delta p \int_{-\infty}^{\infty} dk\ f_S(f)f_S(k) \cdot e^{i(\delta(p)-\delta(k))} e^{i\delta(p-k)[1 - \frac{(p+k)}{2p_0}]/\hbar} \cdot e^{i[z(p-k)/\hbar - t(p^2-k^2)/2M\hbar]}. \tag{27}$$

But this result can be evaluated by the method of stationary phase, and we recover immediately the classical result provided only that $S >> \hbar/p_0$ or, in words, that the distance between the two counters in the collimator is a large number of de Broglie wavelengths. Thus, provided the detectors themselves act as point detectors, when integrated over the active volumes and time gates, we have given a rigorous justification of the classical approximation and have proved that we can measure momentum in a quantum mechanical system to arbitrarily high precision.

The extension of the discussion to three dimensions is straightforward and contains no new conceptual points so far as the time dependence goes, so we will not attempt it here. With regard to the lateral dimensions, physicists will recognize that, in order to avoid diffraction effects, the dimensions W and L will also have to be large when measured in de Broglie wavelengths. In a thought experiment we can always ensure this while still keeping W/S and L/S small simply by scaling everything up in all three dimensions; in any case, the calculation of the <u>intensity</u> of the diffraction patterns across the faces of the counters is a straightforward problem in physical optics, for which standard methods can be used. These can then be used to weight the integrals needed in the probability calculation. Optimization of the counting rates so as to reduce ambiguities due to pileup, now that we have reduced the problem to a classical one, is again a familiar experimental task and provides no conceptual problems. All that remains is to show that, given an initial momentum distribution whose magnitude is knowable to arbitrarily high accuracy, we can indeed construct detectors that will fit into the slits of width w whose diffraction patterns (single and double) we are trying to measure.

IV. CORRECTIONS DUE TO THE STRUCTURE OF THE DETECTOR

The detector we assume to complete our design consists of particles of mass $m \ll M$, struck by the massive particles under study and recoiling. If the recoil energy $p^2/2m$ is greater than some critical energy ε_0, the recoil particle can activate a secondary process, which is amplified by the detector and eventually leads to the signal that we have referred to so far as the firing of the detector. In practice, the energy ε_0 could be the ionization energy of the atoms in the detector, and the amplification process could be an electron and ion cascade, as in a Geiger tube; alternatively, the ionization could produce a photon when the atom that has been ionized recombines, and the amplification could be provided by a photosensitive surface and a photomultiplier tube. But for our purposes all we need is the kinematics of the initiating event, the threshold energy, and the cross section πa_0^2 which each initial event has with respect to the incident beam.

Given this model for the detector, our first step is to calculate the kinematics of the initial scattering process. Assuming that the particles m are free, we can calculate the angular dependence of the momentum and energy of the (scattered) beam and (recoil) detector particles from the conservation of momentum and energy by solving the three equations

$$\begin{aligned} p \sin \phi &= P \sin \theta \\ p \cos \phi + P \cos \theta &= P_0 \\ \frac{p^2}{2m} + \frac{P^2}{2M} &= \frac{P_0^2}{2M} \end{aligned} \tag{28}$$

with the result

$$P_{(\theta)} = \frac{P_0 \cos\theta + \sqrt{\frac{m^2}{M^2} - \sin^2\theta}}{1 + m/M} \tag{29}$$

$$P^2_{(\theta)} = \frac{2mP_0^2\left[\sin^2\theta + \frac{m}{M} - \cos\theta\sqrt{\frac{m^2}{M^2} - \sin^2\theta}\;\right]}{M(1 + m/M)^2} .$$

We see immediately that the maximum angle of scattering of the beam particles is $\theta_x = \sin^{-1} m/M$ and hence that simply by using detector particles of small enough mass we can guarantee that the detector does not introduce significant corrections into our analysis so far as angular distributions go. Since the final detector array has angular aperture Θ, all we need require is that $m/M << \Theta$.

If the density of detector particles in the volume of the detector V is ρ grams/cm^3, the number of particles is $\rho V/m$, and the cross-sectional area they present to the beam is $\pi\rho V a_0^2/m$. Since the volume of the detector is $w\ell\Delta z$, and the area presented to the beam is $w\ell$, the fraction of the time one beam particle incident on this area will activate the counter is $\pi\rho a_0^2\Delta z/m$. If $\rho = \rho_1$, the density of detector particles in D_1, this fraction is simply e_1, i.e., the fraction of the time that, on the average, the detector will fire when (as our procedure for data collection ensures) exposed to one particle passing through the system. Clearly we must make the density low enough so that $e_1 + e_2 < 1$ for the experiment to work as designed, which also ensures that the corrections due to the beam striking more than one particle within the detector will be small. Thus our detector design ensures the conditions assumed in our initial discussion, and justifies the assertion that we can get equivalent results simply by using particles of known mass m, density ρ, and cross section a_0 in the slits without actually recording whether or not individual data points in the final distribution correspond one by one to the three distributions predicted.

A number of corrections are needed before we can directly confront theory with experiment even under these restrictions. The kinematical equations show us that the momentum of the detector particles varies from 0 when the particle of mass M continues to move in the forward direction ($\sin\theta = 0$) to $p_x = (m/M)[2m/M(1 + m/M)]p_0^{1/2}$ when $\sin\theta = m/M = \sin\theta_x$. In order to activate the detector at all, we must clearly require that $(mp_0/M)^2/M(1 + m/M) > \varepsilon_0$, but there will also be a critical angle θ_C, defined by

$$p_0\left[\sin^2\theta_C + \frac{m^2}{M^2} - \cos\theta_C\sqrt{\frac{m^2}{M^2} - \sin^2\theta_C}\right] = M(1 + \frac{m}{M})\varepsilon_0 , \tag{30}$$

within which the detector will not be activated. If the differential cross section for scattering is $\sigma(\theta)$, with $\pi a_0^2 = \int_0^{\theta_x} \sigma(\theta)d\theta$, then the fraction of the time the detector will not be activated even though scattering occurs is

$$f(m,a_0) = \int_0^{\theta_C} \sigma(\theta)d\theta/\pi a_0^2 . \tag{31}$$

But, as we saw in our last section, the scattering destroys the phase coherence necessary for the double slit interference pattern even though the scattering is not directly observed (i.e., does not activate the detector). Consequently, we must modify our initial prediction to read

$$I_0 = N[(1 - e_1 - e_2)f_d + f(m,\theta_0)(e_1 f_w^+ + e_2 f_w^-)]$$
$$I_1 = Ne_1(1 - f(m,a_0))f_w^+ \qquad (32)$$
$$I_2 = Ne_2(1 - f(m,a_0))f_w^- .$$

This makes the comparison with experiment slightly more complicated in practice but does not change anything in principle, except that the correction to I_0 blurs the double slit pattern by including a single slit contamination from those particles which actually scattered from a detector particle in one of the two slits but were not demonstrated to do so because of this detection inefficiency.

In using this expression, we should note that the fractions e_1 and e_2 which occur are the <u>theoretical</u> expressions

$$e_1 = \pi\rho_1 a_0^2 \Delta z/m , \qquad e_2 = \pi\rho_2 a_0^2 \Delta z/m \qquad (33)$$

for any scattering event and not the experimental fractions actually measured in any particular experimental run. The statistical prediction for the experimental fractions (i.e., the number of times the detectors D_1 and D_2 fire compared to the total number of events in which D_S, D_C, and D_3 fire) is obviously

$$e_1^{exp} = e_1(1 - f(m,a_0)), \quad e_2^{exp} = e_2(1 - f(m,a_0)) . \qquad (34)$$

Comparison between these predictions and the experimental results therefore gives us an internal check on the consistency of our assumptions about the detectors; failure of this check would point to systematic error in the overall setup and suggest that we make independent measurements of ρ_i, a_0, $f(m,a_0)$, the momentum distribution emerging from the collimator or arriving at the detector array D_3, etc., in order to resolve them. Since we have already demonstrated that we can, in principle, construct collimators that allow us to measure momenta to arbitrarily high precision in both magnitude and direction, these experimental checks can all be carried out. They are part of the procedure in any carefully designed experiment, and will not concern us further here.

So far we have not discussed the length of the time gate Δt_3 during which we need to keep the counter D_3 open. We saw in the last section that the collimator does not provide a precisely monochromatic (single momentum) beam but introduces a spread in momentum given by Equation 24. This means that the slowest particles will arrive at D_3 still farther behind the center of the pulse than they leave the collimator and the fastest particles still farther ahead of it. This is a classical 'wave packet spreading' produced by our finite spatial and temporal resolutions and has nothing to do with the quantum mechanical effect of the same name. So long as our time gate is centered on $t = rM/p_0$ we will catch the central portion of the pulse, so using the time gate that will catch <u>all</u> particles that have passed our collimation conditions and the two slits can be simply a matter of optimization of design to increase the counting rate of useful events. However, since this time is necessarily longer than Δt, and hence determines the probability that an event we assign to a particular $D_S + D_C$ trigger is in fact correctly assigned, we need this number to optimize the overall intensity of the incident beam. This number is

$$\Delta t_3 = \left(\frac{S + C + r}{S + \Delta z} \right)\left(\Delta t + \frac{M\Delta z}{p_0}\right) . \qquad (35)$$

For similar reasons we should set the time gate Δt_0 on the detectors D_1 and D_2 to be long enough to accommodate the full pulse defined by the collimator. If we make it shorter, we can also make Δt_3 shorter; this suggests that we could simplify the overall design by omitting the counter D_C and using the time gate on the array D_3 to define the momentum. We have not done this, because we thought it simpler conceptually to separate the action of the collimator from the rest of the experiment, rather than present an integrated and optimized design. Another correction, due to the fact that the detection process (even for $\theta < \theta_C$) robs the beam particles of energy and hence slows them down compared to the particles in the double slit pattern, should also be included for completeness. However, this correction, which in principle would require us to use an angle-dependent time gate on D_3 correlated with the firing of D_1 or D_2, and with a corresponding modification of $f(m,a_0)$, would take us into the problem of detecting the time structure in the counts at D_3. It also would give us still another experimental handle on the discrimination between the three patterns I_0, I_1, and I_2; but since we started only with the problem of measuring intensity distributions, we will not pursue the analysis further.

Our treatment so far is not explicitly wave mechanical. We now show that, provided the range R of forces between the beam and detector particles is small compared to the dimensions of the detector, the same approximations suffice. We introduce the relative coordinate $\underline{x} = \underline{r}_M - \underline{r}_m$ between the detector and beam particles, with conjugate momentum $\underline{k} = (mP - Mp)/(M + m)$ and center of mass coordinate $\underline{X} = (M\underline{r}_M + m\underline{r}_m)/(M + m)$ conjugate to the total momentum $\underline{K} = \underline{P} + \underline{p}$. Then the wave function outside the range of forces (i.e., for x > R) is simply

$$\psi(\underline{\bar{x}},\underline{x}) = e^{i\underline{K}\cdot\underline{\bar{X}}/\hbar}\left[e^{i\underline{k}_0\cdot\underline{x}/\hbar} + \frac{f(\hat{k}_0\cdot\hat{\underline{x}})e^{ikx/\hbar}}{x}\right]. \qquad (36)$$

In the usual treatments of scattering we are interested in the second term, which we examine outside the region to which the initial beam is geometrically confined. In fact, this is how we would measure the cross section for the detection process, since $\pi a_0^2 = (\pi/2)\int_{-1}^{\cos\theta_C} d\xi|f(\xi)|^2$.

In order for our detector not to fire every time a particle goes through we must require that $\pi a_0^2 < w\ell$ (even if there is only one particle in the detector) and 1/n of that value if there are n particles. But our geometrical conditions already require us to have $w \lesssim d \ll r$, so that by the time the scattered wave reaches D_3 the scattered amplitude is down by $\sim a_0/r \sim d/r$

(and the intensity by $[a_0/r]^2$) compared to the leading term, which is just the initial free particle wave function (times the detector particle wave function). Further, simply for geometrical reasons, the range of forces R within which Equation 36 has to be modified must be much smaller than w in order to fit the detecting systems into the slit without introducing significant wall effects. We conclude that, if $m/M \ll \Theta$, $a_0 < w$, and $R \ll w$, the detector will function as postulated in Section II and that the calculation of corrections, though complicated in detail, can be carried out in terms of experimentally accessible quantities. Further corrections due to nonuniform distribution of the density ρ of detector particles across the slits, momentum dependence of the cross section, nonuniformity of the amplification process which starts from the recoil particle of mass m and ultimately determines in practice the time-resolution characteristics of the detector, not to mention other woes that will occur to an experimental physicist, can also be included. They enter the calculation in the integration of the probability of detection over the detector volumes and the time gates. Since we have already discussed in the last section how these integrals are to be calculated from a quantum mechanical wave function, assuming uniform distributions, we do not repeat the discussion here using these experimental resolution functions; they are simple multiplicative factors of the integrand from a mathematical point of view.

V. DISCUSSION AND CONCLUSIONS

We have now completed our operational analysis of how the double slit experiment can be divided into three objectively distinguishable classes of events, how these are related to counts in detectors with finite spatial and temporal resolutions, and to a limited extent how the corrections to the simple prediction can be either measured or bounded. Our analysis proceeded by two steps. In the first, we showed that we can, in fact, measure momentum to arbitrarily high precision by an appropriate collimator design and hence make the beam incident on the slits as 'monochromatic' as we wish. We then showed that, provided only that the events which initiate the firing of the detectors are single scatterings from particles of mass m much smaller than the mass M of the beam particles under study, the cross section for these scatterings πa_0^2 is much smaller than the cross-sectional area of the detectors in the slits and that the range of force R between the beam particles and the detector particles is small compared to the smallest limiting geometrical dimension (usually Δz, the thickness of the detectors in the beam direction), we can also make the perturbation due to the detectors as small as we wish.

This last restriction on the operation of the experiment, which requires us, in principle, to be able to imagine the existence of particles with arbitrarily small masses and dimensions, is, we believe, the most interesting result to emerge from our analysis. For a strict operationalist (Bridgman, 1928), the concept of length is not the same at atomic dimensions as for lengths that can be measured by laying down a standard ruler a certain number of times. Our attitude is somewhat different. We are exploring the properties of a model space by performing thought experiments and endeavoring to discover what physical model systems are needed to give operational meaning to the theory in question. We reserve discussion of whether this model space is connectable to physical phenomena until this discussion of the model space is completed. Thus, our preliminary conclusion is that, under the restrictions given above, the double slit experiment can be given an objective, operational description. The reason this is possible is that nonrelativistic quantum mechanics, as a theory, makes no assertions about the values of the masses, the cross sections, the range of forces, or the geometrical dimensions that are allowed; the only universal constant that enters is $\hbar$, and, if we concentrate on a particular type of particle, M. In other words, the structure of the theory is completely independent of the units we use for mass, length, and time; nonrelativistic quantum mechanics is <u>scale invariant</u>. This is not so obvious in terms of historical development, since the theory developed out of study of the Bohr model for the atom using first electrostatic and later electrodynamic forces as an integral part of the theory. You will note that I have carried through the above discussion without invoking any electromagnetic effects, partly for reasons of simplicity, but also just to make this point.

To put the matter more graphically, for the experiment we have described, we could, given world enough and time, perform the experiment with cannonballs, the slits being holes in armor plate, the surfaces covered with tar so as to make them 'absorbing'. For a detector we could shoot a stream of buckshot across the openings for a finite time (the 'time gates' Δt), with a detection corresponding to one of the buckshot being knocked out of the stream and either detected separately or found missing by weighing or counting the ones that arrive on the other side of the slit. Someone may object that with cannonballs (at 18th-century speeds) we could simply see which slit the ball went through and thus force the paradox which some see by finding that even so there is (or is not) a double slit interference pattern. But this does not work. If there is sufficient illumination to see each cannonball as it passes through one or the other slit, the radiation pressure on the cannonball is sufficient to destroy the phase coherence needed for the interference pattern to appear. We could use an optical detector, but only at such low light intensity that there is a finite probability that not a single quantum will be absorbed by the cannonball throughout its entire trip through the apparatus. In other words, we are constrained by the theory to perform the experiment in the dark!

Turning now to physics, the least massive particle known is (and has been for nearly a century during which hosts of other 'particles' have been discovered) the electron. Allowing for the moment the possibility of much less massive particles, our conceptual analysis is still adequate, and we can ask what size apparatus we would need to measure, say, the single and double slit diffraction patterns of two slits 0.1 cm wide and 5 cm apart, using electrons in the lowest easily accessible experimental range of velocities, which gives a de Broglie wavelength of about 10^{-8} cm. In order to give sufficient intensity to the double slit pattern, we ask that the first minimum of the single slit pattern from slit 1 fall at the position of the maximum from slit 2. This requires us to choose $r \approx \pi pwd/2\hbar \approx 78.5$ km! Thus, unless we go to supercooled systems (which would allow us to an order of magnitude shorter wavelength) and microscopic slits, we could hardly expect to raise the funds for a direct experimental test of macroscopic de Broglie wave interference. Microdesign and low velocity pose additional problems, since electrostatic image forces induced in the slits by the charge of the electrons, whether the slits are made of conducting or insulating material, would probably override the quantum effect we are seeking. Even so, we could not perform the experiment as described above, since we lack detecting particles smaller than the electron.

The actual tests of the de Broglie wavelength hypothesis are therefore much less less direct than the experiment we have been discussing. They have to assume that the particles under study can be used as 'detectors' by calculating the probabilities of interaction using the wave functions as functions of space and time coordinates (unmeasurable in operational terms) defined by electromagnetic forces. As Bridgman pointed out, so long as only electromagnetic forces are available, it is not clear that these space-time coordinates that occur in the theory have an 'objective' significance, since the operations that define them can only be performed in one way. Fortunately for us, Colella, Overhauser, and Werner (1975) have recently obtained clear de Broglie interference patterns between neutron beams that have passed through regions of different gravitational potential and are then recombined coherently. Thus, the spatial coordinates of the gravitational potential act empirically in the same way as the spatial coordinates of the electrostatic potential in the usual nonrelativistic treatments of atomic effects. This gives us some confidence in extrapolating macroscopic ideas about space and time, as they occur in the model space of quantum mechanics down to atomic (10^{-8} cm) dimensions.

Granted this, the tests of de Broglie interference patterns performed initially by Davisson and Germer, and repeated for many situations and types of particles (neutral atoms, neutrons, etc., as well as electrons) since, can be invoked as giving evidence close to that of our thought experiment. Typically these are performed not by transmission through slits but by reflection from the surface of a crystal. Since we can use x-rays, calibrated in wavelength against mechanically ruled gratings, to measure the spacing between the atoms on the surface of the crystal, we can argue that the coherent interference from this regular array containing many sites rather than just two is good evidence for the de Broglie wavelength hypothesis. We also can use data from such systems to obtain information analogous to the detectors in the slits in our thought experiment, since we get individual scatterings from individual atomic sites as well as the interference pattern. These are analogous to the single slit patterns of our experiment. We can even perform the analogue of checking this hypothesis by changing the density of detecting particles, since the single atomic patterns (which add incoherently) change with the temperature of the sample in a different way than the blurring that increasing temperature introduces into the coherent interference pattern.

We therefore believe that, although the evidence is indirect, the concept of well-defined geometrical structure and well-defined momenta (subject of course to the restrictions of the uncertainty principle when we try to define both at once) can be self-consistently defined at the atomic level. We think it is probably not often realized, however, that conceptually this implies the existence of particles of much smaller mass and cross section than those under study, if we want to feel comfortable in using all the concepts of the space-time continuum usually assumed without comment in writing down nonrelativistic wave functions. The reason we can get away with this is that we can, by using sufficiently large macroscopic apparatus, define momentum to arbitrarily high precision. This in turn enables us to measure short distances <u>indirectly</u> through quantum mechanical effects which can also be blown up to macroscopic size by going to sufficiently large detecting apparatus. For instance, we might view our double slit experiment as a measurement of the distances w and d. If these were parameters in a molecular model which was free to move under the impact of our beam particles M, rather than being part of the massive baffles which hold the openings under study rigidly in place, the analysis would be less direct but in a reasonably objective sense would still count as a measurement of (quantum mechanical, <u>not</u> classical) distance. From this point of view, the main difference between quantum mechanical and classical model space is that in the quantum mechanical space we can refine either our momentum or our distance measurements as much as we please separately, or sequentially, or accept indirect consequences when our measurement gives some information about each, but not both simultaneously; in the classical model space we are allowed to imagine limiting procedures without these restrictions. But both spaces have limiting procedures and hence conceptually define continua.

The main reason for questioning these limiting procedures, and hence raising doubts about the operational meaningfulness of the space-time concept itself at short distances, comes for me from an effect we have so far not mentioned. This comes from the coupling of relativity to quantum mechanics and was first clearly and succinctly discussed by Wick (1938) in his explanation of the range of nuclear forces given by Yukawa's (1935) meson theory. I have used this insight as one of the starting points in my attempts (Noyes, 1975, 1976, 1977; Note 3) to reformulate elementary particle physics in more operational language. Briefly, the problem is that in a relativistic theory, once one tries to define distances shorter than the compton wavelength (h/mc) of the lightest particle involved, the corresponding momenta are sufficiently energetic to provide the full rest energy (mc^2) of this particle, and additional particles appear in the system. Even if we are well below the energetic threshold for particle creation, the uncertainty principle allows such particles to be present for short times with finite probability. These 'virtual particles' produce indirectly observable effects, such as the nuclear force already alluded to, 'vacuum polarization', etc. Thus, the limiting procedures we have defined above fail at short distance or for sufficiently high momenta.

By now there is considerable experimental evidence that, except for a few absolute conservation laws (charge, baryon number, lepton number, exclusion principle), all known types of particles can be transformed into each other more or less readily at sufficiently high energy. It therefore seems very unlikely that a massive particle less massive than the electron will be discovered in the near future. Thus, in a relativistic quantum theory with this empirical fact added, the units of mass, length, and time are fixed by $\hbar$, c, and m_e, and the theory is no longer scale invariant. So far as we can see, this is a powerful argument against using model spaces that allow indefinitely small (continuum) limits to be taken. This is not a new stricture. Bohr and Rosenfeld (1933) analyzed the much more complicated case of the quantum theory of the electromagnetic field in a manner similar in spirit to my presentation here. In that case one has two universal constants (c and $\hbar$) but is allowed to use arbitrarily large masses and charges. Since the theory is scale invariant, they consider themselves at liberty to invoke an arbitrarily complicated apparatus within a wavelength (and it gets pretty complicated) but do finally succeed in rederiving the commutation relations for the electromagnetic field operators. They note in their paper that the argument breaks down in any theory with an intrinsic mass and hence that the same argument cannot be applied to the second quantization of the matter field. In a sense, this current effort of mine can be considered as an addendum to their paper, in which I attempt to give precise operational meaning to nonrelativistic quantum mechanics but end up with the conclusion that this is inevitably frustrated by the Wick-Yukawa mechanism.

References

Bohr, N., & Rosenfeld, L. On the measurability of the electromagnetic field. Konig Danske Vidensaberns Selskab, Mathematische-Fysik, 1933, 12(8).

Colella, R., Overhauser, A. W., & Werner, S. A. Observation of gravitationally induced quantum interference. Physical Review, 1975, 34, 1472-1474.

Goldberger, M. L., & Watson, K. Collision theory. New York: Wiley, 1964.

Jeffries, H., & Jeffries, B. S. Mathematical physics. Cambridge, England: Cambridge University Press, 1950.

Noyes, H. P. Fixed past and uncertain future: A single time covariant quantum particle mechanics. Foundations of Physics, 1975, 5, 37-43 [Erratum: 6, 125].

Noyes, H. P. A Democritean phenomenology for quantum scattering theory. Foundations of Physics, 1976, 6, 83-100.

Noyes, H. P. A Democritean approach to elementary particle physics. In F. Gilman (Ed.), Proceedings of the 1976 SLAC Summer Institute, SLAC 198, November 1976.

Wick, G. C. On the range of nuclear forces in Yukawa's meson theory. Nature, 1938, 142, 993-994.

Yukawa, H. On the interaction of elementary particles, I. Proceedings of the Physico-Mathematical Society of Japan, 1935, 17, 48-57.

Reference Notes

1. Noyes, H. P. A scientific retrodiction of our past (Stanford Linear Accelerator Center PUB-1380). Unpublished manuscript, Stanford University, 1974.
2. Noyes, H. P. Professional responsibility (Seminar at Stanford for the Social Thought Program). Unpublished manuscript, Stanford University, 1976.
3. Noyes, H. P. Non-locality in particle physics (Stanford Linear Accelerator Center PUB-1405, rev.). Unpublished manuscript, Stanford University, 1975.