COSMOLOGY AND THE PILOT WAVE INTERPRETATION OF QUANTUM MECHANICS

Frank J. TIPLER
Department of Mathematics and Department of Physics, Tulane University, New Orleans, LA 70118, USA

Received 29 June 1983
Revised manuscript received 9 January 1984

Bell has recently revived the pilot wave interpretation of de Broglie and Bohm as a possible scheme for interpreting wave functions in quantum cosmology. I argue that the pilot wave interpretation cannot be applied consistently to systems whose wave functions split into macroscopically distinguishable states. At some stage the pilot wave interpretation must either tacitly invoke wave function reduction in the same manner as the Copenhagen interpretation, or else abandon locality by requiring physical particles to move faster than light. Consequently, the many-worlds interpretation is the only known realist interpretation of the quantum mechanical formalism which can be extended to quantum cosmology.

Interpreting the wave function is one of the most difficult problems in quantum cosmology. There are several reasons for this. First, most interpretations of the wave function in non-relativistic quantum mechanics require a clear division of the total physical system into the observer and the observed, with the former standing clearly outside the latter. Such a split is quite impossible in cosmology, for there is nothing outside the universe. Second, the interpretation of the absolute square of the wave function as a probability density is usually accomplished by analyzing not a single system, but rather an ensemble of similarly prepared systems. Again, such a procedure is impossible in cosmology, because there is only one universe.

Since no direct appeal to experiment can be made in quantum cosmology to interpret the wave function, one must either argue that the interpretation arises from the wave function itself – as is argued by many proponents of the many-worlds interpretation [1–3] – or one must find an interpretation that is consistent with a priori assumptions about the nature of the entities constituting the universe. For example, each classical cosmological model of general relativity describes time evolution of what is essentially a single classical system: a universe can be considered as a point in superspace, and the evolution of such a point is a trajectory in superspace governed by a generalized geodesic equation with a force term. Thus, if an interpretation of the wave function could be found which could reproduce all the usual statistical implications of quantum mechanics by invoking only the classical entities of particles acting under the influence of forces, this interpretation could be extended immediately to the wave function in quantum cosmology. The interpretation problem would disappear, since nothing fundamentally new would be added by quantization.

Such a "classical" construction of quantum mechanics was independently invented by de Broglie [4–6] in the 1920's and by Bohm [7–11] in the 1950's. This construction, called the pilot wave interpretation, has recently been resurrected by Bell [12] for the express purpose of providing quantum cosmology with its long-sought wave function interpretation. Unfortunately, I shall show in this paper that the pilot wave interpretation does not quite attain its goal of providing a purely classical picture of quantum phenomena. Like the Copenhagen interpretation, the pilot wave interpretation must at some point suspend the evolution equations and demand wave function collapse, or else force physical particles (not merely pilot waves) to travel faster than light. Consequently, the pilot wave interpretation cannot be used to interpret wave functions in quantum cosmology.

To see the difficulty, let us review the salient features of the pilot wave interpretation. It will be sufficient for our purposes to restrict attention to the one-particle Schrödinger equation (we set $h/2\pi = 1$):

$$i\partial\psi/\partial t = -(1/2m)\nabla^2\psi + V(x)\psi \ . \qquad (1)$$

If we substitute

$$\psi = R\,\exp(iS)\ , \qquad (2)$$

where the functions $R = R(x, t)$ and $S = S(x, t)$ are real, into (1), we obtain

$$\partial R/\partial t = -(1/2m)[R\nabla^2 S + 2\nabla R\cdot\nabla S]\ , \qquad (3)$$

$$\partial S/\partial t = -(\nabla S)^2/2m - [V(x) - (1/2m)/\nabla^2 R/R]\ . \qquad (4)$$

Eq. (3) can be re-written in terms of $\psi\psi^* = R^2 \equiv \rho$, a quantity which in the conventional interpretation is the probability density, as

$$\partial\rho/\partial t + \nabla\cdot(\rho[\nabla S/m)] = 0\ . \qquad (5)$$

Now the system of equations (4) and (5), a system which is completely equivalent to Schrödinger's equation (1), can be given a classical interpretation. Eq. (4) is just the Hamilton–Jacobi equation for a particle of mass m and velocity $\nabla S/m$, moving in the potential

$$V(x) - (1/2m)\nabla^2 R/R\ . \qquad (6)$$

The additional potential term $-(1/2m)(\nabla^2 R/R)$ was named the "quantum potential" by Bohm [7,8]. From standard Hamilton–Jacobi theory, any ensemble of particles, each of mass m and each with initial velocity $\nabla S/m$, will move on trajectories normal to the S = const surfaces. Such an ensemble of non-interacting particles will define a particle density ρ, which will satisfy an equation of continuity. The appropriate equation of continuity is precisely (5), so it is natural to regard the system of equations (4) and (5) as giving the spacetime evolution of such an ensemble. Thus the pilot wave interpretation is a theory of the evolution of both single particles, and ensembles of non-interacting particles. A single particle in the pilot wave interpretation actually has spacetime coordinates $x(t)$ and velocity

$$\dot{x} = \nabla S/m\ . \qquad (7)$$

The quantities x and $\dot{x}$ are termed "hidden variables". The system of equations (4), (5) and (7), subject to the initial conditions

$$\rho(x, 0) \equiv \rho_0(x)\ , \qquad (8)$$

$$S(x, 0) \equiv S(x)\ , \qquad (9)$$

$$x(0) \equiv x\ , \qquad (10)$$

completely defines the physical system consisting of either a single particle or ensemble of particles acted on by the potential (6), and the time evolution of this system.

As emphasized by Keller [13], probability is supposed to enter the theory through incomplete knowledge of initial data or through the use of ensembles, just as in classical statistical mechanics. If the former, the quantities $\rho_0(x)$ and $S(x)$ are asserted to be known exactly, but x_0 is not, which is why $x(t)$ and $\dot{x}(t)$ are called "hidden variables". In a general statistical theory, an initial probability distribution $\phi_0(x)$ of x is assumed to be given, and by the law of conservation of probability, this probability distribution $\phi(x, t)$ as a function of time satisfies

$$\partial\phi/\partial t + \nabla\cdot(\phi\nabla S/m) = 0\ , \qquad (11)$$

$$\phi(x, 0) = \phi_0(x)\ , \qquad (12)$$

where the $S(r, t)$ in eq. (11) is the same $S(x, t)$ that appeared in (4) and (5).

The key difficulty with the pilot wave interpretation is that although $\rho(x, t)$ and $\phi(x, t)$ satisfy the same equation, they need not be equal, since they in general satisfy independent initial conditions (8) and (12), respectively. The function $\rho_0(x)$ represents the initial value of a real quantum mechanical field which controls the evolution of a particle or ensemble, while $\phi_0(x)$ represents the initial probability or ensemble distribution of x. As pointed out by Keller [13], in standard statistical mechanics $\phi_0(x)$ is arbitrary and need not equal $\rho_0(x)$. Only if one imposes $\phi_0(x) = \rho_0(x)$ as an additional postulate in the pilot wave interpretation can one truly recover the usual statistical interpretation of $\rho(x, t)$. Bohm [9] claims to *derive* the equality $\phi_0(x) = \rho_0(x)$. He asserts that the long-time evolution of the system of eqs. (4), (5), and (7) would result in $\phi(x, t)$ approaching $\rho(x, t)$ as $t \to +\infty$, no matter what $\phi_0(x)$ was. Indeed, one might expect such asymptotic properties of the system (4), (5), and (7), since $\phi(x, t) \to \rho(x, t)$ in classical statistical mechanics if $\rho(x, t)$ is the equilibrium distribution. However, I shall show Bohm's derivation to be incorrect.

It will prove instructive first to review a simple example which illustrates why $\phi(x, t)$ need not equal

$\rho(x, t)$. Consider a stationary state. Such a state will have a wave function of the form $\psi(x, t) = \Psi(x) \times \exp(-iE_n t)$. For many physical situations – e.g. hydrogen atom S-states and standing waves in a one-dimensional box – the function $\Psi(x)$ is a real number, so $\psi(x) = R(x, t)$ and $S(x, t) = 0$. For the standing wave in a one-dimensional box of length L, $R(x, t) = R(x) = (2/L)^{1/2} \sin(\pi n x/L)$. We have $p = \nabla S/m = 0$, so a particle of non-zero energy E_n does not move. As emphasized by Einstein [14] this is a rather paradoxical result, because it is true even if the mass m and the quantum number n are large enough for the quantized particle in the box to be considered a classical object. This would apparently violate the principle of correspondence, because a particle of non-zero energy in a box *would* move. Nevertheless, $p = 0$ is consistent with the interpretation of a particle of energy E_n moving in a potential given by (6), because $V(x) = 0$ but $(-1/2m)(\nabla^2 R)/R = n^2\pi^2/2mL^2$ which is equal to E_n. Thus the particle does not move because all of its energy is in the form of "quantum" potential energy. In his reply to Einstein's criticism, Bohm claimed [8,15] the correspondence principle was not violated because an observation of a particle in the box could be carried out only by changing the particle's standing pilot wave into a traveling wave packet which gave the expected classical notion, on the average.

Bohm's argument could hold only if $\phi_0(x) \approx \rho(x) = (2/L)\sin^2(\pi n x/L)$. Otherwise the notion of the wave packet need not have any correspondence with the average motion of the particle. Furthermore, note that the explicitly quantum mechanical forces, given by $(-1/2m)\nabla^2 R/R$, clearly do *not* act in this simple example to make $\phi(x, t)$ approach $\rho(x, t)$. The quantum force is zero everywhere, since the quantum potential is a constant. A zero force cannot change the initial distribution of particles in a box. One might think that a violent change in the quantum potential could cause $\phi(x, t)$ to approach $\rho(x, t)$. One often sees claims by pilot wave supporters [7,8,12] (see also ref. [6]) that such violent changes will occur at zeros of the wave function, where $R(x, t) = 0$. But zeros of $R(x, t)$ need not result in a divergence in the quantum potential $-(1/2m)\nabla^2 R/R$, because $\nabla^2 R$ could also approach zero at the zeros of R. The behavior $\nabla^2 R \to 0$ as $R \to 0$ is seen to occur by direct calculation in our simple one-dimensional box example; the quantum force is zero at all nodes. In fact, if we make the standard requirement that the fundamental physical fields – in the pilot wave interpretation these include $V(x, t)$, $R(x, t)$ and $S(x, t)$ – be everywhere C^2, then eq. (4) would imply that the quantum potential is also C^2 and hence bounded everywhere, including the wave function zeros. If the system of equations (3), (4) lead to singularities in one of the fields $V(x, t)$, $R(x, t)$, or $S(x, t)$ which were not singular points of $\psi(x, t)$ – and it is possible this could occur [16] – this would give a purely mathematical reason to reject the pilot wave theory. I will assume, however, that this problem does not arise.

If $S(x, t)$ is smooth, then we can show that $\phi(x, t)$ does *not* approach $\rho(x, t)$ if $\rho_0(x) \neq \phi_0(x)$, unless they both approach zero for all x. This follows immediately from (5). If $S(x, t)$ is smooth, then the flow lines of (5) never intersect, so using (7) we can write (5) as

$$\partial\rho/\partial t + \mathbf{v}\cdot\nabla\rho + (\rho/m)\nabla^2 S = 0 ,$$

or

$$D\rho/Dt = -(\rho/m)\nabla^2 S , \tag{13}$$

which can be integrated along the flow lines to give

$$\rho(x(t), t) = \rho_0(x) \exp\left(-\int_0^t \frac{\nabla^2}{m} S(x(\tau), \tau)\, d\tau\right). \tag{14}$$

Since $\phi(x, t)$ and $\rho(x, t)$ both satisfy (13) which is linear, the quantity $|\phi(x, t) - \rho(x, t)|$ would have a non-zero lower bound on every flow line for which $|\phi_0(x) - \rho_0(x)| \neq 0$, and along which the integral $\int_0^t \nabla^2\, d\tau$ is bounded. But this integral will be bounded unless $\phi(x, t)$ and $\rho(x, t)$ both approach zero along that flow line. In particular, $\phi(x, t)$ and $\rho(x, t)$ are not close to zero at the present time, at least for most x in the laboratory. Because of the spreading of wave packets, these functions may approach zero as $t \to +\infty$, but they are definitely non-zero today. Thus, if $\rho_0(x) \neq \phi_0(x)$ for some x, this difference would have persisted until the present epoch. I conclude that Bohm's argument to the contrary [9] is incorrect. Freistadt has also expressed doubts [17] about the mathematical validity of Bohm's proof.

It should be emphasized that the results of taking ensemble averages in the pilot wave interpretation is utterly different from the results obtained by the same procedure in statistical mechanics. In the latter case, a slight change in the microscopic initial conditions would not lead to any observable change in any macroscopic quantity such as the temperature or the pressure. Indeed, it must not if such procedures as coarse graining or assuming random, small, external perturbations on the system (the usual way of proving $\phi \to \rho$ in statistical mechanics [18,19]) are to be used. However, as Bohm's reply to Einstein's criticism makes clear, a tiny change in the initial conditions – a small perturbation due to the measurement – can cause an enormous change in a macroscopic observable, the particle velocity. Befor the measurement, a classical object was not moving. After the measurement, it is moving at a speed which can be arbitrarily large. This instability of the evolution equations in the pilot wave interpretation makes it impossible to validly carry out a coarse-grain calculation in configuration space, as Bohm tried to do to show $\phi \to \rho$.

This instability actually implies that the closer we try to approximate the true value of $\phi(x, t)$ by $\rho(x, t)$, the worse the predictions of the theory become. The true positions of a particle or an ensemble are actually [12]

$$\phi(x, t) = \sum \delta(x - x_N(t)) ,$$

where the summation is over the particles. As $\rho(x, t)$ approaches $\phi(x, t)$, it ceases to be a smooth, slowly varying function, but instead develops sharp spikes at the points where the particle(s) is(are) located. Such a distribution implies a large quantum potential, which actually diverges as $\rho(x, t) \to \phi(x, t)$. This is a consequence of the fact that a wave function which is a δ-function initially is spread out over all space the next instant. A theory which gives radically different results depending on how we choose to approximate it locally is termed "not self-consistent" by Misner et al. [20] (see also ref. [16]). They contend theories which are not self-consistent in this sense must be rejected, and they use their criteria to reject a number of otherwise viable gravitation theories.

The continuity equation (5) [and (7)] is invariant under time reversal, so if $\phi(\boldsymbol{x}, t) \neq \rho(\boldsymbol{x}, t)$ at some time t in the future, this inequality would also hold now, and for any finite time in the past. I claim that standard interpretations of the probability distribution, the assumption of locality, and the fact that definite values are obtained in position measurements for macroscopically distinguishable states imply $\phi(\boldsymbol{x}, t) \neq \rho(\boldsymbol{x}, t)$.

The details of the argument depend on which specific interpretation of the probability distribution is used. Let us first suppose $\phi(\boldsymbol{x}, t)$ measures an "intrinsic" uncertainty about the position of a single particle whose time evolution is described by (4), (5), and (7). For concreteness let us consider, following Bohm [7,8], the inelastic scattering of a single-particle wave packet off a hydrogen atom. Before the collision the wave function is

$$\Psi_\mathrm{i} = \psi_0(\boldsymbol{x}) \exp(-\mathrm{i}E_0 t) f_0(\boldsymbol{y}, t) , \tag{15}$$

where $\psi_0(\boldsymbol{x})$ refers to the initial state of the atom, and

$$f_0(\boldsymbol{y}, t) = \int \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{y}) f(\boldsymbol{k} - \boldsymbol{k}_0) \exp(-\mathrm{i}k^2 t/2m)\, \mathrm{d}\boldsymbol{k} , \tag{16}$$

describes the incident wave packet. The asymptotic form of the wave function after the collision is

$$\Psi_\mathrm{f} = \sum_n \psi_n(\boldsymbol{x}) \exp(-\mathrm{i}E_n t) f_n(\boldsymbol{y}, t) + \Psi_\mathrm{i} , \tag{17}$$

where

$$f_n(\boldsymbol{y}, t) = f(\boldsymbol{k}\cdot\boldsymbol{k}_0) r^{-1} g_n(\theta, \phi, k) \times \exp[\mathrm{i}\boldsymbol{k}_n\cdot\boldsymbol{r} - (k_n^2/2m)t]\ \mathrm{d}\boldsymbol{k} , \tag{18}$$

is an outgoing wave packet with center $\boldsymbol{r}_n = (\boldsymbol{k}_n/m)t$. When t is very large, there is negligible overlap between the various outgoing wave packets. A measurement at time t would reveal the scattered particle in some definite wave packet $f_n(\boldsymbol{y}, t)$. In other words, after the measurement $\phi(\boldsymbol{x}, t)$ is narrowly peaked in one of the wave packets, and is exactly zero in the other packets. Although the $R(\boldsymbol{x}, t)$ and $S(\boldsymbol{x}, t)$ fields are not required to propagate slower than light, physical particles are so restricted [‡1], and so $\phi(\boldsymbol{x}, t)$ must have been narrowly peaked in one of the wave packets and zero elsewhere *before* the measurement. But the identification

[‡1] I assume – as do Bohm, Bohr, and just about everyone else – that the measurement process can be analyzed without reference to relativistic modifications to Schrödinger's equation. It is legitimate to invoke the fact that no particle can move faster than light even in the non-relativistic analysis, because the actual relativistic equations should be closely approximated by the non-relativistic wave equations at the low energies considered here. A similar argument (combining non-relativistic equations with $v \leqslant c$) is made in all discussions of the EPR paradox [22].

$\phi(x, t) = \rho(x, t)$ would by the evolution equations contradict the assumption that (15) was the initial wave function. Hence $\phi(x, t)$ cannot be interpreted as some "intrinsic" uncertainty about the position of a single particle. The situation is even worse if we try to assume that the position of the particle is exactly known, which we must do if we really believe in the pilot wave interpretation. In this case $\phi(x, t) = \delta(x(t))$, which is quite different from $\rho(x, t)$: at *all* times. Thus the pilot wave interpretation has essentially the same wave function reduction problem that the Copenhagen interpretation has [23,24].

One runs into similar problems if one tries to interpret $\phi(x, t)$ as the distribution of position measurements carried out on an ensemble of collisions, or as the spatial distribution of an ensemble of particles. The key difficulty is that one wants to equate an *actual* distribution of values $\phi(x, t)$ to a *real* physical field $\rho(x, t)$. Thus $\phi(x, t)$ cannot refer to an *ideal* ensemble consisting of an actual infinity of position measurements, (or in the case of an initial ensemble of particles, an actual infinity of particles), but rather to the actual distribution one obtains when one runs the experiment a finite number of times, N (or the actual distribution of a finite number of particles). If N is small this distribution will be quite different from an idealized distribution of position requirements in the limit as $N \to \infty$. To equate an ideal distribution to an actual field would be to abandon a realistic interpretation of probability. In contrast to statistical mechanics, $\rho(x, t)$ is in no way an "equilibrium" distribution, so we cannot claim that the distribution $\rho(x, t)$ is the most probable distribution. As shown above, the quantum forces do *not* tend to make $\phi(x, t)$ equal to $\rho(x, t)$. Hence the only way to actually define the particle distribution $\phi(x, t)$ is to set up an experimental arrangement with wave function ψ. After a finite time the experimental equipment is dismantled, and the actual observed $\phi(x, t)$ would be different from that required by $|\psi^2| = \rho(x, t) = \phi(x, t)$, so $\phi(x, t) \neq \rho(x, t)$ even if $\phi_0(x) = \rho_0(x)$. The only way an actual infinity of measurements could have been carried out is if in fact there are an infinity number of universes, in each of which a position measurement is made. But this is the many-worlds interpretation, not the pilot wave interpretation.

I should like to thank E.D. Conway, D. Deutsch and J.A. Goldstein for helpful discussions.

References

[1] B.S. De Witt and N. Graham, The many-worlds interpretation of quantum mechanics (Princeton Univ. Press, Princeton, 1973).
[2] D.N. Page, Phys. Lett. 91A (1982) 57.
[3] D. Deutsch, Quantum mechanics as a universal physical theory, University of Texas preprint (1982).
[4] L. de Broglie, Non-linear wave mechanics: a causal interpretation (Elsevier, Amsterdam, 1960).
[5] L. de Broglie, The current interpretation of wave mechanics: a critical study (Elsevier, Amsterdam, 1964).
[6] M. Jammer, The philosophy of quantum mechanics (Wiley, New York, 1974).
[7] D. Bohm, Phys. Rev. 85 (1952) 166.
[8] D. Bohm, Phys. Rev. 85 (1952) 180.
[9] D. Bohm, Phys. Rev. 89 (1953) 458.
[10] D. Bohm and J.P. Vigier, Phys. Rev. 96 (1954) 208.
[11] D. Bohm, B. Schiller and J. Tiomno, Nuovo Cimento Suppl. 1 (1955) 48.
[12] J.S. Bell, in: Quantum gravity 2: A Second Oxford Symposium, eds. C.J. Isham, R. Penrose and D.W. Sciama (Oxford Univ. Press, London, 1981).
[13] J.B. Keller, Phys. Rev. 89 (1953) 1040.
[14] A. Einstein, in: Scientific papers presented to Max Born (Hafner, New York, 1953) p. 33.
[15] D. Bohm, in: Scientific papers presented to Max Born (Hafner, New York, 1953) p. 13.
[16] J. Smoller, Shock waves and reaction–diffusion equations (Springer, Berlin, 1983).
[17] H. Freistadt, Nuovo Cimento, Suppl. 5 (1957) 1.
[18] I.E. Farquhar, Ergodic theory in statistical mechanics (Wiley, New York, 1964).
[19] K. Huang, Statistical mechanics (Wiley, New York, 1963).
[20] C.W. Misner, K.S. Thorne and J.A. Wheeler, Gravitation (Freeman, San Francisco, 1973) p. 1066.
[21] C.M. Will, Phys. Today 25 (1972) 23.
[22] A. Einstein, B. Podolsky and N. Rosen, Phys. Rev. 47 (1935) 777.
[23] J.M. Jauch, E.P. Wigner and M.M. Yanase, Nuovo Cimento 48 (1967) 144.
[24] A. Loinger, Nuc. Phys. A108 (1968) 245.