International Journal of Theoretical Physics, Vol. 25, No. 5, 1986

Why Is the Universe So Large?

Don N. Page[1]

Received October 2, 1985

S. W. Hawking's proposal for the wave function of the universe, if correct, determines the conditional probabilities for all properties of the universe. In a simple minisuperspace model it predicts that at any given nonzero energy density, the universe is most probably infinitely large.

Physical models for the universe consist of three parts: (1) physical variables, (2) dynamical laws, and (3) boundary conditions. Generally theoretical physics has concentrated on the first two parts and regards a complete account of them to be a unified field theory, while considering the third part to be arbitrary. But we know that the boundary conditions must be very special in order to explain the flatness, homogeneity, isotropy, and arrow of time of our universe.

Hawking (1982, 1984a,b; Hartle and Hawking, 1983) has proposed that "The boundary conditions of the universe are that it has no boundary," by which he means that the probability or square of the wave function for any three-geometry and three-dimensional matter field configuration on it is given by a path integral over all compact positive-definite four-geometries and four-dimensional matter field configurations, without boundary, containing the desired three-dimensional one. For many purposes it is more convenient to work with the wave function ψ itself, the square root of the probability. It is given by a path integral over all compact four-geometries to one side of the desired three-geometry, that is, having it as its one and only boundary. Hartle and Hawking (1983) showed that this wave function obeys the Wheeler-DeWitt equation, a second-order hyperbolic differential equation in superspace, the infinite-dimensional space whose coordinates describe the three-dimensional geometry and matter field configuration. Hawking's proposal for the wave function then amounts to giving boundary

[1]Department of Physics, Pennsylvania State University, University Park, Pennsylvania 16802.

0020-7748/86/0500-0545$05.00/0

conditions for ψ on a Cauchy surface in superspace and using the Wheeler–DeWitt equation to evolve the wave function to other points of superspace.

The interpretation of the wave function ψ is that $|\psi|^2$ is proportional to the probability density or measure for the three-geometry and matter fields to be at a particular point of superspace. The probability or measure for the universe to have some property A is then

$$P(A) \propto \int \bar{\psi} \mathbb{P}_A \psi * 1 \tag{1}$$

where $\mathbb{P}_A$ is the projection operator onto the property A and $*1$ is the volume element in superspace. However, the absolute probability or measure is untestable by observations made within the universe, and this is reflected in the fact that the wave function is not normalizable over superspace—the integral in (1) diverges if $\mathbb{P}_A$ is the identity operator. Instead, only conditional probabilities are both normalizable and testable by observations. With condition B suitably defining what observation is made, the conditional probability of the observational result A is then

$$P(A|B) = \frac{\int \bar{\psi} \mathbb{P}_B \mathbb{P}_A \mathbb{P}_B \psi * 1}{\int \bar{\psi} \mathbb{P}_B \psi * 1} \tag{2}$$

For example, if B is the condition that an individual exists with the appropriate faculties and looks at the sky at night, one can ask for the conditional probability of some result A, say that individual's seeing the sky to be in all directions as bright as the sun. If $P(A|B)$ is calculated to be much less than unity, and if one does not get A given B, then the theory passes this test. If, on the other hand, one had used a wave function incorporating the naive expectation from Olbers paradox so that $P(A|B)$ was calculated to be very near unity, then the failure to observe A given B would refute this theory at the confidence level given by $P(A|B)$. Thus, $P(A|B)$ is testable (with a confidence that depends on how close it is to 0 or 1), whereas $P(A)$ is not, because with no way to measure $P(B)$ observationally, there would be no way to check the absolute size of $P(A)$.

One important test of Hawking's proposed wave function would be whether it predicts a universe as large and spatially flat as what we observe. This is the flatness problem of cosmology—how is it that the universe has expanded to at least 10^{183} Planck volumes and yet the density has not dropped so low that gravity becomes unimportant? It would be difficult to give a complete, precise answer to this question, first because it would be formidable to calculate the complete Hawking wave function for all the degrees of freedom of the three-geometry and matter fields, and second because it would be arduous to specify completely the condition defining our existing and making observations of the size and flatness of the universe.

However, we can check what is predicted for a highly idealized model and condition.

Hence we go to a minisuperspace model, in which the wave function depends only on a finite number of degrees of freedom. We need at least one variable representing the property we are testing (size of the universe) and another representing the condition of our making observations, so we consider a two-parameter minisuperspace model consisting of a homogeneous, isotropic geometry (Friedmann–Robertson–Walker) minimally coupled to a self-interacting homogeneous scalar field (Hawking, 1984a,b; Page, 1985a; Hawking and Page, 1986). Following Hawking and Page (1986), the three-geometry will be given by the three-metric

$$g_{ij} = a^2\sigma^2\tilde{g}_{ij} \tag{3}$$

where $\tilde{g}_{ij}$ is a metric for a compact three-space of constant unit curvature

$${}^3\tilde{R}_{ijkl} = k(\tilde{g}_{ik}\tilde{g}_{jl} - \tilde{g}_{il}\tilde{g}_{jk}) \tag{4}$$

with $k = +1$, 0, or -1, where σ is a normalization constant chosen so that the three-volume of this unit-curvature space is $4\pi G/(3\sigma^2)$, and where a is the scale parameter giving the size. The homogeneous scalar field (constant on each three-space) will be expressed as

$$\Phi = (4\pi G/3)^{-1/2}\phi \tag{5}$$

and its self-interaction potential will be written as

$$U(\Phi) = \frac{3}{4\pi G\sigma^2}\tilde{U}(\phi) \tag{6}$$

where ϕ is taken as the second parameter of this two-dimensional minisuperspace, so $\psi = \psi(a, \phi)$.

Now Hawking's path-integral proposal for $\psi(a, \phi)$ is equivalent to giving the Wheeler–DeWitt equation it obeys and appropriate boundary conditions to select the correct solution of this equation. The results depend on the measure chosen for the path integral, which affects the factor ordering of the Wheeler–DeWitt equation. Hawking and I proposed that the differential operator in this equation should take the form of the Laplacian in a natural metric in the superspace (Hawking and Page, 1986), which in this two-dimensional case (with the lapse function N chosen to be the constant σ) is the flat metric

$$ds^2 = -a\,da^2 + a^3\,d\phi^2 = -du\,dv \tag{7}$$

with null coordinates

$$u = \tfrac{2}{3}a^{3/2}e^{-3\phi/2}, \qquad v = \tfrac{2}{3}a^{3/2}e^{+3\phi/2} \tag{8}$$

Then the Wheeler–DeWitt equation becomes

$$(-\tfrac{1}{2}\nabla^2+V)\psi\equiv\left(2\frac{\partial^2}{\partial u\,\partial v}+a^3\tilde{U}-\tfrac{1}{2}ka\right)\psi$$

$$\equiv\frac{1}{2a^3}\left(a\frac{\partial}{\partial a}a\frac{\partial}{\partial a}-\frac{\partial^2}{\partial\phi^2}+2a^6\tilde{U}-ka^4\right)\psi(a,\phi)=0 \qquad (9)$$

which is the Klein–Gordon equation for a particle of variable mass-squared $2V=2a^3\tilde{U}-ka$.

The boundary conditions necessary to select a particular solution of (9) may be chosen to be the value of ψ along the null cone $uv\equiv\frac{4}{9}a^3=0$ that is the past causal boundary of the minisuperspace. One can then estimate that, so long as $\tilde{U}(\phi)$ rises slower than $e^{6|\phi|}$ for large $|\phi|$, the dominant contribution to the path integral for points near the boundary will come from paths having negligible action, so ψ should be nearly constant along this boundary (Page, 1985a; Hawking and Page, 1986). Then for many [but not all (Hawking, 1985; Page, 1985b)] purposes it may be a sufficiently good approximation to set $\psi=1$ along the boundary, remembering that the overall scale of ψ does not affect the testable conditional probabilities given by (2).

To find $\psi(a,\phi)$ for $a>0$, we hence integrate (9) forward in the timelike variable a of the metric (7) from the approximate boundary condition $\psi(0,\phi)=1$. Assuming that $\tilde{U}(\phi)$ increases monotonically well above unity for large $|\phi|$ but always at a much slower logarithmic rate than $e^{6|\phi|}$ does (e.g., as the free massive self-interaction $\tilde{U}=\frac{1}{2}m^2\phi^2$ does for $|\phi|\gg 2^{1/2}m^{-1}+\frac{1}{3}$), we have that the wave function in this large-$|\phi|$ region has the approximate Bessel-function form

$$\psi=J_0[\tfrac{1}{3}a^3(2\tilde{U})^{1/2}] \qquad (10)$$

When the argument of the Bessel function is large, one can write this in the WKB form

$$\psi=\psi_+ + \psi_- = C\,e^{iS}+\bar{C}\,e^{-iS} \qquad (11)$$

where the prefactor and phase have the values

$$C=\left(\frac{2\pi}{3}a^3\right)^{-1/2}(2\tilde{U})^{-1/4},\qquad S=\frac{\pi}{4}-\frac{1}{3}a^3(2\tilde{U})^{1/2} \qquad (12)$$

In this WKB regime where the wave function oscillates rapidly, it may be interpreted as a superposition of an ensemble of classical wave packets which move along trajectories normal to the surfaces of constant phase S.

These trajectories obey the classical Friedmann–Robertson–Walker scalar-field equations

$$\dot{a}^2 = a^2\dot{\phi}^2 + 2a^2\tilde{U}(\phi) - k \tag{13}$$

$$\ddot{\phi} + 3a^{-1}\dot{a}\dot{\phi} + d\tilde{U}/d\phi = 0 \tag{14}$$

where the overdot represents d/dt in the classical four-dimensional Lorentzian-signature metric

$$ds^2 = \sigma^2(-dt^2 + a^2\tilde{g}_{ij}\,dx^i\,dx^j) \tag{15}$$

represented by each trajectory. In the large-$|\phi|$ region where (10)–(12) are valid, the trajectories have

$$\frac{\dot{a}}{a} \approx (2\tilde{U})^{1/2} \tag{16}$$

$$\dot{\phi} \approx -\frac{1}{3}\frac{d}{d\phi}(2\tilde{U})^{1/2} \tag{17}$$

which gives a long period of roughly exponential expansion or inflation.

Although there is a two-parameter family of solutions of the classical FRW equations (13) and (14) (not counting the choice of origin for the time parameter t), the trajectories given by the wave function (10) form only a one-parameter set obeying (16) and (17). Thus in the WKB regime ψ does represent an ensemble of many classical worlds, but it does not represent all possible worlds. Its predictive power is contained in the probability distribution it gives over the set of classical trajectories it represents, and in the fact that the probability for other trajectories is very small (zero in the WKB approximation).

For large $|\phi|$, this one-parameter set of trajectories may be labeled by ϕ_0, the value of ϕ they have where ψ first crosses zero as a is increased from zero [roughly where $S = -\pi/2$ or $a = (9\pi/4)^{1/3}(2\tilde{U})^{-1/6}$]. With our suggested factor ordering in the Wheeler–DeWitt equation and with the probability measure (1) or (2) in the natural minisuperspace metric (7), the probability contributed to a certain region of minisuperspace by the trajectories of a WKB component such as $\psi_+ = C\,e^{iS}$ is proportional to the proper time $\sigma\Delta t$ that the trajectories spend in the region, multiplied by the magnitude of the Klein–Gordon flux carried through the region by that WKB component (Hawking and Page, 1986). For the WKB component ψ_+ given by (11) and (12), the flux carried along a pencil of trajectories labeled by large $|\phi_0|$ works out to be simply proportional to $\Delta\phi_0$, the spread in ϕ_0 of these trajectories. The WKB component ψ_- will give an opposite flux but an equal contribution to the probability (1) of being in a certain region of minisuperspace, and the interference effects between ψ_+ and ψ_- in $|\psi|^2$ will average out if one integrates over a region large compared with the size of the oscillations of ψ, which is quite small (even in Planck units) in the WKB regime.

Now one can extend the WKB solution starting in the large-$|\phi|$ region by following the solutions of (13) and (14) into the small-$|\phi|$ region where (16) and (17) are no longer valid. If $\tilde{U}(\phi)$ decreases monotonically to an absolute minimum value of zero at $\phi = \phi_m$ and has a quadratic dependence on ϕ near that minimum, the classical trajectories will eventually undergo damped oscillations around that minimum. Averaged over several oscillations, the trajectories will behave as a dust-filled FRW model with

$$\dot{a}^2 = Ea^{-1} - k \tag{18}$$

where

$$E = a^3(\dot{\phi}^2 + 2\tilde{U}) \tag{19}$$

is approximately conserved during the oscillatory phase and represents the equivalent amount of dustlike "mass" the solution has. During the inflationary phase when $|\phi|$ is large, E grows rapidly, and one can integrate this growth to estimate that when a trajectory that starts at ϕ_0 enters its oscillatory phase, it will then have

$$\ln E(\phi_0) \approx 18 \int_{\phi_m}^{\phi_0} \left(\frac{d\tilde{U}}{d\phi}\right)^{-1} \tilde{U}(\phi)\, d\phi + \text{log terms} \tag{20}$$

Hence if ϕ starts at large ϕ_0, E becomes enormous.

Once we have the classical trajectories given by the wave function in the WKB regime and the probability distribution for them, we can calculate conditional probabilities by (2). In the simple two-dimensional minisuperspace model, we cannot really have sufficient conditions for an observer. However, we can argue that an observer could exist in a more realistic model only if the matter density were in some reasonable range, bounded away from zero and from excessively high values. Hence in our simplified model we will take the condition B to be that the energy density lies in some finite range above zero but well below the Planck value. Then property A can be that the size of the universe lie within some range consistent with observations, and the question is whether $P(A|B)$ is not too small to make such observations unlikely.

For simplicity, express the size of the universe in terms of a, and let ρ be $8\pi G\sigma^2/3$ times the density, so that

$$\rho = E/a^3 \tag{21}$$

As discussed above, the probability dP contributed by trajectories that started at large ϕ_0 is proportional purely to the spread $d\phi_0$ of the trajectories and to dt along the trajectories, so $dP/dt\, d\phi_0$ is constant. Then changing variables to a and ρ in the dustlike regime (where $\rho \ll 1$ and $\rho a^2 - k > 0$)

by using (18), (20), and (21), and using the fact that the integrand of (20) at ϕ_0 has only a very weak dependence on $E(\phi_0)=\rho a^3$ when it is very large, leads to a probability distribution of the approximate form

$$dP \propto \rho^{-1}(\rho a^2 - k)^{-1/2}\, d\rho\, da \tag{22}$$

for $a^{-3} \ll \rho \ll 1$ (and $\rho a^2 - k > 0$).

One sees that for B giving a fixed range of ρ, the unnormalized conditional probability distribution for a integrates to a logarithmic divergence at infinite a. This is simply because the probability distribution over trajectories is flat as a function of ϕ_0, and each trajectory spends roughly the same amount of time (and hence contributes the same probability) passing through the fixed range of ρ (if ϕ_0 is large enough that $\rho a^2 \gg 1$ there). Hence the contribution diverges as $|\phi_0|$ (and therefore also a, at fixed ρ) is taken to infinity. Thus the normalized conditional probability $P(A|B)$ is zero if the range of a given by A does not include $a=\infty$ and is unity if it does include $a=\infty$. In other words, at fixed energy density, the universe given by Hawking's wave function in this minisuperspace model most probably is infinitely large. In this homogeneous, isotropic case, it would also most probably be precisely flat (e.g., ${}^3R = 6k\sigma^{-2}a^{-2} = 0$), though in a more realistic model including inhomogeneous and anisotropic modes (Halliwell and Hawking, 1985) there would be perturbations of the spatial curvature around the average value even if the spatial average remained exactly zero.

Obviously this calculation (Hawking and Page, 1986) is only a first step toward showing that Hawking's proposed wave function predicts an observed universe that is very large. First, the choice of the superspace measure and factor ordering needs to be better justified, and the difficulty mentioned by Hawking and Page (1986) sorted out. Next, more complex and realistic models than the two-dimensional minisuperspace need to be analyzed. Its validity may be especially questioned when one deduces that most trajectories come from very large ϕ_0, for when ϕ is so large that the energy density surpasses the Planck value, one might expect the quantum fluctuations in the modes that are ignored to make a significant correction to the behavior deduced from the homogeneous, isotropic mode alone (A. Vilenkin, personal communication). But it is exciting to see that something like Hawking's proposal for the wave function of the universe may actually be able to explain from first principles why the universe is so large.

ACKNOWLEDGMENTS

I appreciate the hospitality of the Aspen Center for Physics, where this paper was begun. Financial support was also provided by National Science Foundation grant PHY-8316811 and by the Alfred P. Sloan Foundation.

REFERENCES

Halliwell, J. J., and Hawking, S. W. (1985). *Physical Review D*, **31**, 1777.

Hartle, J. B., and Hawking, S. W. (1983). *Physical Review D*, **28**, 2960.

Hawking, S. W. (1982). In *Astrophysical Cosmology: Proceedings of the Study Week on Cosmology and Fundamental Physics*, p. 563, H. A. Brück, G. V. Coyne, and M. S. Longair, eds., Pontificiae Academiae Scientiarum Scripta Varia, Vatican.

Hawking, S. W. (1984a). In *Relativity, Groups and Topology II*, p. 333, B. S. DeWitt and R. Stora, eds., North-Holland, Amsterdam.

Hawking, S. W. (1984b). *Nuclear Physics B*, **239**, 257.

Hawking, S. W. (1985). *Physical Reviews D*, **32**, 2489.

Hawking, S. W., and Page, D. N. (1986). *Nuclear Physics B*, **264**, 185.

Page, D. N. (1985a). In *Quantum Concepts in Space and Time*, p. 274, C. J. Isham and R. Penrose, eds., Oxford University Press.

Page, D. N. (1985b). *Physical Reviews D*, **32**, 2496.