Is the universe transparent or opaque?

1972 J. Phys. A: Gen. Phys. 5 1722

(http://iopscience.iop.org/0022-3689/5/12/012)

J. Phys. A: Gen. Phys., Vol. 5, December 1972. Printed in Great Britain. © 1972

Is the universe transparent or opaque?

P C W DAVIES†

Institute of Theoretical Astronomy, University of Cambridge, Cambridge CB3 OE2, UK

MS received 20 June 1972

Abstract. We determine the maximal class of conformally flat cosmological models which eventually absorb all the electromagnetic radiation they contain. This is the requirement of the Wheeler–Feynman absorber theory of radiation. Existing literature on the subject is not only unnecessarily voluminous and complicated, but also contains many serious errors. In this paper the condition for complete absorption is more or less written down immediately, and the various cosmological models tested by inspection. Some of the existing errors concerning the refractive index of the intergalactic medium, and the past absorber are examined in some detail. Of the Friedmann cosmologies, only the oscillating models are opaque. A number of ever-expanding nonrelativistic cosmologies, such as the Dirac and the latest Hoyle–Narlikar models, are also opaque. An assessment of the present value of the Wheeler–Feynman theory is given in the light of recent work on its quantization.

1. General remarks

Much effort has been expended in determining which cosmological models are transparent and which are opaque to light on the future null-cone (Hogarth 1962, Hoyle and Narlikar 1964, Roe 1969, Burman 1971a, 1971b). This is a precondition for the application of Wheeler–Feynman electrodynamics. In this paper a very much simplified and more accurate determination of this opaqueness is presented, together with a discussion on the current status of the Wheeler–Feynman theory in the light of its recent quantization by Hoyle and Narlikar (1969, 1971) and Davies (1971, 1972a).

The starting point of most discussion on this topic is the cosmological principle (homogeneity and isotropy). Models which obey this principle are conformally flat, which means that electromagnetic waves propagate along cosmological null geodesics with a sharp wavefront; that is, they do not grow a 'tail' on the inside of the light cone. This is a consequence of the massless nature of the photon (conformal invariance of Maxwell's equations). This assumption enables one to calculate absorptive processes in Minkowski space with an appropriate redshift factor. It is hard to imagine that the existence of a 'tail' will affect the outcome very much,‡ but we shall also assume the cosmological principle here for simplicity, because the model can then be described in terms of a single parameter, the scale factor $R(t)$ which is a function of the cosmic time t alone and appears in the usual Robertson–Walker metric. In order to include in our discussion models in which matter is continually created, we shall also introduce

† From 1 October 1972 at Mathematics Department, King's College, Strand, London WC2, UK.

‡ These tails will in any case be created by local inhomogeneities, that is, small scale 'bumps' in the cosmological metric.

an independent parameter, the matter particle density ρ. For the usual matter conserving models ρ and R are connected by the equation

$$\rho R^3 = \text{constant}. \tag{1}$$

We shall not necessarily assume the field equations of general relativity, but allow for completely general functions $R(t)$ and $\rho(t)$. We can, however, divide the models into two classes. The ever-expanding models, in which $R(t) \to \infty$ as $t \to \infty$ and the oscillating models in which $R(t) \to 0$ for some finite $t > 0$. We are not, of course, interested here in the behaviour in the region $t \simeq 0$.

We shall first of all deal with the ever-expanding models. The thermal evolution of these models for large t has been fully discussed in Davies (1972b, to be referred to as I). The galaxies, which we shall assume have a geometric cross section which tends to a finite limit (maybe in the form of black holes) as $t \to \infty$, will eventually burn themselves out, after which the total photon number per unit comoving volume can only decrease. The intergalactic medium, if it exists, can only absorb radiation at long wavelengths if it contains some free ions.† As we have an infinite time available for the absorption to occur the actual proportion of ionized atoms is not important, only the way in which their number density decreases with time. In I it was shown that for models which expand more slowly than $t^{1/2}$ all ions eventually recombine and the gas becomes totally transparent. This conclusion can only be avoided if we postulate some sort of charge excess, either locally or globally, or alternatively resort to the continual creation of ions.

As is well known, under the adiabatic expansion of the universe, the matter cools faster than the radiation (see, for example, Sciama 1971):

$$T_{\mathrm{m}} \sim \frac{1}{R^2} \tag{2}$$

$$T_\gamma \sim \frac{1}{R} \tag{3}$$

where T_{m} and T_γ are the respective temperatures of the matter and radiation. The neutral matter is uncoupled, but the ionized matter will be coupled to the radiation to a greater or lesser extent depending on the interaction process and the expansion rate. Generally, therefore, the ions will tend to cool more slowly than the neutral matter by absorbing radiation:

$$T_{\mathrm{i}} \sim \frac{1}{R^p}, \qquad 1 \leqslant p \leqslant 2 \tag{4}$$

where T_{i} is the temperature of the ions.

We shall assume that the radiation is black body, with a characteristic frequency ω, which obeys the redshift relation

$$\omega \sim \frac{1}{R}. \tag{5}$$

We know also, of course, that $\omega \simeq kT_\gamma$ where k is the Boltzmann constant and we are using units such that $\hbar = c = 1$. The black-body nature of the radiation is a good assumption at the present time (most of it is in the form of 3 K background), but may

† As $R \to \infty$ all photons become long wavelength because of the redshift.

not be so good if the accumulated starlight eventually dominates. In any case the ratio of the density of photons ρ_γ to the density of material particles (neutral and ionized) is very large at this time ($\rho_\gamma/\rho \simeq 10^7$) and is unaffected by the expansion ($\rho_\gamma R^3$ = constant also). After the galaxies have cooled and most of the intergalactic matter recombined, the photon–ion ratio will be enormous, and it is usually a very good approximation to assume that the photon number $\rho_\gamma R^3$ is constant when discussing absorption, that is, we can ignore the heat loss to the radiation by absorption.

2. Search for opaque models

The major criticism of all previous work on this subject is its extraordinary complexity. Firstly, all the above workers transform the Robertson–Walker metric to flat space, and work in an unfamiliar coordinate system which is, moreover, different for each model. The transformed time coordinate is not intuitively suggestive, and its range of values varies from model to model. It is necessary to introduce both the space curvature and the notion of horizons explicitly into these discussions to obtain the appropriate functions and limits. In fact, as we shall see, both curvature and horizons are completely irrelevant to the issue of complete absorption.

Having complicated matters with an unnecessary transformation, these authors then proceed to discuss the absorption process in terms of refractive index, which introduces further unnecessary theory, and a batch of complex looking equations, with a lot of riemannian geometry.

The notion of an opaque universe was discussed in I, where it was shown how one could more or less write down the absorption condition straight away from thermodynamic considerations, and check each cosmological model by inspection.

The argument is trivial. The ions lose thermal energy by the expansion at a rate $\frac{3}{2}k\,\mathrm{d}T_i/\mathrm{d}t \propto -\dot{R}/R^{p+1}$ per ion. If they gain energy from the radiation faster than this then all the radiation will eventually be absorbed, and the universe will be opaque. Under these conditions we shall have thermal equilibrium between the matter and the radiation, so we can put $p = 1$ (remembering the high thermal capacity of the radiation compared to the ions; the ions follow the radiation temperature rather than *vice versa*). The absorption rate is $E\sigma$ where E is the energy density of radiation and σ is the absorption cross section. For long wavelengths only inverse bremsstrahlung (collisional damping) is important. This has an effective mean cross section given by (see I for a derivation of this):

$$\sigma \sim \frac{\rho}{T_i^{1/2}\omega^3}\left\{1-\exp\left(-\frac{\omega}{RT_i}\right)\right\}. \tag{6}$$

For $p = 1$, the exponential factor in (6) is constant, so $\sigma \sim \rho R^{7/2}$. Also $E \sim R^{-4}$, which follows from the relation (3) and the Stefan–Boltzmann law $E \propto T_\gamma^4$. We have, therefore, for the matter conserving models

$$E\sigma \sim \frac{\rho_i}{R^{1/2}} \sim R^{-7/2}. \tag{7}$$

The condition for complete absorption is therefore that, after a sufficient time $R^{-7/2} \geqslant \text{constant} \times \dot{R}/R^2$. Integrating this we obtain

$$R \sim t^{2/5} \qquad \text{(or slower).} \tag{8}$$

Thus, in a few lines, we can reproduce the result of Hogarth's original paper, except that the treatment here applies not only to the zero curvature models considered by Hogarth, but to all models, even those with curvature. Notice incidentally, that Hogarth obtained the wrong answer. He found the limit $R \sim t^{1/4}$ instead of (8) because he ignored the temperature dependence of the collisional damping.

The most important feature of the result (8) is that all such models are below our limit $t^{1/2}$ for complete *recombination*, so in the absence of some special postulate such as creation we have the important conclusion: *the intergalactic medium of all ever-expanding cosmologies is transparent for large R.*

In fact, it is very strange that all previous writers have concentrated on intergalactic absorption when the obvious place for absorption to occur is in the galaxies. There is in fact, no evidence for any intergalactic absorption and, actually, no evidence for an intergalactic medium.

Absorption by galaxies is even easier to discuss. The macroscopic objects (stars, black holes, etc) of each galaxy will absorb their geometric cross section of photon flux, irrespective of frequency ω. A given object will absorb photons at a rate proportional to $\rho_\gamma \sim 1/R^3$ so that provided R does not increase faster than $t^{1/3}$ the total number of photons absorbed will diverge, that is, absorption will be complete. Note that this is also the limit for Thomson scattering of each photon, because the cross section for this is also independent of frequency, being simply the geometric classical size of the charged particle. These remarks are unaffected by recombination arguments, so we have the second important conclusion: *all cosmologies which expand asymptotically faster than* $t^{1/3}$ *are transparent.*

In the absence of creation etc, the only model of interest which behaves as an opaque cavity is the Dirac model, $R \sim t^{1/3}$.

It may be felt that we have made some unjustifiable assumption in the above treatment of the absorption theory, so we will now derive the same results in a different way which is even simpler, and, moreover, easier to compare with the earlier treatments. In fact one can easily write down the condition for an opaque universe straight away. The probability that a given photon will be absorbed in a time $\mathrm{d}t$ is $\rho\sigma\,\mathrm{d}t$ (strictly $1-\mathrm{e}^{-\rho\sigma\mathrm{d}t}$) where ρ is now the density of absorbing matter only and σ its absorption cross section. Absorption will be complete, and the model opaque, if the integral of this quantity diverges

$$\int^{\infty} \rho\sigma\,\mathrm{d}t = \infty. \tag{9}$$

Provided $\rho\sigma$ decreases slower than $1/t$ then the model is opaque. For the matter conserving cosmologies this requires σ decreasing slower than R^3/t by (1). In the case of intergalactic absorption we have seen that $\sigma \sim \rho R^{7/2} \sim R^{1/2}$ so that we have complete absorption if $R^{1/2} \geqslant \text{constant} \times R^3/t$, that is $R \sim t^{2/5}$ or less, which is just the condition (8). For galactic absorption $\sigma = \text{constant}$, so we require $\rho \sim t^{-1}$ or less, that is $R \sim t^{1/3}$ in the matter conserving case. Indeed, it is now easy to see why the Dirac model is the limiting case here. This model is founded on the requirement that the well known large number coincidence $\sqrt{N} \simeq mt$ be true for all t, where m is an elementary particle mass unit and N the total particle content of the world. Converting this to a particle density, we have $\rho \simeq m^2/t$, that is $\rho \sim t^{-1}$ which is just the above absorber condition.

There is a helpful interpretation of (9) available. Suppose we have a spherical shell of radius r, thickness $\mathrm{d}r$, concentric on a source of radiation. The proportion of the area of this shell occupied by the absorbing cross section of matter is $\rho\sigma\,dr$. Note here that this proportion is independent of the space curvature, even though this curvature affects the *total* area of the surface at a given distance r. This is because the solid angle distortion for the absorbing region is the same as for the whole surface†. Replacing dr by dt, the light travel time across the shell, the criterion for perfect absorption is that the integrated cross section cover the sky, that is, just condition (9). Put this way the problem is rather like Olber's paradox in reverse. The reader should not be confused by the equality $dr = dt$. We are working in a *coordinate* volume not a *comoving* volume. In the latter case we should have $dr = dt/R$, ρ = constant, but σ (which does not expand with the universe) would be multiplied by a factor $1/R^2$ for area, giving the same result.

Inspection of (6) and (9) also shows that the steady-state model‡ is a complete absorber (though perhaps in only a restricted sense to be discussed below) because ρ and T_i are both constant. This result, which also follows immediately from inspection of an expression which can simply be written down, has taken many pages of working in previous treatments.

There is one issue which has to be checked in the above argument. If we substitute the general behaviour of the ion temperature described by (4) into the cross section (6), then use *this* cross section in our condition (9), the absorber requirement becomes (noting that $\exp(-\omega/kT_i) \to 0$ for large R when $p < 1$) $R^{p/2-3} \sim t^{-1}$ or slower. If we assume $R \sim t^n$ then this becomes the inequality

$$n \leqslant \frac{1}{3-\frac{1}{2}p}. \tag{10}$$

When $p = 1$ we recover the $t^{2/5}$ law from (10). If we ignore for the moment the heating effect of absorption and put $p = 2$ in (10) we have $R \sim t^{1/2}$. It may be imagined that by expanding faster than $t^{2/5}$, the increased cooling rate of the ions might raise the cross section sufficiently in expression (6) to complete the absorption. That this is not so follows from a simple argument given in I which shows that the cooling rate for $R \sim t^n$ where $1 > n > \frac{2}{5}$ is given by

$$\tfrac{3}{2}p = 4 - \frac{1}{n}. \tag{11}$$

The inequality (10) and equation (11) only have a single solution: $n \leqslant \frac{2}{5}$, $p = 1$. This is shown diagrammatically in figure 1 where we have plotted n against p for (10) and (11). Note that $p = 1$ for all $n \leqslant \frac{2}{5}$ and $p = 2$ for all $n \geqslant 1$. Of course, if the initial thermal capacity of the radiation was not so high compared to that of the ions, the cooling rate (10), which is derived by neglecting the heat loss by the radiation, is invalid, and it is plausible that complete absorption could take place for all $n \leqslant \frac{1}{2}$. In practice, the approximation is a very good one and the $t^{2/5}$ limit very nearly correct.

3. The refractive index

The integrand $\rho\sigma$ of (9) is called the absorption coefficient. It is sometimes expressed in terms of the imaginary part of the refractive index, K by the relation $\rho\sigma = K\omega$.

† If half a rubber sheet is painted white, and the sheet is stretched, half of it is still white.

‡ We refer here to the strict steady-state cosmology of Bondi and Gold, not the C-field cosmology of Hoyle, which, as we shall see, is of no interest here.

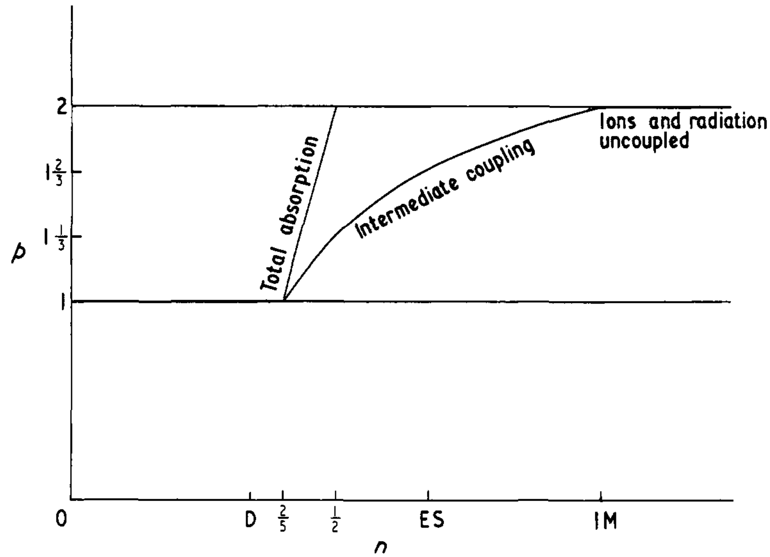

Figure 1. Ionic cooling rates: $T_i \sim R^{-p}$ where $R \sim t^n$.

Suppose we rewrite expression (6) in the form

$$\sigma = \frac{\omega_p^2 \nu}{\omega^2 \rho} \tag{12}$$

where ω_p is the plasma frequency of the ionized medium, given by $\omega_p^2 = 4\pi\rho e^2/m$ and

$$\nu \propto \frac{e^4 \rho}{m^{1/2} T_i^{1/2} \omega}\left\{1 - \exp\left(-\frac{\omega}{kT_i}\right)\right\} \tag{13}$$

and is the effective collision frequency of a given ion with heavy particles†. The numerical factors associated with σ are also absorbed into ν, so that (12) may be written as an equality.

Conventionally, however, K not only includes true absorption processes such as inverse bremsstrahlung, but also scattering processes, such as Thomson scattering, which removes energy from the primary beam, but does not represent an overall energy loss. We must therefore add to σ the Thomson scattering cross section (the only important source of scattering to survive in the limit $\omega \to 0$) $8\pi e^4/3m^2$. We shall rewrite this as

$$\frac{1}{\rho}\left(\frac{4\pi\rho e^2}{m}\right) \times \frac{2e^2}{3m} = \frac{\omega_p^2 \tau}{\rho}$$

where

$$\tau = \frac{2e^2}{3m}$$

is the electron relaxation time ($\simeq 10^{-24}$ s). The total effective cross section σ_{eff} is then

$$\sigma_{\text{eff}} = \frac{1}{\rho}\left\{\frac{\omega_p^2}{\omega}\left(\omega\tau + \frac{\nu}{\omega}\right)\right\} \tag{14}$$

† The overwhelming contribution to the opacity comes from electrons.

so that

$$K = \frac{\rho\sigma_{\text{eff}}}{\omega} = \frac{\omega_p^2}{\omega^2}\left(\omega\tau + \frac{\nu}{\omega}\right). \tag{15}$$

This is the formula used by previous workers who derive it from the classical theory of the damped motion of charged particles oscillating in a periodic electromagnetic field. The formalism is not really appropriate to the region that is being considered here, that is $\omega \simeq kT_i$ or $\omega \gg kT_i$, because ν is then a function of ω. However, for $\omega \ll kT_i$, which is in fact the usual situation of interest in propagation through plasmas, we may expand the exponential in (13) to give, approximately

$$\nu \propto \frac{e^4 \rho}{m^{1/2} T_i^{3/2}} \tag{16}$$

which is independent of ω, although it is true that for completely opaque models ω/kT_i is constant, so that both (13) and (16) have the same dependence on R, that is $\nu \sim R^{-3/2}$.

When $\omega \to 0$ (ie $R \to \infty$), ν/ω tends to zero like $\omega^{1/2}$, so that (15) becomes

$$K \to \frac{\omega_p^2 \nu}{\omega^3}. \tag{17}$$

This formula was used by Hogarth (1962) who referred to inverse bremsstrahlung as collisional damping. He ignored the T_i dependence of ν, simply putting $\nu \propto \rho$. Consequently he obtained $K\omega \propto \omega^4 \sim R^{-4}$ which, using (9), gives the limit for complete absorption as $R \sim t^{1/4}$, rather than the correct result of $R \sim t^{2/5}$. The situation was improved by Burman (1971), who included a temperature dependence of ν, but who took the wrong limit $\omega \gg kT_i$. Consequently he used the formula (16) instead of (13), which would not have mattered if he had not also used the wrong temperature dependence of T_i, that is he took $T_i \sim 1/R^2$. This gives $K\omega$ = constant, which leads to the impressive but erroneous result that all models that expand slower than $R \sim t$, that is all models with particle horizons, are complete absorbers. This includes, of course, nearly all the general relativity models; in particular, the Einstein–de Sitter model.

Further confusion was introduced by Hoyle and Narlikar, who omitted the second term in the brackets of the right hand side of equation (15) entirely, simply taking $K = \omega_p^2\tau/\omega^2$ and referring to this as radiative damping. We have already seen how this term does not represent true absorption at all, but merely the Thomson scattering of the radiation. It is not clear that this is sufficient for the Wheeler–Feynman theory.

In fact, what is normally understood to be radiative damping has not yet been included in any of the above expressions (it is normally completely negligible). To do so we introduce into the denominator of (15) the familiar damping term to give

$$K = \frac{(\omega_p^2/\omega^2)\{\omega\tau + (\nu/\omega)\}}{1 + \{\omega\tau + (\nu/\omega)\}^2}. \tag{18}$$

If we put $\nu = 0$ in (18) we see that for Thomson scattering alone the radiative damping term $(\omega\tau)^2$ tends to zero in the low frequency region and is hence inconsequential. However, we now have a good reason why we cannot neglect the collisional damping in the steady-state theory, as did Hoyle and Narlikar. As $\omega \to 0$ the $(\nu/\omega)^2$ term in the denominator of (18) will dominate, because ν = constant in the steady-state theory. In the matter conserving cosmologies there is no problem as $\nu \to 0$ faster than ω (in fact, like $\omega^{3/2}$).

Actually, formula (18) only applies in the region $\omega \gg \omega_p$. For low frequencies it must be replaced by the general formula for the complex refractive index (eg Ginzburg 1964)

$$(\mu - \mathrm{i}K)^2 = \left(1 - \frac{\omega_p^2/\omega^2}{1+\{\omega\tau+(\nu/\omega)\}^2}\right) - \frac{\mathrm{i}(\omega_p^2/\omega^2)\{\omega\tau+(\nu/\omega)\}}{1+\{\omega\tau+(\nu/\omega)\}^2}. \tag{19}$$

When $\omega < \omega_p$ the real part of the right hand side of (19) becomes negative definite, indicating that reflection is taking place. The wave cannot propagate through the plasma. Suppose for a moment that $\nu = 0$. There is an evanescent wave which exponentially decays into the plasma with skin depth ω_p^{-1}. This does *not* represent absorption by the plasma, but was taken as such by Hoyle and Narlikar (1963)†. To obtain the true absorption we must include a nonzero damping term ν in (19). We then obtain, in the extreme low frequency limit

$$K \simeq \frac{\omega_p}{\sqrt{(2\omega\nu)}} \tag{20}$$

a result which follows from taking the square root of (19) in the limit $\omega \to 0$.

If we accept formula (20) as the true expression for K as $\omega \to 0$ in the steady-state model, then this model is apparently *not* opaque in the strict sense. This is clear from the condition (9) by putting $\rho\sigma = K\omega \sim \omega^{1/2} \sim R^{-1/2}$ and noting that $R \sim e^{Ht}$. Physically we interpret this situation as follows. The radiation becomes increasingly redshifted until it drops below the constant plasma frequency ω_p. It then becomes trapped by reflection in the intergalactic cloud, slowly dissipating by collisional absorption, but apparently not completely. Thus, although the radiation cannot freely propagate to spatial infinity, nevertheless *complete* absorption for this model has not yet been demonstrated.

4. Oscillating cosmologies

These models can be immediately dealt with. Transforming condition (12) to an integral over R:

$$\int^0 \frac{\sigma}{R^3\dot{R}}\,\mathrm{d}R = \infty. \tag{21}$$

As we approach the singularity $R \to 0$ and $\omega \to \infty$. At very high photon energies pair production is the dominant absorption process, the cross section of which remains constant as $\omega \to \infty$. The integral in (21) therefore diverges if R collapses faster than $t^{1/3}$. In the general relativistic radiation dominated Universe, which is applicable in this region, $R \sim t^{1/2}$ and we clearly have complete absorption as the singularity is approached. This describes the limiting behaviour of the oscillating Friedmann models. The result is in agreement with that of Roe (1969) who however, rejected the oscillating models as candidates for Wheeler–Feynman theory on other grounds (see discussion below on the past absorber). Once again, Roe's paper is very lengthy and complicated, requiring explicit discussion of both curvature and horizons, whereas in fact the result follows immediately from the integral in (21).

† This situation does not arise in the matter conserving cosmologies where $\omega_p/\omega \sim \omega^{1/2} \to 0$ so ω is always $\gg \omega_p$ eventually.

5. Exotic models

Although the oscillating Friedmann models apparently satisfy the absorber condition, none of the ever-expanding models do. Of the nonrelativistic cosmologies which do only the Dirac model has any sort of theoretical foundation. This paucity of opaque evolutionary models suggests that we explore certain exotic possibilities.

5.1. Particle creation

As the absorption coefficient is proportional to the density of absorbing material, the continual creation of new matter will obviously permit opaqueness in more rapidly expanding models. To discuss this topic, we assume that $R \sim t^n$, $\rho \sim t^{-m}$; $n, m > 0$. For the matter conserving models $m = 3n$.

Consider first the intergalactic absorption. We have seen that the integrand of (9) is proportional to $\rho^2 R^{7/2}$ in this case, so that the absorber requirement becomes

$$2m - \tfrac{7}{2}n \leqslant 1. \tag{22}$$

In the case of galactic absorption we have already seen how the limiting cases are those models which obey the large number coincidence $\sqrt{N} \simeq mt$. In the case of the Dirac model, the increase in N demanded by this relation is due to particles coming over the particle horizon. The distance of this horizon is proportional to $\dot{R}^{-1}$, so that $\dot{R} \sim t^{2/3}$, that is $R \sim t^{1/3}$. We could also maintain the large number coincidence by creating particles at a faster rate of expansion. For instance, in the latest Hoyle–Narlikar cosmology (Hoyle and Narlikar 1972) $R \sim t^{1/2}$ so that, using the coincidence in the form $\rho \propto 1/t$ we have $\rho R^3 \sim t^{1/2}$. Thus, new particles are continually created in a comoving volume at a rate proportional to $t^{1/2}$ (for large t we may neglect the contribution from the particle horizon). In fact, one can construct a whole class of models from the Hoyle–Narlikar theory which lie on the line

$$n = \tfrac{1}{6}m + \tfrac{1}{3}. \tag{23}$$

The line (23) is easily deduced. The Hoyle–Narlikar conformal gravitational theory requires that the 'coincidence' $G\rho t^2 \simeq 1$ be satisfied for all times t. A second requirement is that the newtonian force law between macroscopic bodies remains unchanged, that is $G\rho^2 R^6 \sim$ constant. Eliminating G between these two relations gives a connection between R and ρ: $R^3 \sim t\rho^{-1/2}$. Putting $R \sim t^n$, $\rho \sim t^{-m}$ leads to (23). The two adjustable parameters of this theory are G and ρR^3 (rate of creation), which may display arbitrary time dependence. The theory includes the Einstein–de Sitter ($G, \rho R^3 =$ constant) and the new model discussed above, chosen from the line (23) by its intersection with the second number 'coincidence' $\sqrt{N} \simeq mt$ (ie $\rho \sim 1/t$).

5.2. Time dependence of the fine structure constant

The inverse bremsstrahlung cross section is proportional to the fine structure constant e^2. A time increase in this quantity would enable the more rapidly expanding cosmologies to absorb. However, as radiative recombination and inverse bremsstrahlung are inverse processes, they depend on the same power of e, and hence a sufficient rate of increase to render, say, the Einstein–de Sitter model opaque, would also bring about total recombination in this model.

The only way in which this total recombination can be avoided is by postulating some form of charge excess, for instance as an inequality in electron and proton numbers, or a polarization due to electric fields 'frozen in' to the cosmological metric.

6. Results

In figure 2 we have plotted R against ρ in the form n against m ($R \sim t^n$, $\rho \sim t^{-m}$). The matter conserving models lie on the oblique line $m = 3n$. The regions above and below this line contain the matter creating and annihilating models respectively. The horizontal line $n = 1$ denotes the uniformally expanding model (Milne model). Above this line the models accelerate and possess an event horizon; below the line they decelerate and possess a particle horizon. The horizontal line $n = \frac{1}{2}$ divides the models which do and do not display total recombination (this does not apply to those with creation, so this line is terminated on the $m = 3n$ line). Several well known models are marked for reference: Milne (M), Einstein–de Sitter (ES), Dirac (D) and Hoyle–Narlikar new model (HN). The broken oblique line divides the transparent from the opaque models.

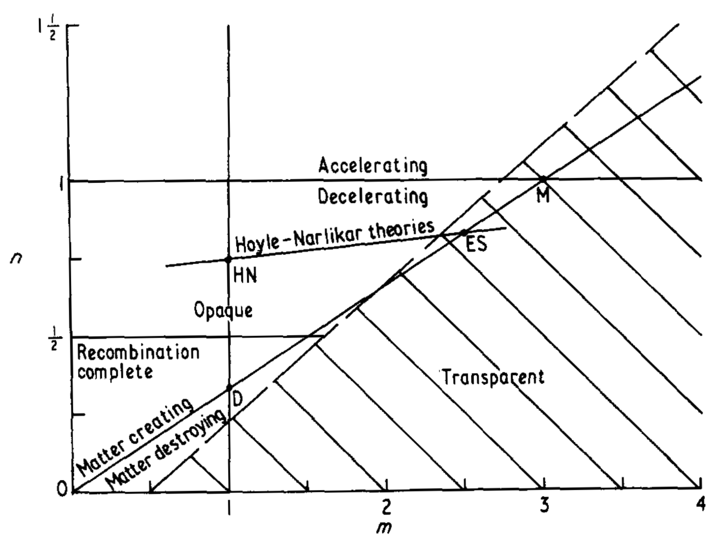

Figure 2. The cosmological models are divided into transparent and opaque classes. The closed Friedmann model is not included. Only the Dirac and latest Hoyle–Narlikar models are opaque†.

7. Wheeler–Feynman absorber theory

It has been known for many years that the photon propagator contains the information necessary to construct amplitudes for real photon processes (Feynman 1948, 1950). This arises from the observation that what may be an external photon line on a Feynman graph describing a limited system, may become an internal line when a larger system is considered. The appellation 'real' is an expression or preference that we wish to ignore distant matter that has no relevance to the particular process being contemplated.

† Correction: the gradient of the broken line should be greater, so that it intersects the solid oblique line at $n = \frac{2}{5}$.

The mathematical redundance of the real photon is only translated into a physical redundance if all such photons are absorbed somewhere. That is to say, if the universe behaves like the inside of an absorbing cavity.

If the universe does behave like a light-tight box, we may make a remarkable mathematical transformation. The pole in the photon propagator, which normally describes the propagation of real (on shell) photons between currents may be subtracted out without any change in the results, provided it is subtracted from *all* such propagators in the whole system. Physically this means that the photons become only auxiliary particles—they do not have their own degrees of freedom. The use of real photons is then merely a way of keeping track of the influence of very distant particles. A classical version of such a direct interaction electrodynamics was developed by Wheeler and Feynman (1945, 1949) and enjoys the distinct advantage of being free of self-energy divergences. This is because self-action can be described in this theory in terms of the response of the absorber, rather than the action of a world line on itself. A fundamental feature of the theory is that the propagator contains symmetrically both advanced and retarded solutions, thus replacing the usual *ad hoc* boundary conditions on the electromagnetic field with a boundary condition on the absorbing particles instead. That is, the apparent nonexistence of advanced radiation imposes a constraint on our cosmological models. This is a very attractive situation, and has led some cosmologists to favour the Wheeler–Feynman theory over the Maxwell theory. However, original hopes that the ultraviolet divergences would disappear from the quantized version of the theory have not been realized, because it appears that some form of self-action must be introduced in the quantized version to account for effects such as the Lamb shift. This self-action immediately re-introduces the divergences (Davies 1971, 1972a).

However, in the author's view the theory still possesses certain advantages over the conventional one:

(i) There is no infinite zero-point field energy.

(ii) We may choose a modified photon propagator which damps out the divergencies without having to develop an appropriate dynamical theory of the field, as in the conventional case.

(iii) It is more economical in basic postulates. In particular, if the universe is a light-tight box, there is no need for an *ad hoc* assumption about free field solutions of the Maxwell equations.

(iv) The theory possibly provides a connection between the thermodynamic, cosmological and electrodynamic arrows of time.

(v) We may learn something about quantum measurement theory by coupling microsystems to cosmology and irreversible thermodynamics in this way.

In connection with item (v) it is interesting to note that the quantum theory was developed originally because classical electrodynamics predicted an infinite energy residing in the electromagnetic field, and nonstationary atomic orbits in which the electron would collapse into the nucleus emitting an unbounded amount of energy in an ultraviolet catastrophe. But in Wheeler–Feynman theory an atom can only emit bounded amounts of energy to the absorber, and the quantum theory need not be invoked for a cavity in thermal equilibrium.

8. The past absorber

In their original paper, Wheeler and Feynman found that a static euclidean universe permitted the existence of both selfconsistent fully advanced as well as retarded solutions†. They appealed to thermodynamics in order to discard the advanced solutions. Thus they assert that the electrodynamic arrow of time is a consequence of the thermodynamic arrow. Of course such a universe would eventually reach thermal equilibrium at a rather high temperature, and the distinction between advanced and retarded radiation would disappear, because all correlations would be destroyed. In practice the cosmological expansion prevents this situation from coming about. However, subsequent authors (Hogarth 1962, Hoyle and Narlikar 1964, Roe 1969, Burman 1971a, 1971b) have assumed a direct connection between the cosmological expansion and the existence of a radiation asymmetry, without any appeal to thermodynamics. Indeed, it is asserted that the thermodynamic asymmetry follows as a consequence of the electrodynamic asymmetry. In these treatments, the usual absorption formulae are applied unmodified to the advanced radiation which travels into the past, as well as to the usual retarded radiation so that the required transparency of the *past* absorber is not assumed to be a consequence of thermodynamics, but depends only upon such features as the Doppler shift and cosmological expansion. Thus, for instance, although a closed oscillating universe provides a perfect future absorber in this treatment, it is also considered to provide a perfect past absorber, leading to a mixture of advanced and retarded radiation in their view.

There are several reasons why the author finds this point of view unacceptable. Firstly, the existence of thermodynamic asymmetries has nothing to do with electrodynamics. Neutral particles obey the second law just as much as charged particles. Secondly, most thermodynamical theory is nonrelativistic, and the speed of light is treated as infinite. There is then no distinction between 'advanced' and 'retarded', but this does not affect the second law. Thirdly, the notion of retarded and advanced waves is an essentially classical concept. In quantum electrodynamics the photon propagator D_F is perfectly symmetric under time inversion, which merely expresses the fact that if we interchange the labels of 'emitter' and 'absorber' of a photon when we carry out the time inversion then the new situation occurs with equal probability to the old. But in the classical case time inversion changes a diverging to a converging wave. That is, radiation processes in classical theory involve large numbers of photons travelling in different directions in a correlated fashion. The absence of converging waves is a consequence of the lack of correlation between different parts of the universe in the past. This has been called the law of conditional independence by Penrose and Percival (1962).

Consider the well known undergraduate problem of two isolated two level atoms in nonrelativistic interaction, with one atom excited at time zero. There is a well determined probability that after a time t the excitation will have been transferred from the first atom to the second. This probability is of the form

$$\frac{\sin^2\frac{1}{2}(E_1-E_2)t}{(E_1-E_2)^2} \tag{24}$$

where E_1 and E_2 are the initial and final energies of the whole system. Clearly the excitation oscillates between the two forever. It is perfectly time symmetric, and after a long time there is equal probability that either atom is excited. In order to obtain an irreversible transition we must perform an incoherent sum of expression (24) over a

† Such a model is obviously symmetric in time cosmologically.

large number of unexcited final states. Then, by the usual law of statistics, it is overwhelmingly probable that the initially excited atom is in its ground state after a long time ($t \gg (E_1 - E_2)^{-1}$). An essential requirement for this result is, of course, that the interacting systems are in their ground states initially. This is an expression of thermodynamic disequilibrium. If these states were instead arranged in some Boltzmann type distribution corresponding to overall thermal equilibrium of the system, then the transition probability would be balanced exactly by its inverse.

Because this analysis is nonrelativistic there is no discussion of retardation necessary. Nevertheless, the purely *statistical* tendency of energy dissipation (and entropy increase) in nonequilibrium situations has already emerged. When we come to consider a relativistic system, it is clear that the basic requirement that energy radiate away from a hot object will not be affected.

Now consider figures 3, 4 and 5. A represents a bundle of initially excited world lines, and B, C, D represent a cold environment. All three diagrams describe the loss of energy by A to B, C, D, ... In figure 3 this occurs by retarded radiation and in figure 4 by advanced radiation. These are chosen to reduce to figure 5 in the nonrelativistic limit. In figure 3 positive energy is transferred through space, and in figure 4 it is negative energy. It is important to note that the possibility of positive energy converging from distant matter B, C, D, ... on to A is not allowed by the statistics.

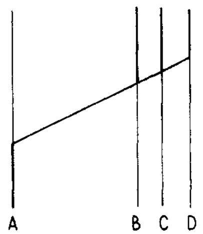

Figure 3. Positive energy radiation passing into the future.

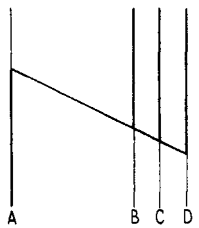

Figure 4. Negative energy radiation passing into the future.

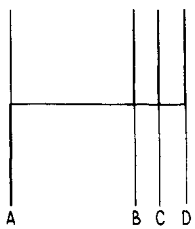

Figure 5. In the nonrelativistic case figures 3 and 4 are indistinguishable.

We only have to explain why figure 4 is not observed. In the author's view, the only reason is that it can lead to paradoxical situations, such as that depicted by figure 6. Suppose we excite the bundle A at time 0, by some nonelectromagnetic process perhaps, causing A to emit a signal (full oblique line) to B, C, D, ... We can now arrange that B, C, D, ... pre-empt the excitation at 0 by signalling into 0's past (broken oblique line). But the excitation is only pre-empted if the original signal is sent, etc and we fall into the usual causality paradox. The solution is therefore not selfconsistent and not allowed. Unfortunately, this does not prove that *all* solutions based on figure 4 are paradoxical.

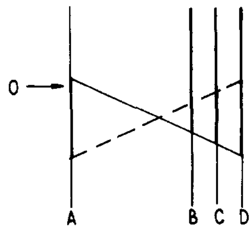

Figure 6. Negative energy radiation leads to paradoxical situations.

The conclusion is that the thermodynamic arrow determines the electrodynamic arrow, and we need not consider the notion of a past absorber. The whole concept of the directionality of time is redundant unless there is a thermodynamic disequilibrium. In practice this tends to be maintained by the cosmological expansion, which thereby becomes a sufficient (but not necessary) reason for the arrow of time to be observed, but in no way *causes* it.

There is, however, one situation in which the past absorber cannot be ignored. This is in a symmetric universe in which entropy increases in both time directions about a given instant t_0 (figure 7). Boltzmann was the first to suggest a cosmology of

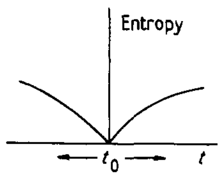

Figure 7. In a time-symmetric universe the entropy increases as we move away from t_0 in both directions of time.

this type, but it also arises in the C-field universe (Hoyle and Narlikar 1964), which tends to the steady-state world asymptotically. Such cosmologies will obviously include mixtures of advanced and retarded radiation. To calculate the mixture we can use a result derived by Hogarth (1962) for imperfect absorption. If the fraction of intensity absorbed on the future and past light cones is $1-\Delta_f$ and $1-\Delta_p$† respectively, then the effective field of a given particle is

$$F_{ret}-\tfrac{1}{2}(F_{ret}-F_{adv})\left(\frac{\Delta_f(2-\Delta_p)}{\Delta_p(1-\Delta_f+\Delta_f/\Delta_p)}\right). \qquad (25)$$

† Note our changes of sign.

For perfect future and past absorbers $\Delta_p \to 0$ and $\Delta_f \to 0$ so that expression (25) becomes indeterminate. If we specify, however, that $\Delta_f/\Delta_p \to \alpha$ where α is some fixed number, then expression (2) becomes

$$\frac{1}{1+\alpha}F_{\text{ret}} + \frac{\alpha}{1+\alpha}F_{\text{adv}}. \tag{26}$$

Figure 8 describes the C-field cosmology which has the following behaviour. At $t = -\infty$ the scale factor $R = \infty$. The Universe then contracts to a minimum value of R, after which it expands again to $t = \infty$. There is continuous creation in this model,

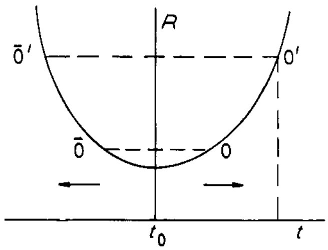

Figure 8. The C-field cosmological model has the entropic behaviour shown in figure 7, which leads to *advanced* radiation in the Wheeler–Feynman theory.

generated by the C-field, and in the regions $t \to \pm\infty$ the situation tends to a steady state. The time t_0 is a symmetry point, and it is usually conjectured that it represents an entropy minimum similar to that depicted in figure 7. Thus entropy would appear to increase as we move away from t_0 in either direction in time so that all observers not at t_0 would claim that the thermodynamic arrow of time was 'normal'. In their 1964 paper, Hoyle and Narlikar claim that retarded potentials occur to the right of t_0 and advanced potentials to the left. If this were so, then the thermodynamic, cosmological and electrodynamic arrows of time would all point the same way, and observers in each limb would apparently see the same situation. To see that this cannot be so consider an observer at 0, on the right hand limb of figure 8. Retarded waves travel to the right, and advanced waves to the left from 0.

Suppose that after a certain time the retarded wave has reached a point 0′ at which the absorption effectively shuts off leaving a fraction Δ_f ($\ll 1$) unabsorbed. Then the advanced wave will also undergo absorption only as far as the mirror point $\bar{0}'$. However this wave will also have undergone some additional absorption between 0 and $\bar{0}$ (we are assuming here, though it does not affect the present argument, that this extra absorption is restricted to the region between the minimum at time t_0 and $\bar{0}$). The unabsorbed fraction at $\bar{0}$ is β, say, and at $\bar{0}'$ is $\Delta_p = \Delta_f\beta$. Thus $\alpha = 1/\beta$, and we see from expression (26) that a fraction $1/(1+\beta)$ of advanced radiation is present, even in the limit $\Delta_f \to 0$ (in fact, it was shown that Δ_f does not tend to zero, though it is exceedingly small). At t_0, $\beta = 1$ and the effective field is $\frac{1}{2}\text{F}_{\text{ret}} + \frac{1}{2}F_{\text{adv}}$. As we move away from t_0 in either direction, $\beta \to 0$ and $\alpha/(\alpha+1) \to 1$, so that the effective field tends to the fully advanced solution. In the asymptotic steady-state region therefore, only fully *advanced* solutions are consistent. It should, however, be pointed out that the advanced waves are blue shifted until they reach the minimum at t_0, and for sources in the asymptotic region this would cause a breakdown in the linearity assumption in the $t = t_0$ region when gravitational effects of these blue shifted photons become important. This might disrupt the whole model so much as to change its basic character. What is clear is that if the C-field cosmology could exist, it could not have substantial retarded radiation anywhere.

References

Burman R 1971a *Observatory* **90** 240–9
—— 1971b *Observatory* **91** 141–54
Davies P C W 1971 *J. Phys.* A: *Gen. Phys.* **4** 836–45
—— 1972a *J. Phys.* A: *Gen. Phys.* **5** 1025–36
—— 1972b *R. Astron. Soc.* in press
Feynman R P 1948 *Rev. mod. Phys.* **20** 367–87
—— 1950 *Phys. Rev.* **80** 440–57
Ginzburg V L 1964 *The Propagation of Electromagnetic Waves in Plasmas* (Oxford: Pergamon)
Hogarth J E 1962 *Proc. R. Soc.* **267** 365–83
Hoyle F and Narlikar J V 1963 *Proc. R. Soc.* **273** 1–11
—— 1964 *Proc. R. Soc.* **277** 1–23
—— 1969 *Ann. Phys., NY* **54** 207–39
—— 1971 *Ann. Phys., NY* **62** 44–97
—— 1972 *Mon. Not. R. Astron. Soc.* **155** 323–35
Penrose O and Percival I C 1962 *Proc. Phys. Soc.* **79** 605–16
Roe P E 1969 *Mon. Not. R. Astron. Soc.* **144** 219–30
Sciama D W 1971 *Modern Cosmology* (Cambridge: Cambridge University Press) chap 12
Wheeler J A and Feynman R P 1945 *Rev. mod. Phys.* **17** 157–81
—— 1949 *Rev. mod. Phys.* **21** 425–33