TIME ASYMMETRY IN ELECTRODYNAMICS AND COSMOLOGY

J. V. Narlikar

Institute of Theoretical Astronomy, University of Cambridge, Cambridge, England

ABSTRACT

The concept of the 'arrow of time' is discussed in relation to thermodynamics, electrodynamics and cosmology. Time symmetric electrodynamics, the absorber theory of radiation and quantum transitions receive attention to develop a working theory applicable to a universe with a perfect future absorber and an imperfect past absorber, such as the steady state cosmological model. It is stated that the approach adopted herein establishes a strong connection between the electrodynamic and cosmological arrows of time, and points a way towards linking these arrows with the thermodynamic one.

---

1. THE ARROWS OF TIME IN PHYSICS

What do we mean by the arrow of time? In a somewhat vague manner, we may associate it with the sense in which we relate our subjective experiences to the space-time manifold. It would be rash to attempt a precise definition that will satisfy all scientists and philosophers! It is possible, however, to discuss the subject in a rigorous manner in the more restricted field of fundamental physics.

Given a space/time diagram, we can arbitrarily choose a direction along the time axis, by identifying one end with the past and the other with the future*. This enables us to arrange events in a chronological order. Consider a set of events A,B,C,... say, which can occur again and again. If these events always occur with the same chronological order we are able to use these events themselves to fix the time sense. This would not be possible if these events occurred in a random order at different times. Our physical experience reveals events of both types—although it is the events of the former type that are responsible for the arrow of time. What are these events?

These events can be broadly grouped under three categories, thermodynamics, electrodynamics and cosmology. As an example of the events in the first category, consider a hot body and a cold body in contact with each other. If we observe the system 'a little later', we find the hot body a little cooler and the cold body a little hotter than 'before'. In other words, if we measure the temperatures of the two bodies at different instants we can order these instants chronologically. Such events are usually known as

---

* I am excluding here cosmological models with closed time-like lines.

irreversible events. In electrodynamics, if we observe an accelerated electric charge we notice that it loses energy. Thus the energy flow from the charge can be used to fix a chronological order. In cosmology the expansion of the universe enables us to do the same. If we photograph two nebulae at instants t_1 and t_2, their separation from each other would tell us whether $t_2 > t_1$ or $t_2 < t_1$. Thus events in each category determine an arrow of time.

The three arrows described above appear to arise from quite different sources. We may raise the question: 'Why do these arrows point the way they do?' In other words, is there any connection between these different arrows of time? Before we search for a solution of the more fundamental question: 'Why an arrow of time?' it may well be profitable to investigate whether the arrows of time in thermodynamics, electrodynamics and cosmology are mutually related.

Opinions differ as to how deep this connection is. Some physicists believe that the connection, if any, is very superficial. I want to discuss the opposite point of view, particularly in connection with electrodynamics and cosmology.

2. TIME SYMMETRIC ELECTRODYNAMICS

Returning to the problem of the radiating charge, we may enquire as to the origin of the time asymmetry. Maxwell's equations are time symmetric. Corresponding to any solution of these equations we may obtain another by changing t to $-t$. Now the problem of radiating charge is solved by means of retarded solutions of Maxwell's equations. If instead we use advanced solutions we would describe a charge *receiving* energy. Mathematically both the solutions (and any linear combination of them) describe the electromagnetic fields associated with an accelerated charge. Only the retarded solution is taken to represent the actual situation. So long as there is no other direction of time we cannot distinguish between the pure retarded and pure advanced solutions. Given another arrow of time, however, say from cosmology, the distinction between the retarded and advanced solutions becomes real. Why do we choose retarded solutions? The usual answer is 'causality'. However, to say that we choose retarded solutions in order to conform with causality does not take us any further in our understanding of the electrodynamic arrow of time. So long as the choice of retarded solutions is made on such arbitrary grounds, we will not be able to make progress in this direction. It appears that to achieve this we have to resort to a different theory of electrodynamics which does not permit so much freedom of choice.

Such a theory already exists in literature, and in its crude form even preceded Maxwell's theory. This is the theory of direct interparticle action. In this theory electric charges interact directly with one another and the electromagnetic field has no independent existence. Mathematically we may describe the theory through an action principle due to Fokker[1]. The action is given by

$$S = -\sum_a \int m_a \, da - \sum_{a < b} e_a e_b \iint \delta(s_{AB}^2) \, da_i \, db^2 \tag{1}$$

Here a, b, c, ... label the charged particles, e_a, m_a being the charge and mass of particle a. A denotes a typical point on the world line of a, da^i denoting

the coordinate displacements and da the proper time element at A, along the world line. We have used flat space coordinates with metric $\eta_{ik} = \text{diag. } (-1, -1, -1, +1)$. The vector indices are raised and lowered according to η_{ik}. The delta-function in equation 1 has as its argument s_{AB}^2, the square of the interval between A and B, two typical points on the world lines of a and b. Thus A and B interact only if they are connected by a light ray. This is the relativistic generalization of the Newtonian concept of instantaneous action at-a-distance. The double sum in 1 does not include the terms $a = b$; the self-action is therefore absent. We have taken the velocity of light as unity.

The 4-potential generated by charge a at a point X is given by

$$A_i^{(a)}(X) = e_a \int \delta(s_{AX}^2) \, da_i \tag{2}$$

and it satisfies *identically* the gauge condition

$$A_{;i}^{(a)i} = 0 \tag{3}$$

and the wave equation

$$\Box A_i^{(a)} = 4\pi j_i^{(a)} \tag{4}$$

where $j_i^{(a)}(X)$ is the current due to charge a at X

$$j_i^{(a)}(X) = e_a \int \delta_4(A, X) \, da_i \tag{5}$$

We may also define the field tensor corresponding to 2

$$F_{ik}^{(a)} = A_{k;i}^{(a)} - A_{i;k}^{(a)} \tag{6}$$

By virtue of relations 3 and 4 this satisfies the Maxwell equations

identically.

This is a fundamental difference between the above theory and the Maxwell field theory. In the latter, Maxwell equations are genuine equations. Thus to a solution of equation 7 we can add *any* solution of

$$F^{(a)iki}_{\;\;;} = 0 \tag{8}$$

On the other hand in the above theory equation 7 is an identity. The $F^{(a)ik}$ are really given by equations 2 and 6. If we denote the usual retarded solution generated by charge a to be $F_{\text{ret}}^{(a)}$, equation 6 corresponds to

$$F^{(a)} = \tfrac{1}{2}[F_{\text{ret}}^{(a)} + F_{\text{adv}}^{(a)}] \tag{9}$$

Thus there is no freedom of choice; we are *forced* to accept equation 9.

It was this lack of freedom that proved to be a stumbling block to the theory for a long time. Charges are *observed* to produce retarded fields; not the time symmetric fields given in 9. If we consider the equation of motion of charge a, obtained by setting $\delta S = 0$ for a small variation of the world line of a, we find that a force like the Lorentz force acts on a. But it arises from the fields of all other particles in the universe, i.e. from a field

$$\sum_{b \neq a} \tfrac{1}{2}[F_{\text{ret}}^{(b)} + F_{\text{adv}}^{(b)}] \tag{10}$$

The advanced fields in equation 10 appear to present an embarrassment to the theory.

However, an ingenious way out of this difficulty was found by Wheeler and Feynman[2] with their absorber theory of radiation.

3. THE ABSORBER THEORY OF RADIATION

Wheeler and Feynman argued that the interaction is time symmetric only between pairs of particles. The possibility still exists that the collective interaction of a large number of particles could lead to time-asymmetric results. The actual universe is a collection of large numbers of charged particles. An asymmetry in the large scale behaviour of the universe is therefore expected to lead to a local asymmetry in electrodynamics. According to Wheeler and Feynman the requirement to be met by the universe in order to produce the necessary asymmetry is that it shall act as a perfect absorber of all electromagnetic disturbances generated within it. That is, if we set the charge a in motion, all the disturbances related to this should be absorbed by the universe:

$$\sum_b \tfrac{1}{2}[F_{\text{ret}}^{(b)} + F_{\text{adv}}^{(b)}] \to 0 \qquad \text{as} \quad r \to \infty \tag{11}$$

faster than $1/r$, where r is the distance from a.

Since advanced and retarded fields have supports on different branches of the light core, expression 11 implies that

$$\sum_b \tfrac{1}{2}F_{\text{ret}}^{(b)} \to 0, \quad \sum_b \tfrac{1}{2}F_{\text{adv}}^{(b)} \to 0 \qquad \text{as} \quad r \to \infty \tag{12}$$

and hence that

$$\sum_b [F_{\text{ret}}^{(b)} - F_{\text{adv}}^{(b)}] \to 0 \qquad \text{as} \quad r \to \infty \tag{13}$$

However, the quantity in 13 has no sources and hence it satisfies the homogeneous wave equation. Since it vanishes at infinity faster than $1/r$, it must vanish identically. Therefore the field 10 acting on the particle a has the form

$$\sum_{b \neq a} \tfrac{1}{2}[F_{\text{ret}}^{(b)} + F_{\text{adv}}^{(b)}] = \sum_{b \neq a} F_{\text{ret}}^{(b)} + \tfrac{1}{2}[F_{\text{ret}}^{(a)} - F_{\text{adv}}^{(a)}] \tag{14}$$

The first term on the RHS denotes the retarded interaction of all other particles while the second term is the radiative reaction in the form given by Dirac[3]. Thus the theory appears to provide the correct answer if the condition 11 of perfect absorption is satisfied.

There was, however, one snag in the argument developed so far. Wheeler and Feynman applied this to the static Euclidean universe with a uniform distribution of charges. Such a universe satisfies 11 and hence leads to 14. However, this universe is time symmetric and hence cannot distinguish between the past and the future light cones. Hence the time-reversed form of 14, viz.

$$\sum_{b \neq a} \tfrac{1}{2}[F_{\text{ret}}^{(b)} + F_{\text{adv}}^{(b)}] = \sum_{b \neq a} F_{\text{adv}}^{(b)} + \tfrac{1}{2}[F_{\text{adv}}^{(a)} - F_{\text{ret}}^{(a)}] \tag{15}$$

is also valid.

Wheeler and Feynman recognized this and argued that the asymmetry

which allows one to distinguish between 14 and 15 comes from thermodynamics. The thermodynamic arrow of time is ascribed to asymmetrical initial conditions. It is these conditions which make 14 highly probable and 15 highly improbable in the sense of statistical mechanics. Thus according to Wheeler and Feynman the electrodynamic and thermodynamic arrows of time are strongly related.

A new feature into this argument was introduced by Hogarth[4] many years later. He pointed out that the above argument ignores the cosmological time arrow. In an expanding universe the past and future light cones do not behave symmetrically where absorption is concerned. Electromagnetic waves going into the future are red-shifted while those going into the past are blue-shifted. Also, in some models of the universe the matter densities in the past and the future are different. Thus one cannot argue that if 14 holds so would 15 or vice versa. Hogarth's calculations show that in certain models of the universe, e.g. the open Friedmann models, 14 does not hold but 15 does. Also in the steady state model 14 holds but not 15. In other words, the observed alignment of electrodynamic and cosmological time arrows could be explained only in certain cosmological models, and not others.

Hogarth's argument had to be improved in two ways, however. First, he had taken over the Fokker action 1 in flat space and used it in the Robertson–Walker cosmological spaces of the expanding cosmological models. Since these spaces are conformally flat, and the electromagnetic equations are conformally invariant, the above step was justified. For completeness it was necessary, however, to rewrite equation 1 in the curved space of general relativity. Secondly, the use of the refractive index, which plays an important part in the calculation, involved thermodynamics which Hogarth was trying to avoid. In a later work Hoyle and Narlikar[5] took account of both these points, although their conclusion did not differ from Hogarth's. I summarize below the main points of the argument.

Suppose we want to arrive at 14. Then the condition to be satisfied by the universe is that it must be a perfect absorber of all disturbances connected with the motion of a along the future light cone of a. That is

$$\sum_b F_{\text{ret}}^{(b)} \to 0 \qquad \text{as} \quad r \to \infty \tag{16}$$

faster than $1/r$. However, expression 16 requires that we have a definite process of absorption. If we stick to pure electrodynamics this is provided by radiative damping. Now the damping term has the correct sign if 14 holds, and the wrong sign (implying growth of the disturbance) if 15 holds. To avoid ambiguity therefore it is necessary to assume that 14 holds to begin with and then to show that the theory is consistent with this assumption. Thus one has to work within a self-consistent cycle of argument.

When the charge a is set in motion, its retarded field sets other charges in motion along its future light cone. Since 16 implies a damping of the retarded field of a as the wave goes away from it, this also implies that the advanced fields arising from the motion of all other charges b provoked by the retarded field of a must also be damped. Hence we have

$$\sum_b F_{\text{adv}}^{(b)} \to 0 \qquad \text{as} \quad r \to \infty \tag{17}$$

faster than $1/r$. Notice that 17 arises from 16 which is the more basic condition. If we were discussing the consistency of 15 we would require 17 to hold along the past light cone of a and then deduce 16. Also, for 17 to follow from 16 it is essential that the motion of charge a be bounded in space-time, a requirement not met by the 'run-away' solutions. Such solutions are therefore not possible in this theory.

From 16 and 17 we then arrive at 14 through steps similar to those used by Wheeler and Feynman. It is instructive, however, to see the precise part played by the universe.

Suppose we are in a universe in which 14 holds. Consider a charge a at time $t = 0$. The future light cone of a has $t > 0$, and the past light cone $t < 0$. Now we rewrite the LHS of 14 in the following form

$$\sum_{b \neq a} \tfrac{1}{2}[F_{\text{ret}}^{(b)} + F_{\text{adv}}^{(b)}] = \tfrac{1}{2} \sum_{\substack{b \neq a \\ t < 0}} F_{\text{ret}}^{(b)} + \tfrac{1}{2} \sum_{\substack{b \neq a \\ t > 0}} F_{\text{adv}}^{(b)} \tag{18}$$

The second term on the RHS of 18 is the advanced contribution of all the particles b on the future light cone of a, which are excited by the retarded disturbance from a. This contribution arrives instantaneously as the retarded wave leaves a. We shall call it the 'response' of the universe and denote it by $R^{(a)}$. A comparison with the RHS of 14 shows that

$$\begin{aligned} R^{(a)} &= \tfrac{1}{2} \sum_{\substack{b \neq a \\ t < 0}} F_{\text{ret}}^{(b)} + \tfrac{1}{2}[F_{\text{ret}}^{(a)} - F_{\text{adv}}^{(a)}] \\ &= \sum_{\substack{b \neq a \\ t < 0}} \tfrac{1}{2}[F_{\text{ret}}^{(b)} - F_{\text{adv}}^{(b)}] + \tfrac{1}{2}[F_{\text{ret}}^{(a)} - F_{\text{adv}}^{(a)}] \\ &= \sum_{\substack{b \\ t \leqslant 0}} \tfrac{1}{2}[F_{\text{ret}}^{(b)} - F_{\text{adv}}^{(b)}] \end{aligned} \tag{19}$$

In 19 we have used the fact that $F_{\text{adv}}^{(b)} = 0$ for $t < 0$.

Equations 19 denotes the response required from the universe in order to arrive at 14 which is in conformity with observations. As mentioned before the steady state universe provides the appropriate response, but not the usual big bang models.

4. QUANTUM CONSIDERATIONS

We have so far considered classical electrodynamics. Can these ideas be extended to quantum electrodynamics? Recent work[6] shows that it is possible to have a quantum theory of direct interparticle action. The shortage of space does not permit a detailed discussion of these new ideas. I will be content with a brief mention of salient features.

The quantum analogue of the radiating electron occurs in the spontaneous transition of an atomic electron. While the induced transitions are both upward and downward with symmetrical rates, the spontaneous transitions are always downward, accompanied by emission of radiation. Can this be related to the expansion of the universe?

In the Maxwellian quantum electrodynamics the phenomenon of spontaneous transition is explained by first quantizing the electromagnetic field. The vacuum state in this theory has non-trivial properties as a result of the

field quantization rules, and these play an important part in arriving at the spontaneous transition probability. For example, if the electron is making transitions between two energy levels E_m and E_n, $E_n > E_m$, in the presence of a field of n quanta of frequency $v = (E_n - E_m)/\hbar$ the probabilities of downward and upward transitions are related by

$$\frac{P(E_n \to E_m)}{P(E_m \to E_n)} = \frac{n+1}{n} \tag{20}$$

In the direct particle theory there are no fields and hence no field quantization or quanta in the above sense. So far as induced transitions are concerned the theory gives the same result as field theory. But what about spontaneous transitions? Returning to the classical expression 14, the first term on the RHS is responsible for induced transitions whereas the second term leads to spontaneous transitions. Since the second term is part of the response $R^{(a)}$ of the universe, it is to the universe that we must look for an explanation of spontaneous transitions.

It turns out that the calculations are most easily performed by using the path integral method of Feynman[7, 8]. This provides the quantum analogue of 14 and 19. For example, the response of the universe is non-zero even when there is no net electric field from all other particles. Thus the electron is influenced by the universe and makes a transition (spontaneously).

The physical reason for this may be understood in the following way. A downward spontaneous transition is accompanied by the emission of an electromagnetic wave which causes transitions in the future absorber. An expanding universe acts as a 'cold environment', with all absorber particles in their ground states. The wave emitted by the source particle a is absorbed by the absorber particles which make upward transitions, induced by this wave. Complete absorption is necessary if the whole cycle is to work self-consistently. It is also easy to see why spontaneous upward transitions do not occur. These would require augmentation instead of absorption of the electromagnetic waves emitted by the source. The universe in the future light cone does not permit this—since the particles are mostly in their ground state and cannot jump down farther to enhance the incoming radiation.

Apart from absorption, the universe also plays an important part through dispersion. If we Fourier analyse the outgoing wave from the source, the different components travel at different velocities through the intergalactic medium. The intergalactic density and other parameters are such that large phase differences are built up between neighbouring frequencies, with the result that the different components become randomly phased. This random phasing is usually *assumed* in field quantization. In the above work this extra assumption is not required.

I have so far not mentioned thermodynamics explicitly. But a connection between the thermodynamic and cosmological arrows now begins to appear. An expanding universe acts as a sink—a fact which appears to be connected with the asymmetry of upward and downward transitions considered above. This asymmetry leads to a result like 20 even in the direct particle theory. It is well known that the law of black body radiation follows from 20. As

yet these ideas are somewhat tentative, but they serve to indicate that a connection between thermodynamics and cosmology may be built up.

5. CONCLUSION

The ideas described above work in a universe with a perfect future absorber and an imperfect past absorber. Of the well known cosmological models only the steady state model meets this requirement. If we regard the present astronomical evidence as unambiguously against the steady state theory we must abandon this approach altogether. On past record of astronomical observations, and with the present uncertainties attached to the recent data, such an unequivocal conclusion is not justified[9].

On the other hand we may ask whether there are any gains in the above approach. As discussed above this approach establishes a strong connection between the electrodynamic and cosmological arrows of time, and also points a way towards linking these arrows with the thermodynamic one. Such a connection cannot be established with Maxwellian electrodynamics which allows too much freedom of choice. Moreover, as the recent work[10] shows this approach is better than the usual quantum electrodynamics in handling the so-called radiative corrections.

I end with a few remarks on the question 'why an arrow of time'. If we were able to establish a strong connection between the different arrows of time, physical, biological etc., we could then argue in the following way. An absolute direction is immaterial: the important concept is the relative orientation of the different arrows. For example, if we can say 'we grow old because the universe expands', then this is equivalent to saying that we grow young as the universe contracts! The question 'Why do we grow older, not younger?' without reference to other arrows of time has no meaning.

REFERENCES

[1] A. D. Fokker, *Z.Phys.* **58**, 386 (1929).
    A. D. Fokker, *Physica,* **9**, 83 (1929); **12**, 145 (1932).
[2] J. A. Wheeler and R. P. Feynman, *Rev. Mod. Phys.* **17**, 157 (1945).
[3] P. A. M. Dirac, *Proc. Roy. Soc. A*, **167**, 148 (1938).
[4] J. E. Hogarth, *Proc. Roy. Soc. A*, **267**, 365 (1962).
[5] F. Hoyle and J. V. Narlikar, *Proc. Roy. Soc. A*, **277**, 1 (1963).
[6] F. Hoyle and J. V. Narlikar, *Ann. Phys. (N.Y.)*, **54**, 207 (1969).
[7] R. P. Feynman, *Rev. Mod. Phys.* **20**, 367 (1948).
[8] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals.* McGraw-Hill: New York (1965).
[9] F. Hoyle, *Proc. Roy. Soc. A*, **308**, 1 (1968).
[10] F. Hoyle and J. V. Narlikar, *Nature, London,* **222**, 1040 (1969).