Q.Jl R. astr. Soc. (1975) 16, 245–264

Cosmology and Verifiability*

G.F.R.Ellis

(Physics Department, Boston University and Department of Applied Mathematics
and Theoretical Physics, Cambridge)†

(Received 1975 April 14)‡

SUMMARY

The way in which one obtains a cosmological model from basic principles is discussed. It is shown that some of these principles are of necessity unverifiable. The observed isotropy of the universe leads to a uniformity assumption which can be codified in various ways, leading to slightly different world models. The overall structure of these world models is discussed in detail, and the different ways in which different parts of the universe are accessible to observation can then be made explicit.

THE COSMOLOGICAL PROBLEM

Relativistic Cosmology aims to determine the structure of the universe from a fusion of the results of astronomical observations with knowledge derived from local physical experiments. The problem of determining this structure (1) is centred on the fact that there is only one universe to be observed, and that we effectively can only observe it from one space-time point. Because it is a unique object, we cannot infer its probable nature by comparing it with similar objects; and (on the scale we are considering) we are unable to choose the time or position from which we view it. Our predicament is analogous to that of a premaritime man living on a small island in an ocean, who observes around him a host of other small islands apparently scattered at random on a seemingly limitless sea. Unable to move from his island, his theory of the world in which he lives can only be based on this partial view.

UNVERIFIABLE ASSUMPTIONS

Given this situation, *we are unable to obtain a model of the universe without some specifically cosmological assumptions which are completely*

---

*Paper delivered at the Boston Colloquium for the Philosophy of Science, 1973 November 6.

†Present Address: Department of Applied Mathematics, University of Cape Town, South Africa.

‡Received in original form 1974 December 30.

unverifiable. Because we wish to talk about regions we cannot directly influence or experiment on, our theory is at the mercy of the assumptions we make. To illustrate this, consider the possibility that our friend on the island might be of a theological disposition; and might have decided that there was in fact only one island in the world, surrounded by an ocean which ended at a beautifully painted diorama constructed by an artistic and kindly God, giving him the illusion of a limitless sea covered by islands. In an exactly analogous way, a modern cosmologist who was also a theologian with strict fundamentalist views could construct a universe model which began 6000 years ago in time and whose edge was at a distance of 6000 light years from the solar system. A benevolent God could easily arrange the creation of this universe not only so that suitable fossils would be present in the Earth (having been created, together with the rest of the universe, 6000 years ago) to imply a long geological history, but also so that suitable radiation was travelling towards us from the edge of the universe to give the illusion of a vastly older and larger expanding universe. It would be impossible for any other scientist on the Earth to refute this world picture experimentally or observationally; all that he could do would be to disagree with the author's cosmological premises.

What has been violated here is the expectation that the ordinary laws of everyday physics, carefully and correctly applied, will lead to correct inferences about what exists; for such application of these laws leads (in these unusual universes) to expectations different from what would actually be there. To exclude this possibility, we invoke the first of our unverifiable assumptions about the universe: *whenever normal physical laws can be applied, they correctly predict the structure of the universe.* I shall call this the *local predictability assumption.* There are two facets to this requirement: firstly, that the normal physical laws we determine in our space-time vicinity are applicable at all other space-time points. This demand for uniformity in Nature is necessary for reasonable predictions to be made about distant parts of the universe; for otherwise there is too much arbitrariness in what we can suppose. Without this guide, we have no suitable set of rules to tell us what to expect. In any case we have a set of physical laws which are locally valid, and the established scientific policy—based on the 'Occam's razor' or 'minimal assumption' attitude— is to continue extrapolating, applying these laws in larger domains and to more distant points, unless something makes it clear that this is the wrong procedure. Note that this does not exclude, for example, theories in which the gravitational constant varies; all this amounts to (assuming we know the law which determines how this 'constant' varies) is that local physical laws are rather more complex than we might originally have supposed. The second aspect of this statement is the implication that we keep on applying these laws as long as this is possible; and the resulting expectations are fulfilled. It is this aspect (expressed mathematically in the requirement that space-time be inextendible (2)) which prevents our universe model beginning or ending at an edge such as the one described above.

Having adopted this principle, one might hope it would not be necessary to make further unverifiable assumptions in order to obtain a reasonably unique cosmological model from our observations. However, the nature of the observations we are able to make prevents fulfilment of this hope. Two facts lead to this conclusion. First, we are unable to examine directly space-time itself, or the distribution of matter in it; rather we observe particular objects—stars, galaxies, quasi-stellar objects, dust, and so on—in space-time, and only when we have somehow determined their intrinsic properties can we deduce their distribution and the properties of the intervening space-time (3). Second, we simply do not have the astrophysical information needed to determine their nature sufficiently accurately. This is partly because there is a wide variation in the properties of individual objects in each class; partly because we simply do not understand the nature of some of the classes of objects we are observing; and to a very large extent because the light we receive from the more distant objects was emitted a long time ago. Thus we need to have a satisfactory theory of their time-evolution in order to determine their intrinsic properties at the time they emit the light by which we observe them. We do not have such a theory. So, for example, having obtained measurements of the radio brightness of a radio source, we are unable to determine directly from our measurements whether we are receiving radiation emitted from a bright source a long time ago, or from a weaker source which is relatively nearby, or from a weak source which is very far away but appears anomalously bright because of the curvature of the intervening space-time. The situation is similar to that of the isolated man on the island if he is able to measure accurately the apparent sizes of the other islands, but does not know their intrinsic sizes. Any particular island he sees might be a small one nearby, or a much larger one a long way off. A new principle is needed to order the observations.

As presented thus far, the argument may sound rather weak; it may seem that introduction of a new principle is a counsel of expediency rather than necessity. Might it not be that given sufficient time for increased understanding of the astrophysics involved, the problem would eventually simply go away; for then we would have sufficient information to use the observed objects as ‘standard candles’ which could be reliably used to chart the universe? This is most unlikely to be the case, not only because of the nature of the difficulties encountered in astrophysics, but because of one fundamental aspect of our present knowledge of the universe which has not been mentioned so far.

This crucial feature is that the universe appears to be *isotropic about us* to an extraordinary accuracy. In particular the number counts of distant radio sources show that their average distribution is the same in all directions; the X-ray background radiation is isotropic to better than 5 per cent; and the microwave background radiation is isotropic to better than 0·2 per cent (4). No matter what direction we choose in order to obtain information about the large-scale structure of the universe, we obtain the same answer as for any other direction. Thus there seem (on a cosmological scale) to be no preferred directions about us; we are unable to point in a certain direction and say 'the centre of the universe lies over there'; in fact we are unable to say that any direction is particularly different from any other.

To consider the consequences of this, suppose our astonished friend on his island found that his observations lead to the same conclusion. He would then be able to use this fact to construct for himself models of his world, even though he did not know the distances of the islands he observed. He would, after a while, discover there were two possible situations. Either the islands could be scattered uniformly over a uniform ocean in such a way that all islands were roughly the same distance from the island nearest to them, and so that the world looked very much the same to any observer on any island (*cf.* Fig. 1(a)); or they could be distributed in some other way, for example with all the islands that looked smaller a much smaller distance from their nearest neighbours than all the islands that looked larger (*cf.* Fig. 1(b)). The common feature of all these other ways of arranging the islands would be that they were all centred on his own island; by measuring the positions of all the islands in the sea one would with complete certainty deduce that his own island was at the centre of the visible part of the world. Although he himself would not be able to point out any direction as the direction to the centre of the world, an intelligent observer on any of the other islands he could see would indeed be able to do so; and all such observers would point at his own island!

The situation in relativistic cosmology is precisely similar. We can construct all space-times which would give exactly isotropic observations about one particular galaxy; and they are either exactly spatially homogeneous and isotropic space-times, which are isotropic about every galaxy—in this case, all galaxies are equivalent—or they are centred on that one galaxy. This galaxy is then at the centre of the universe (5). The actual universe, which is not exactly isotropic about us, may then be expected to be very similar to one or other of these idealized possibilities.

In ages gone by, the assumption that the Earth was at the centre of the universe was taken for granted. As we know, the pendulum has now swung to the opposite extreme; this is a concept that is anathema

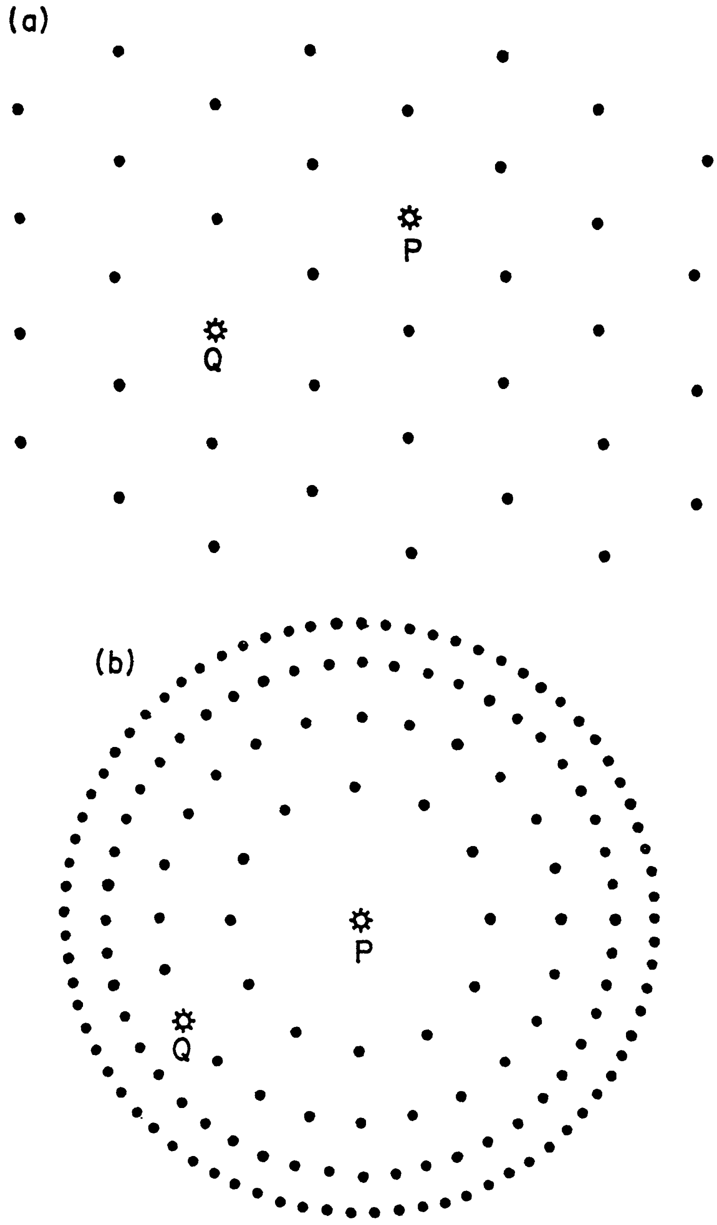

FIG. 1(a). An arrangement of particles which is statistically isotropic about every particle (P and Q are equivalent). (b) An arrangement of particles which is statistically isotropic about P, but not about any other particle (P and Q are not equivalent).

to almost all thinking men. This is partly because we now believe that our Galaxy is no different from millions of others; more fundamentally, it is due to the Copernican–Darwinian revolution in our understanding of the nature of man and his position in the universe. He has been dethroned from the exalted position he was once considered to hold.

It would certainly be consistent with the present observations that we were at the centre of the universe, and that, for example, radio sources were distributed spherically symmetrically about us in shells characterized by increasing source density and brightness as their distance from us increased (6). Although mathematical models for such Earth-centred cosmologies have occasionally been investigated, they have not been taken seriously; in fact, the most striking feature of the radio source counts is how this obvious possibility has been completely discounted. The assumption of spatial homogeneity has inevitably been made, and has led to the conclusion that the population of radio sources evolves extremely rapidly (7). What has therefore happened is that an unproven cosmological assumption has been completely accepted and used to obtain rather unexpected information about astrophysical processes.

It seems likely that reasonable theories will continue to be based on this assumption. One may adopt this view simply because our own Galaxy seems a rather undistinguished place to be the centre of the universe, or because of deeper philosophic reasons. In any case, we shall accept the implied attitude, and turn to consider the different ways it can be formalized. Important differences in our concept of the universe arise if we formalize it in different ways.

UNIFORMITY PRINCIPLES

The traditional way of codifying the view that we occupy an average, rather than a highly special, position in the universe is to adopt the *Cosmological Principle* (8): that is, the assumption that *the universe is spatially homogeneous*. This principle implies the existence of a cosmic time, and states that all measurable properties are the same at the same cosmic time. In particular, our observations of the isotropy of the universe would mean that all other observers viewing the universe at the same time would find their observations equally isotropic. Hence one obtains as idealized universe models the exactly spatially homogeneous and isotropic (or *Robertson–Walker*) space-times (9). These are supposed to represent the smoothed-out structure of the universe: a more realistic universe model can be obtained by superimposing small perturbations on this completely smooth substratum.

The Cosmological Principle is a positive statement with far-reaching consequences. An alternative way of proceeding is to make a negative statement. Thus we might make the assumption: *we are not at the centre of the universe*. (I shall refer to this as the *Copernican Principle*.) As has been indicated above, this principle together with the observed isotropy of the universe about us again leads us to perturbed Robertson–Walker space-times as models of the observed universe.

To illustrate the differences between these two approaches, consider once again our marooned natural philosopher. Having formulated for himself a 'Cosmological Principle'—that every part of the world is identical with every other part—he triumphantly announces his homogeneous and isotropic world model: the world is a completely smooth ball. Not only are all points equivalent to each other, but for every point, observations made in any direction are equivalent to observations made in any other direction. His lady-friend—who has been around all the time, but engaged on other enterprises—now correctly but somewhat unkindly points out that the world does not look very uniform to her. This necessitates him explaining that the world model was not meant to be an exact model of the world, but only an approximate one showing its basic, overall structure; a more adequate model would be obtained by thinking of a lot of islands scattered all over the idealized smooth ball. The homogeneity is meant to be understood in some unspecified statistical sense.

As his friend's reaction is not completely positive, he broods overnight and the next day formulates his 'Copernican Principle'— that their own island is not at the centre of the world. He then easily convinces her that this principle—not being stated as an exact requirement of uniformity—is readily amenable to a statistical discussion; and that (because of the isotropy of the world about their island) it leads to the conclusion that the world they see is approximately a smooth ball with islands scattered over it in a uniform way. He is delighted to find she accepts the principle as compelling, and the resulting world-model as an obvious consequence. The new formulation has the advantage that unlike the Cosmological Principle which only applies to highly idealized models of the world, the Copernican Principle can be applied to realistic world models; and so is a more satisfactory way of formalizing the assumption.

Nevertheless, in practice these principles may be interpreted so as to lead to the same ideas about the observed universe. The problem lies elsewhere, as our friend realizes with a sinking feeling when his companion asks him 'Gee, does that mean there are islands just like ours in all the parts of the world we can't see?'. This question puts him in a quandary. His Cosmological Principle made a definite prediction about all the unobservable areas over his horizon, namely that conditions there are the same as conditions near him. But he has no observational information whatever about these regions, nor will he ever obtain such information in the foreseeable future; so this conclusion is a direct result of his completely unverifiable assumption about the world. If he merely assumes the Copernican Principle, this orders his world the same way in the observable region, because he knows that in this region the world is nearly isotropic about him. But he does not have any such information about the unobservable regions, and accordingly the Copernican Principle (as formulated here) makes no particular prediction about these hidden regions. Indeed, according to the available evidence they could be totally different from the areas near him. Thus there could be many more islands, or many fewer, or no islands, or perhaps a continent in some part or other; or perhaps his whole concept of the world as a roughly uniform ball might be wrong, for while it might have that form near him, it could be, for example, that the region he saw was just the top of a mountain based on some landform of completely unknown shape.

The situation in cosmology is essentially the same. The Cosmological Principle determines a complete universe model; the Copernican Principle only a model of the observed part of the universe. The first model is satisfying because it is complete, but unsatisfying because it makes predictions about parts of the universe which are beyond observation; one has only one's faith in the integrity of this principle to validate these predictions. The second model is satisfying in that it only attempts to state conditions in the observable parts of the universe, but is therefore also unsatisfying, as there are further regions of the universe which it does not attempt to describe. Attempts to resolve this by setting up some intermediate principle seem unlikely to help. For example, we could postulate *a weak cosmological principle*: that *we are at a typical position in the universe*; but the effect is essentially the same as that of adopting the strict form of this principle (which is the form actually used in most writing on the subject). One could alternatively give different formulations of the Copernican Principle, such as the *Sample Principle* (our observed region of the universe, if sufficiently large, constitutes a fair sample of the universe); or the statement 'we do not occupy a privileged position in the universe'. We obtain essentially the same world-models as when using the original form of the Copernican Principle (a problem arising here is that it is not absolutely clear what 'privileged' should be understood to mean). These alternative formulations do not solve the essential dilemma.

UNIVERSE MODELS

In order to be more precise, I shall briefly sketch the universe models obtained when these two principles are used. In doing so, I shall take General Relativity with vanishing cosmological constant (Λ) as the theory correctly describing the effect of gravity and determining the structure of space-time; similar models would be obtained if $\Lambda \neq 0$, and from closely related theories, such as the Brans–Dicke theory. I shall also take the conventional interpretation (**10**) of the observations, rather than one of the more exotic alternatives (**11**) (which explain certain puzzling features at the expense of introducing various new problems of interpretation). To give an idea of these universe models, consider the picture of the (curved) space-time obtained when two space dimensions are suppressed; the resulting diagram is a space-time diagram with only one space dimension, representing the total history of the universe. The curvature of space-time results in the space-sections being represented by curved lines; at each point we may think of the local time direction as being orthogonal to the space-section at that point.

When the Cosmological Principle is applied, the exactly homogeneous and isotropic world models resulting can, in general, be represented as in Fig. 2 (there is an exceptional case which we shall return to later).

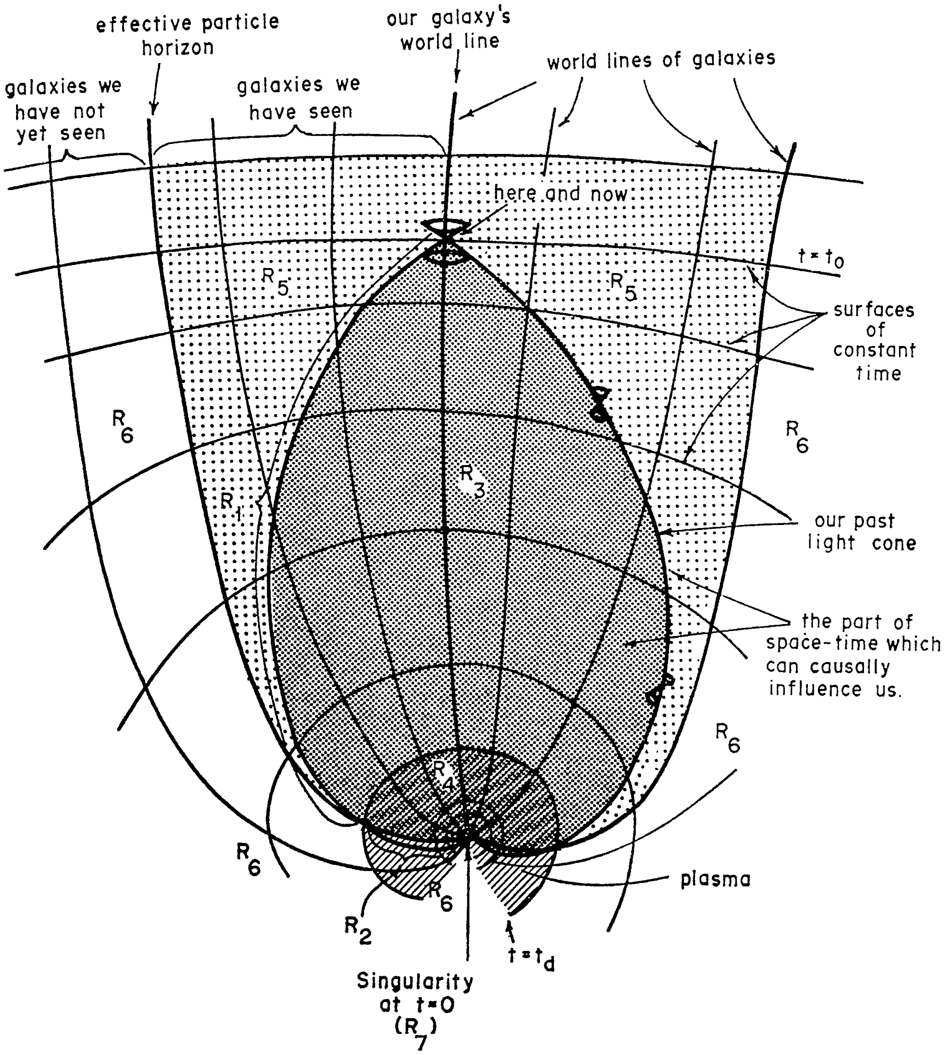

FIG. 2. Universe model based on Cosmological Principle.

The surfaces of constant time are surfaces of spatial homogeneity, so all physical quantities are identical at each of the events on any one of these surfaces. The complete histories of galaxies are represented by world lines, showing the spatial position of the galaxy at each time; the separation of these world lines at any one time represents the distance between the galaxies at that time. The present time is represented by the surface $t = t_0$. The present expansion of the universe is represented by the way the world lines intersect later surfaces of constant time at wider separations; conversely, as one considers earlier times, the galaxies are closer together. Continuing back in time, the matter particles are ever closer together, and consequently the density of the matter is ever higher; and within a finite time in the past, the density and temperature of the matter become infinite at the initial 'big bang'. It is convenient to choose the time parameter so this occurs when $t = 0$. Mathematically, a singularity in our world model occurs here; physically, known local theory breaks down here: we are unable to predict to earlier times. Thus our universe model is *finite in past time*: the matter and radiation, and space-time itself, do not exist at earlier times, so this represents the beginning, or origin, of the universe model.

The second important feature of these universes is that, like the matter, the radiation in the universe is compressed in the past, and becomes indefinitely hot at early times. This means there is a finite time, t_d, $(t_0 > t_d > 0)$ when the radiation is sufficiently intense to ionize the matter, and at earlier times (i.e. for $t_d > t > 0$) the universe is filled with a *plasma which is opaque to electromagnetic radiation*. The third important feature is our *past light cone*, i.e. the history in our past of light which we see at this instant (**12**). This represents the part of space-time from which it is now possible for us to receive directly signals in the form of electromagnetic, gravitational or any other type of radiation. It bounds the part of space-time from which we could have received any signal or other form of communication because particles and signals are unable to travel faster than light, and the light cone represents signals impinging on us at the speed of light. Thus we can only be influenced by events lying in, or on, this past light cone.

We can immediately identify seven regions in this idealized universe model, which have essentially different observational status. Region R_1 is the part of our past light cone since the recombination of the primeval plasma (i.e. for $t_0 > t > t_d$). This is the set of events from which we may receive information by means of electromagnetic waves, in particular by light or radio observations, and by any other form of radiation (such as gravitational waves). It is the maximal part of the universe we can actually hope to *see* (part of this light cone is blocked off from our view by intervening matter; we can only actually receive radiation from a particular event on our past light cone if nothing opaque intervenes). Region R_2 is the part of our past light cone prior to recombination (i.e. for $t_d > t > 0$). While we cannot receive information from these events by electromagnetic radiation, because the plasma absorbs or scatters photons passing through it, we can in principle 'see' these events by using very sensitive gravitational wave and neutrino telescopes. Thus we can, in principle, directly probe this region by observing radiation other than electromagnetic radiation.

Region R_3 (the interior of our past light cone since decoupling) is a part of space-time which we cannot observe by any form of radiation. However, we have sufficient information available (from our direct observations on our light cone, and from other kinds of information such as geological data) to be able to have a general idea of what conditions are like here. For example, we see the Andromeda galaxy at a certain time in its history; by determining its velocity, we can with reasonable certainty plot its previous motion, that is, determine its world line in R_3. In principle, the same applies to the region R_4, the interior of our past light cone prior to decoupling; but, in practice, we are unable to form a very precise picture of what is happening here because of the damping effect of the plasma: fairly arbitrary initial conditions lead to much the same final state, so, conversely, observation of the final state gives rather little information as to the initial conditions.

Regions R_5 and R_6 are parts of the universe with which we can have had no causal communication. The difference between these two regions is that the galaxies whose histories are represented by world lines in R_5 are galaxies we could possibly have observed by light or radio waves emitted at some stage in their history; whereas we could never have received such signals from the galaxies whose world lines lie in R_6. Thus while we can predict something about the matter in R_5 by extrapolating from our observations of this matter at earlier stages of its history, we have virtually no information about the matter in R_6 at any stage of its history, and so are quite simply unable to predict the state of this matter from any observational information available to us. Some of this matter could in principle have been observed by gravitational wave or neutrino telescopes; but even when such observations are feasible in practice, we will most probably obtain very little information about the distribution of the matter from these observations. The rest of the matter in this region could not have been observed by us by any form of radiation whatsoever; nevertheless, we could in principle detect that some of this matter exists because of the effects (such as that due to its gravitational Coulomb field) it has on our past light cone. However, no way is known to decode these effects to determine what distribution of matter is causing them; thus we cannot decide if a particular distortion is due to a large distant object or a smaller nearby object (this is a difficult problem involving the 'constraint' equations of General Relativity); so we cannot determine the detailed distribution of matter anywhere in R_6. Finally R_7 represents a different form of unpredictability; it denotes the singularity at the origin of this universe model, where the ability of known physical laws to predict breaks down. This breakdown arises not because of a lack of data, but because attempts to predict using the local predictability principle and presently established local physical laws lead to a contradiction. The universe model may therefore be thought of as beginning at this time; the picture we obtain throws no further light on this origin.

Provided we make one further assumption, the Copernican Principle leads to a rather similar universe model. The extra assumption we have to make to ensure our universe model is reasonable is the *Causality Assumption* (13): *it is not possible for an observer to encounter himself.* Obvious problems concerning the nature of free will arise in a space-time in which an observer's world line can twice approach the same space-time point, so that he (as a young man) meets himself there (as an old man). The assumption that this cannot happen has to be made explicity in this case (14) (it was automatically fulfilled in the exactly homogeneous and isotropic space-times). Having made this assumption, we obtain a universe model (see Fig. 3) divided into regions R_1 – R_7 with the same significance as in the previous model. The regions R_1 and R_3,

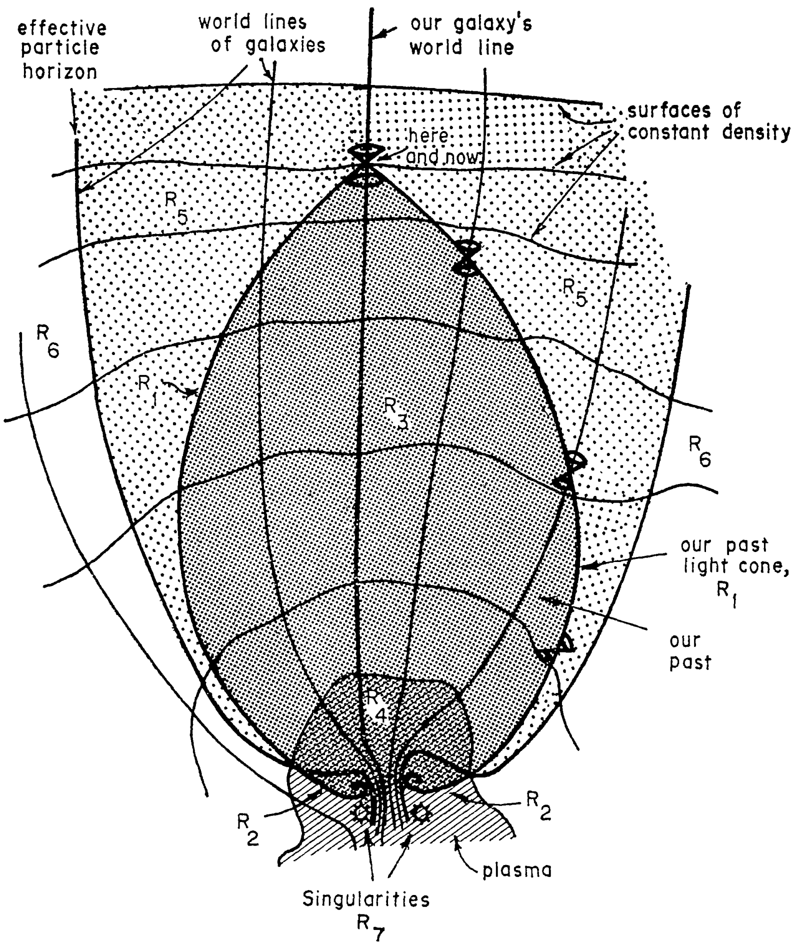

FIG. 3. Universe model based on Copernican Principle.

and the parts of R_5 near R_1, are very similar to the corresponding regions in the previous case; the present picture is just a more accurate (so 'bumpy') representation of these regions than the idealized (smooth) one. Thus the past light cone R_1 may develop caustics and self-intersections; after such caustics or self-intersections it no longer bounds R_3, but rather lies inside R_3(15). Our information is not good enough to let us have precise geometrical concepts of regions R_2 (the part of our past light cone prior to recombination) and R_4 (the part of our chronological past (16) prior to recombination) on the basis of the available data (again, part of R_2 may now lie inside R_4). However, we may be confident of certain features; particularly that there are very high density and temperature regions in R_4, and that there exists at least one singularity, R_7, here (17). The nature of this singularity is not known; in particular, it is not clear whether a large portion of the matter crossing our past light cone in the universe model should be regarded as originating at the singularity or not. There are some hints that these regions are rather *unlike* a Robertson–Walker universe model (18), so this picture of the universe will (when we have sufficient data available to make definite statements about this region) probably differ considerably, in regions R_2 and R_4, from the first picture. Finally, we simply do not have enough information available, on the basis of present or possible future observational data, to say anything much definite about the region R_6 and the parts of R_5 distant from our past light cone. Continuity suggests that conditions in the parts of space-time outside our light cone, but near it, will be similar to conditions inside it; but does not justify our holding the same expectation for more distant regions.

Comparing these two pictures, it is clear that, in the second, far less sweeping use is made of the unverifiable assumption (the Copernican Principle) than in the first (when the Cosmological Principle is the unverifiable assumption). We need to make *some* unverifiable assumption to order our ideas about regions R_1 and R_3, as discussed previously; but the Cosmological Principle has also ordered the overall structure (but not the details) not only in the infinitely distant parts of R_6, but also in parts of R_5 infinitely far away from us in time. These extravagant predictions about parts of space-time completely out of reach of any form of observation seem unreasonable; the impossibility of obtaining any relevant observational data makes adoption of this ordering principle seem suspect, and rather arbitrary (19). In any case, as mentioned above, there are some indications that the uniform models might be wrong in regions R_2 and R_4; it may be that we will be able eventually to prove that the Cosmological Principle is misleading if applied to these regions. It therefore seems that adoption of the Copernican Principle is the better procedure.

PREDICTABILITY

Suppose then that we adopt the Copernican Principle, and so obtain a universe model whose principle features are as sketched in Fig. 3. We know of the *existence* of the regions R_5 and R_6 because of the local predictability assumption; but considerable uncertainty enters into what we will *ever* be able to say, with reasonable confidence, about these regions, and about the early parts of R_4. The farther away in space or time an event is, the less we can reasonably hope to predict about conditions there (**20**). If some regions of the universe model are effectively beyond observational and experimental reach, what scientific status should we assign to these regions? How much significance can we assign to these regions in our universe model, in this situation?

When it was realized that knowledge of the microscopic domain was limited by a fundamental principle of impotence (the 'Uncertainty Principle'), physicists took this principle seriously and made it the basic feature of quantum theory. Should cosmology perhaps, as suggested by W.H.McCrea (**21**), similarly take seriously the fundamental limitations on what we can say about the universe, and turn them into the basic feature of our cosmological theory? It seems likely that this is what we shall, in the end, have to do: to acknowledge our inability ever to determine many features of the universe, and to incorporate this indeterminacy as a basic feature in our universe models.

At present we have no detailed proposal to hand for the implementation of this idea. However, what we can do is to go back to our classical picture of the universe, and examine in more detail the sorts of uncertainty that arise. This will then provide the starting point from which further developments can proceed.

Let us refer back, then, to Fig. 3. The initial uncertainty sets in on our light cone: the farther down our light cone we observe, the less we can say about the objects we are observing. This is partly because of interference by intervening matter, but primarily because of the distance involved: the object has a smaller, fainter, more red-shifted image if it is farther away. The amount of information we can obtain from observations of any particular object in a given time with a particular telescope is limited by optical and quantum considerations; and the farther down our light cone the observed object is, the less we can find out about it in a given time (**22**). Despite the general information we may eventually get from gravitational wave or neutrino observations, it seems that we will never get a detailed picture of R_2.

Our knowledge of local physics enables us to extrapolate from these observations into R_3; if we consider the regions nearer to us in space and time, the data on R_1 are better, and we can extrapolate back with more certainty to determine the previous conditions which have led to what we observe. Some uncertainty arises because of uncertainties in the initial data, and some because of the statistical nature of prediction, both in quantum mechanics and in statistical pyhsics. We are also able to extrapolate back in R_3 to regions near our Galaxy's world line in the very distant past, by use of geological and geophysical data which tell us about the very early history of our Galaxy. This kind of information will probably give us our best indications as to conditions in R_4, where physical conditions can be very extreme and difficult to understand. In particular, reasoning based on the 'cognizability principle' developed especially by Dicke and Carter (23), (24) (we observe the universe; so conditions in our past must have been of such a nature as to have allowed the development of intelligent life) provides an interesting way of deducing limits on conditions in the past.

Prediction in R_5 is more of a problem. In principle, we should be unable to predict anything here; for this is our observational future. It has not yet been accessible to observation, and we do not have complete data available to predict what will happen here; indeed, some of this region lies in our own chronological future. General Relativity allows the possibility that arbitrary electromagnetic or gravitational shock waves can impinge on us without any warning being given by data on R_1; completely nullifying any prediction we might make. In practice, this has not yet happened; and we may regard it as unlikely that it will happen, primarily because of the plasma in R_6 which shields us from any primeval disturbances. If some laser-type wave were emitted towards us from the initial singularity, the plasma would attenuate it and protect us from it; at the very least, the diffusive effect of the plasma would give us some warning of the approaching threat, in the form of a highly increased blackbody radiation temperature or distortions from a blackbody spectrum. In any case, the large redshifts involved decrease intensity by an enormous factor. We have seen all the matter in the region R_5, and can therefore estimate what its future behaviour is likely to be, and could hope to deduce if it were likely to send high-intensity signals towards us. Thus the overall structure of the cosmological model is such that local prediction into the future *is* possible: the data we have (on R_1) are in practice sufficient to predict into parts of R_5 near R_1, on an astronomical scale, because few unexpected data arrive here from elsewhere in the universe (25). Thus we have no more difficulty in predicting where the Moon will be in 5000 years time, than we have in determining where it was 5000 years ago. The main effect of extrapolating into the future rather than the past is that our uncertainties are somewhat greater; for example, men might have destroyed the Moon in 5000 years time, and a complete prediction has to include estimates of the probability of this sort of eventuality (the larger the scale of object considered, the less the problem of such possible interference is likely to be). Having accepted this somewhat greater uncertainty, no major difference arises in predicting on a cosmological scale into the part of R_5 near R_1 rather than the part of R_3 near R_1, from data on R_1. As we have only the idea of continuity to help in predicting into R_6 and the distant regions of R_5, we know very little about these regions.

Putting this together, we arrive at an idea of how certain our knowledge of various regions can be eventually (we are more certain if we have more information available, or if the reliability of our information is improved). The sort of picture we obtain is shown schematically in Fig. 4; this indicates the kind of accuracy with which we can hope to

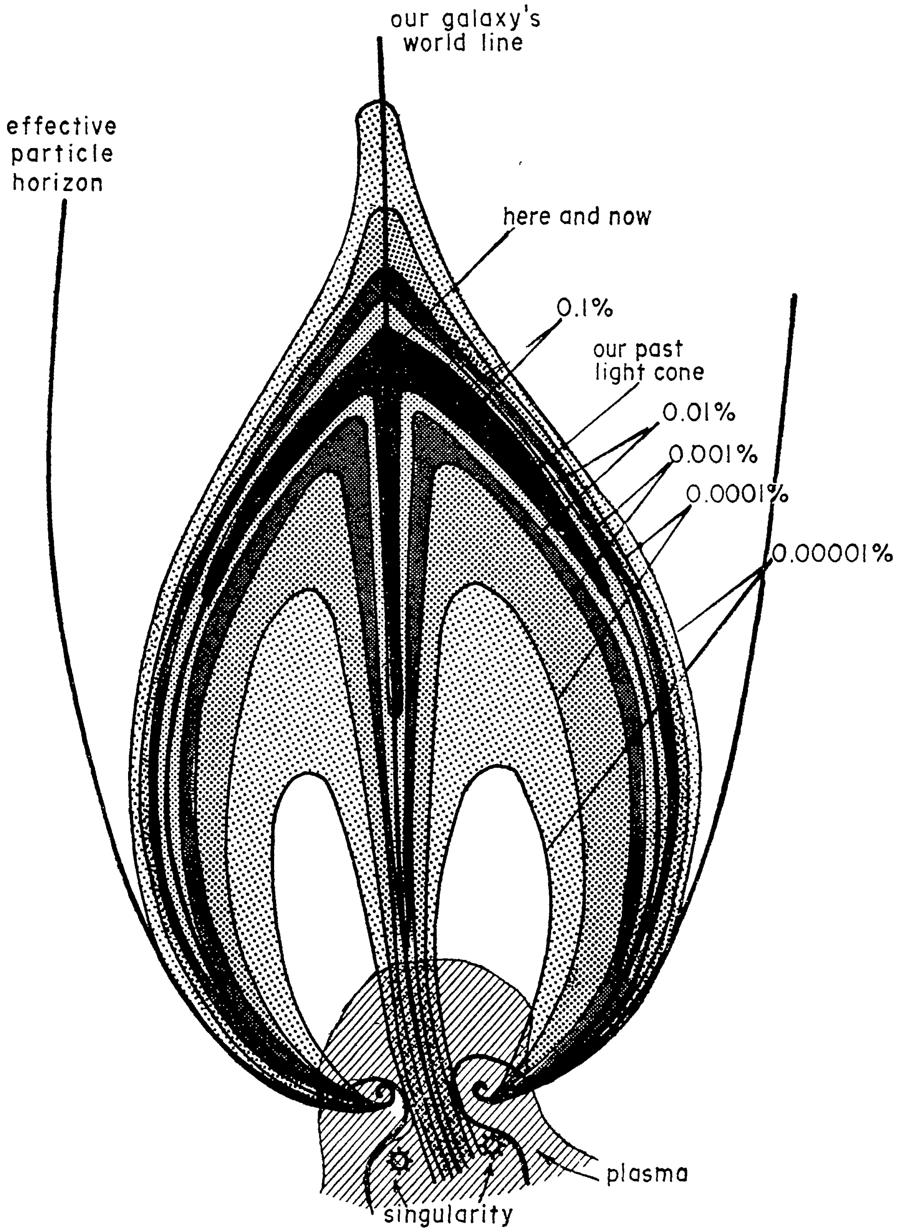

FIG. 4. Prediction certainty in the universe (schematic).

determine the structure of the universe from observations, without introducing further unverifiable assumptions. The mere acquisition of more detailed data will not change this overall picture, but rather will fill in some of the details; it will only be upset by some completely unexpected observational result changing our overall view either of the concept of space-time inherent in General Relativity, or of the cosmological data on which our universe picture is based (**26**); or by the satisfactory inclusion in our theory of a new and compelling principle (such as Harrison's 'bootstrap' principle (**27**)) which necessarily orders those parts of the universe beyond our powers of observation. Without some such complete overthrow of our present conceptual scheme, the diagram we have obtained depicts the main features of the observable universe and the accuracy with which we can hope to determine its details; the aim of observational cosmology, in this context, is to obtain the information necessary to fill in the details of this picture up to the maximum attainable accuracy. The diagram therefore represents the goal of observational cosmology, rather than its present status (**28**). It indicates the extent of knowledge of the universe one can aim to have supported by observational evidence; and is therefore the proper setting in which to consider how empiricist a view one should take in constructing a cosmological theory. The somewhat mundane astrophysical considerations we have taken into account are, in my view, essential to a proper discussion of the epistemology and ontology of cosmology; I do not claim to give such a discussion here, but rather to set the stage for it.

AN EXCEPTIONAL CASE

So far we have considered the general situation. Finally we should consider the exceptional situation which arises if the universe has finite spatial sections (this necessarily occurs in the Robertson–Walker universe with $k = +1$, but can also occur (**29**) if $k = 0$ or -1). The essential difference is that then there are only a *finite number of galaxies* in the universe, and the universe has at any one time only a finite volume; and in many cases the universe has only a finite future ahead of it. This alleviates the problem of discussing regions R_6 and R_5; for there may not then be parts of the universe infinitely far to the future of us, and there are no events indefinitely far away from us spatially. In a restricted subclass of these idealized universes, (in particular if $k = -1$ and Λ is very large) there may be no region R_6 at all: our past light cone may 'close up' on itself so that we could, in principle, obtain sufficient data by optical and radio observations to determine completely the future and past of the universe. In this case, no expected signals from unseen sources would surprise us, as we would have sampled the history of all the matter in the universe and no unexpected information can come from previously hidden parts of the singularity R_7. There would, in practice, still be considerable uncertainties as to what we could predict, but we would be able to relate a major part of the universe to our observations in these cases; the observable universe would be almost the whole universe.

How could we tell if this was the situation in the actual universe? The crucial effect if the light cones closed up would be that *we would be able to see each galaxy in at least two different directions in the sky* (unless some opaque matter intervened). Thus, in principle, we have a simple way of seeing if this is the case or not: we simply compare the images of galaxies in different directions, and see if they might represent the same object or not. Unfortunately, this would be very difficult to establish in practice: not merely because of the extreme difficulty of making the requisite observations (we would expect this effect to occur at the limit of possible observations), but also because the different images we would obtain of one galaxy would be images resulting from light leaving the galaxy in different directions and at different times in its history. Thus we would effectively be looking at the galaxy from different directions and at different stages of its evolution. This would make it virtually impossible to tell if one was in fact seeing the same object in different directions in the sky, or not. Nevertheless, it is certainly something one could attempt to do, even if there is little hope that one would obtain either a definite confirmations or a definite denial.

It is, in fact, more likely that the question of whether these exceptional situations arise in realistic (**30**) universe models, or not, will eventually be decided indirectly from evidence concerning the amount of matter in the universe (**31**) and the value of the cosmological constant: certain combinations of these quantities make the exceptional situations inevitable, and others make them implausible (but not impossible). This is an important question because of the major difference it makes to the verifiability status of the universe. Various authors (including Einstein and Wheeler) have argued on ideological grounds that there must be finite space sections; and for a time it was strongly argued, particularly by Sandage, that observational evidence supported this view. The evidence from observations is not now widely regarded as being conclusive either way, and the question remains open; my own somewhat biased view is that the present evidence makes it rather unlikely that there are compact space-sections, and very unlikely that the past light cone closes up on itself. If this is correct, then the observable part of the universe is a rather small part of the whole universe.

THE BEGINNING

Perhaps the most intriguing question of all is the relation of cosmogenesis—the nature of the singularity where at least part of the matter in the universe is, in some sense, created (**32**)—to observational tests. I suspect that definite views on this will have to wait for a far more advanced theory combining a quantum description of matter with gravitation than any we have at present. If some limits could be placed on the possible nature of the initial singularity by some such theory, this might provide a further way of examining the nature of those distant parts of the universe we have been considering. For, in this case, we would be beginning to understand the creation process itself, and that ought to give us some ideas as to the limits of what might be created. This, or a 'bootstrap' argument (33), would enable us to progress from merely *observing* the universe, to *explaining* it in some sense. At present, this is just a faint and distant hope—a gleam in the eye which may some day become a reality. Such a change in the technological or conceptual apparatus available for examination of the problem couldchan ge the whole situation, and our whole certainty as to the nature of the universe. In fact, our friend on the island was last heard muttering profanely to himself as he chopped down a tree and proceeded to fashion its trunk into a rudimentary but serviceable boat.

REFERENCES

(1) An informative discussion of the theoretical basis of cosmology is given in H.Bondi, *Cosmology* (Cambridge University Press, 1960). A useful discussion of the history of the subject is in J.D.North, *The Measure of the Universe* (Oxford University Press, 1965).

(2) I shall, as discussed later, assume that General Relativity is the correct theory describing space-time and gravitation.

(3) The way this can be done locally has been carefully described by Kristian and Sachs, *Astrophys. J.*, **143**, 379 (1966).

(4) A very readable discussion of recent observations and their interpretation is in D.W.Sciama, *Modern Cosmology* (Cambridge University Press, 1971). A more detailed discussion of the physics involved is in S.Weinberg, *Gravitation and Cosmology* (Wiley and Sons, 1972); a detailed review of recent observational data has been given by Longair, *Rept Progr. Phys.*, **34**, 1125 (1971).

(5) In certain such universes, there could be *two* centres. The argument then proceeds unchanged; there are two galaxies whose situation is completely different from that of all other galaxies in the universe. In an expanding universe, isotropy of the *world picture* (the universe as *seen* by the observer) implies either anthropocentricity, or homogeneity and isotropy of the *world map* (the instantaneous map of the universe he constructs, utilizing his knowledge that light travels at a finite speed).

(6) Systematic redshifts could be observed in such an Earth-centred universe even if it were static.

(7) See Reference (4) above.

(8) The various principles are discussed in detail by Bondi and North (1). Many other reviews of cosmology discuss them, but in less detail, see e.g. the reviews in (4); for a recent reappraisal, see Harrison's articles in *Comm. Astrophys. Space Phys.*, **6**, 23; 29 (1974). We have not here elaborated on the *Perfect Cosmological Principle* (the assumption that the universe is homogeneous in space *and* time) as experimental evidence seems to be against it.

(9) See the references in (1) (4), or the review articles in *General Relativity and Gravitation*, Ed. R.K.Sachs (Academic Press, 1971) (Proceedings of the 47th International School of Physics "Enrico Fermi"); or in *Cargese Lectures in Physics*, *Volume 6*, Ed. E.Schatzmann (Gordon and Breach, 1973).

(10) See Reference (9) above.
(11) See, for example, the discussions in *Nature*, **241**, 109, 338–340 (1973); **242**, 108 (1973).
(12) On a cosmological scale, 1000 years is effectively an 'instant'.
(13) More precisely, the 'strong causality assumption'. See S.W.Hawking and G.F.R.Ellis, *The Large Scale Structure of Space-Time* (Cambridge University Press, 1973), for a discussion of these principles.
(14) See K.Gödel's article in *Einstein: Philosopher–Scientist*, Ed. P.A.Schilpp (Harper Torchbooks, 1959), for some thoughts on causality violation and its consequences.
(15) There may also be parts of the boundary of R_3 and R_4 disjoint from the past light cone R_1 and R_2.
(16) See *The Large Scale Structure of Space-Time* (13) for a detailed discussion of causal concepts and the singularity theorems of Penrose and Hawking.
(17) See Reference (16) above.
(18) See Reference (9) above.
(19) There is a fairly widespread tendency to adopt the view that assuming there is no change in conditions (as in the Cosmological Principle) is effectively making no assumption at all. This is clearly untrue.
(20) See O.Heckmann and E.Schücking in *Onzième Conseil de Physique Solvay: La Structure et l'evolution de l'univers*, Editions Stoops, Brussels (1959), and F. Hoyle in *Rendiconti Scuola Enrico Fermi, XX corso* (Academic Press, 1960) for discussions of this feature.
(21) W.H.McCrea, *Nature*, **186**, 1035 (1960); **187**, 583 (1960); La nuova critica, Cosmologia, 11e serie (1960–1961).
(22) See, for example, A.W.K.Metzner and P.Morrison, *Mon. Not. R. astr. Soc.*, **119**, 657 (1959); G.J.Whitrow and B.D.Yallop, *Mon. Not. R. astr. Soc.*, **127**, 315 (1964); D.H.Gudehus, *Publ. astr. Soc. Pacific*, **84**, 818 (1972).
(23) M.J.Rees, *Comm. Astrophys. Space Phys.*, **4**, 179 (1972). B.Carter, unpublished; R.H.Dicke, *Nature*, **192**, 440 (1960).
(24) E.R.Harrison, see (8).
(25) Local physics is therefore a meaningful enterprise. (Local physics relies completely on the concept of an 'isolated system'; but we cannot isolate any system from gravitational radiation.)
(26) The picture would have to be seriously revised if, for example, observations eventually prove the universe has a hierarchical structure, as has been suggested by de Vaucouleurs (*Science*, **167**, 1203 (1970)) and others; see, for example, Gold in *Nature*, **242**, 24 (1973).
(27) See Reference (24) above.
(28) It is salutary to realize that different observers' estimates of the value of the Hubble constant still disagree to a marked extent about the probable range of values for this number.
(29) O.Heckmann and E.Schücking in *Gravitation*, ed. L.Witten (Wiley, 1962); G.F.R.Ellis, *J. Gen. Rel. Grav.*, **2**, 7 (1971).
(30) But idealized in being smoothed out: a really detailed universe model, describing details of possible black holes, wormholes and so on, may be expected to have a non-trivial region R_6.
(31) See the discussions in the references in (4).
(32) An interesting discussion has been given by N.R.Hanson, *Philosophy of Science*, The Delaware Seminar, **2**, 465 (1962–1963); ed. B.Baumrin (Interscience, 1963).
(33) See Reference (24) above. The basic idea is to make predictions by consistent application of the concept 'Nature is as it is because it is the only possible Nature consistent with itself'.

Q. Jl R. astr. Soc. (1975) **16**, 265–281

Does Astronomy Need 'New Physics'?

V.L.Ginzberg

(George Darwin Lecture delivered on 1975 April 11)

(P.N.Lebedev Physical Institute, Academy of Sciences, USSR, Moscow)

INTRODUCTION

Astronomy and physics have always been closely connected but in some periods this connection has been particularly intimate and, one could almost say, personal. One such period began approximately in 1945 and is still going on—it is caused by the process of transformation of astronomy from a subject based on the optical wave-band to one in which all wave-bands are important. The development of radio astronomy and cosmic ray astrophysics and in the last few years X-ray and gamma-ray astronomy (not to mention neutrino and gravitational wave astronomy which are just beginning) is naturally connected with an intense influx of new people into astronomy, for the most part physicists. In some respects these newcomers differ from astronomers by education. Differences in terminology are striking, as well as the fact that many physicists are ignorant of the elementary facts from classical astronomy. When these neophytes are reminded that in astronomy they observe but do not carry out experiments, that is, of course, of no importance. But it is different if misunderstandings arise concerning questions of principle. Such misunderstandings between physicists and astronomers and as a consequence between astromers themselves do arise and most important of all they concern the answer to the question figuring in the title: 'Does astronomy need "new physics"?'.

A discussion of such general problems cannot influence in any important way the development of astronomy which proceeds mainly as a result of new observations and theoretical investigations and not by the declaration of principles. However, in small doses, a general discussion concerning the connection between physics and astronomy is of some interest and may prove useful. This question is one which is particularly close to my heart since I have been a physicist by education and professional experience since 1938, and then since 1945 I have also been engaged in astrophysics. By the way, my initiation happened merely by