INSTITUT INTERNATIONAL DE PHYSIQUE SOLVAY

ONZIÈME CONSEIL DE PHYSIQUE
tenu à l'Université de Bruxelles
du 9 au 13 juin 1958

LA STRUCTURE
ET
L'ÉVOLUTION DE L'UNIVERS

RAPPORTS ET DISCUSSIONS

publiés sous les auspices
du Comité Scientifique de l'Institut

R. STOOPS
Editeur
76-78, COUDENBERG, BRUXELLES, BELGIQUE
—
1958

LA STRUCTURE
ET
L'ÉVOLUTION DE L'UNIVERS

INSTITUT INTERNATIONAL DE PHYSIQUE SOLVAY

ONZIÈME CONSEIL DE PHYSIQUE
tenu à l'Université de Bruxelles
du 9 au 13 juin 1958

LA STRUCTURE
ET
L'ÉVOLUTION DE L'UNIVERS

RAPPORTS ET DISCUSSIONS

publiés sous les auspices
du Comité Scientifique de l'Institut

R. STOOPS
Editeur
76-78, COUDENBERG, BRUXELLES, BELGIQUE
—
1958

INTRODUCTION

LE ONZIÈME CONSEIL DE PHYSIQUE SOLVAY

Le XI[e] Conseil de Physique Solvay s'est tenu à Bruxelles, dans les locaux de l'Université, du lundi 9 au vendredi 13 juin 1958, suivant les dispositions arrêtées par la Commission Administrative, composée comme suit :

Président :

M. Jules BORDET, prix Nobel.

Membres :

MM. Ernest-J. SOLVAY, Paul DE GROOTE, Ilya PRIGOGINE.

Secrétaire :

M. Frans-H. van den DUNGEN.

Le Comité Scientifique qui avait arrêté le programme des rapports était formé de :

Président :

Sir W. Lawrence BRAGG, prix Nobel.

Membres :

Prof. C.J. GORTER, Prof. C. MØLLER, Prof. J.R. OPPENHEIMER, Prof. W. PAULI, prix Nobel, Prof. Fr. PERRIN.

Le septième membre, M. MOTT, avait fait part de son regret de ne pouvoir être présent.

Les membres rapporteurs du Conseil étaient :

MM. V. A. AMBARTSUMIAN, Academy of Sciences of Armenia.
W. BAADE, Mt Wilson and Palomar Observatories.
F. HOYLE, St. John's College.
O.B. KLEIN, Stockholm University.
G. LEMAITRE, Université de Louvain.
A.C.B. LOVELL, Jodrell Bank Experimental Station.
J.H. OORT, Sterrewacht te Leiden.
A.T. SANDAGE, Mt Wilson and Palomar Observatories.
H.C. VAN DE HULST, Sterrewacht te Leiden.

Les communications de M. A.T. SANDAGE et M. W. BAADE étaient orales.

Les membres invités étaient :

MM. M. FIERZ, University of Basle.
T. GOLD, Harvard University.
O. HECKMANN, Observatory of Hamburg.
B.V. KUKARKIN, Sternberg Institute of Moscow.
P. LEDOUX, Université de Liège.
W.H. Mac CREA, Royal Holloway College.
W.W. MORGAN, Yerkes Observatory.
L. ROSENFELD, Manchester University.
E. SCHATZMAN, Faculté des Sciences de Paris.
H. SHAPLEY, Harvard University.
P. SWINGS, Université de Liège.
J.A. WHEELER, University of Princeton.
H. ZANSTRA, University of Amsterdam.
H. BONDI, King's College, London (Secretary)

Les membres auditeurs, membres du corps professoral de l'Université de Bruxelles étaient :

MM. J. COX, R. DEBEVER, M. DEMEUR, J. GEHENIAU.

MM. GOLD, HECKMANN, MORGAN et WHEELER ont présenté des notes écrites. Le Prof. ROSSLAND, de l'Université d'Oslo, n'a pu se joindre aux invités.

Le secrétaire a été aidé dans sa tâche par les secrétaires adjoints: M^{mes} R. PANKOWSKI-FERN, A. PEETERS-SPITAELS, M^{lle} A. HULEUX, MM. Ch. LAFLEUR, J. HOUGARDY, R. VAN GEEN, membres du personnel scientifique de l'Université.

Le Conseil Scientifique a chargé le Prof. J. GEHENIAU de diriger l'édition du volume contenant les rapports et discussions; il a été aidé dans cette tâche par M. R. VAN GEEN, assistant à l'Université de Bruxelles.

Les Autorités de l'Université ont reçu les membres du Conseil le lundi 9 à 17 heures dans la salle du Conseil. Le jeudi 12, à 21 heures, un dîner a réuni les participants au Resturant de l'Atomium dans le cadre de l'Exposition Universelle de Bruxelles 1958.

ACKNOWLEDGEMENT.

At the end of the general discussion of Friday, June 13, Dr. Harlow SHAPLEY made this speech :

Mr. Chairman, may I have a moment for a benediction? I desire to assure you and your colleagues on the Solvay Committee that we, your guests, are indeed grateful for the opportunity of convening and conferring. We have, of course nothing but enthusiasm for the magical chemistry of Solvay cuisine — enthusiasm for the processing of potables and comestibles. We enjoyed the high living — especially last nights' high feeding in the uppermost ball of the Atomium.

Our thanks are also due to Professor F.H. van den Dungen and his assistants for the organisation of the Congress.

In particular, our thanks go to you, Sir Lawrence for your sympathetic management, and for your skill in genially presiding over these sessions. You have maintained a neutral — I might say a neutron — pose during the turbulence, during the negative and positive charges and counter-charges, the explosions and implosions of gas and argument. We wish you, Sir Lawrence, you and your colleagues, many more pleasant and useful Solvay Conferences.

RAPPORTS
ET
DISCUSSIONS

I. GENERAL STATEMENT OF COSMOLOGICAL THEORY

The Primaeval Atom Hypothesis and the problem of the Clusters of Galaxies

by **G. LEMAITRE**

This report is divided in three parts.

In the first one, we discuss the general aspects of cosmology and point out their relations with Geometry, quantum theory and even with their inevitable philosophical background.

This is done without any intention to polemize against other points of view than the one which we have adopted, but, simply to make clear what are our own assumptions.

The second part is an exposition of the successive processes which arise in our cosmology and lead, in the order stated, from cosmic rays to gaseous clouds and then to proto-galaxies, with stars formation within them, and, finally, to the arrangement of the galaxies in clusters, with large peculiar velocities of their individual galaxies.

This makes possible to define what we call the problem of the clusters of galaxies and the third part reports on some results which point towards its solution.

1. GENERAL ASPECTS

Space.

Cosmology is essentially related to Geometry, and, in fact geometry in its deeper foundations : topology.

We may therefore begin with its topogical aspects, both in space and in space-time.

The question is : Is physical space a « compact » ? Can it be covered by a finite number of «neighbourhoods » ? One of these neighbourhoods would be the 10^9 light year accessible to astronomical observation.

How can we infer that, beyond this observed neighbourhood there are other ones with similar properties and how can we know if these neighbourhoods are in finite number and so form a finite compact or closed space?

One possible attitude of mind, in regard to this problem, is to rely on some « Cosmological Principle » which is dogmatically asserted and adhered to. Such principle would declare that any observer, at any place in the universe, would have essentially the same image of the universe around him.

An opposite way of approaching the problem would be to infer the content of the invisible neighbourhoods, which presumably surround our own one, from indirect consequences of observed facts. Matter is connected with curvature of space-time. Therefore any evidence that matter exists farther than what we directly know, results in producing the geometrical properties of our neighbourhood to larger regions and possibly far enough to close the whole space of finite and not too large radius inferred from the density in our neighbourhood.

Some indirect observation of this kind may come from the cosmic rays or from radio-waves coming from regions inaccessible to our telescopes.

As far as I see, the inclination to rely on an a priori principle is related to Leibnitz philosophical attitude which made him to believe that there is some esthetical design in the Universe or even that the Universe is determined as being the best possible one.

Of couse, this attitude would not necessary degenerate into mere idealism, with no contact with observation. Consequences of the principle should be deduced with the hope that they may come in agreement with observed facts, or, at least, not to be in contradiction with them. But, nevertheless, as a consequence of this philosophical attitude of mind, the a priori probability would be maintained for the principle.

It might be thought that the difference between these two points of view is not, in fact, very great. The one who relies on a cosmological principle would surely allow for local lack of homogenity which will relax the too strong rigidity of the principle. This would refer only to some kind of a « substratum ». Finally his position would not be essentially different from the recognition, by empirical deduction, of a large scale uniformity in the distribution of matter and therefore in the related geometry.

In fact, the difference is very great.

It can be seen in the attitude of some cosmologists who affect to consider on the same footing models with hyperbolic geometry as those with elliptical geometry. It is just a matter of a change of sign in a formula.

Now, if we rely on a cosmological principle there are in fact no essential differences in adopting one sign or the other for the parameter which determines hyperbolic or elliptical space. On the contrary, if we rely on inference from observation the determinant element of judgment is the relative size of the observed sample and that of the whole space which is inferred from it.

When space is open, with an infinite volume, the poor 10^9 light-years observed are a negligeable trifle and nothing can be inferred from that. On the contrary, for a compact finite elliptic space of a radius somewhat ten times the dimension of the region surveyed, the extrapolation may be found not to be too risky.

It is true that locally hyperbolic space is not necessarily open. It is possible to construct with such space polytrops, i.e. Klein's forms of finite volume. This is true even for euclidean space.

In fact, gravitational equations relate material properties to the local properties of space. Even for compact homogeneous space, it is well known that it is impossible to decide, from relativistic considerations, if space is spherical or simply elliptical. A similar situation, more artificial in some aspects, does exists for hyperbolic geometry.

It remains that, if space is found to be locally elliptical, and if it can be inferred that material distribution is not too different farther on, it must be inferred that space is a compact, although we may have some hesitation between its spherical or elliptical form.

Space-time.

A similar situation arises when we consider the topology of space-time.

We may remember that, before the discovery and the interpretation of the red-shift, the only solutions which were looked for, were the statical ones. Even de Sitter's space-time was originally disguised as a static universe.

There was, obviously, some philosophical unconscious attitude of mind in this preference for static solutions.

Later on, there has been an attraction for homothetic solutions for which there are really no change as in statical solution, except for a mere change of scale. This was a concession to the observed red-shift. A change was accepted but it was minimised as far as possible.

In this model, there is a beginning, but the model is always relatively at the same distance from it.

There is obviously some paradoxical flavour in this argument which reminds of Zenon's paradox.

In any way, strict adhesion to the homothetic solution proceeds from the same philosophical attitude than the one we have described above. It is expressed in what has been called the « perfect cosmological principle » and it has lead to the assumption of a continuous formation from nothing, or creation, of matter, in fact of hydrogen.

We shall explain the opposite attitude of mind as regards to the beginning of space-time in the next section together with its connection with quantum-theory.

Quantum theory.

It might be thought that quantum theory has nothing to do with cosmology. It is commonly believed that quantum theory refers only to microscopic phenomena and that macroscopic events, and, of course, the universe is one of them, seems to be outside its realm.

But, this point of view has been strongly opposed by Schrödinger and later on by Einstein. The consequences of quantum theory are not generally conspicuous in macroscopic phenomena, but, in principle, there is a real continuity between atomic phenomena and macroscopic phenomena as is seen, for example, in the Brownian motion. The same occurs for other quantum effects.

Therefore Cosmology must take into account the change of outlook brought about by quantum theory.

The classical outlook was that some « beginning » may be described by « initial conditions » which have just the same degree of freedom as those that will evolve from them. Future results could be inferred from these initial conditions.

With this outlook, it was difficult to understand how the initial conditions could be really a beginning. The same determinist laws which could predict the future, could as well be used to compute some more remote conditions from which the adopted « initial conditions » might have evolved.

This is one of Kant's antinomies.

Any contemplated beginning could be considered itself as a product of evolution. A real beginning could be introduced only by a kind of supernatural agency which would cut down arbitrarily the natural course of events.

Such a cut-down was described by Laplace « chiquenaude initiale » or by Jeans' «finger of God agitating the ether ».

Now, the advent of quantum physics has essentially modified the outlook. The indeterminacy principle opens essentially new possibilities for Cosmology.

Any physical system, and therefore the universe, is described as an assembly of potential « states » which can be or, not to be, occupied.

The most probable distribution, and therefore the final one is equal occupation of all possible states. Then, energy is distributed in as many distinct packets as it is conceivable.

In fact « entropy » may be considered as a measure of the total number of individual packets. For instance, in a radiation in equilibrium, entropy is the number of existing protons.

Degradation of energy, or the approach to maximum entropy, describes the natural tendency of energy to split into numerous bits.

A state of minimum entropy would be a state in which energy is condensed into as few packets as it is conceivable.

Such a state would be a beginning of multiplicity from which evolution would proceed by splitting. It would be a « natural beginning » in the sense that it could not have evolved from some simpler former beginning.

Of course such a state is too simple to afford enough marks from which the subsequent evolution could be inferred. But this is no more necessary, as classical determinism does not stand any more. Initial conditions need not to have the degree of freedom of the universe which will evolve from them.

From the same beginning, widely different universes could have evolved. The actual course of events has been progressively precised, while matter has split into more numerous distinct packets, in some unpredictable way.

Of course, when the number of individual packets became very large, the essential undeterminacy became ineffective and was replaced by the practical determinism characteristic of macroscopic phenomena, which simply arises from the law of large numbers of unpredictable phenomena.

These considerations, besides providing a natural beginning, supply what can be called an inaccessible beginning.

I mean a beginning which cannot be reached, even by thought, but which can only be approached in some assymptotic manner.

In absolute simplicity, no physical questions can be raised.

Beginning of multiplicity really means beginning of the very meaning of any notion which involves a great number of individuals. Space and Time are such notions.

It stands just before the beginning of space and time which acquires progressively a meaning while multiplicity increased enough. As space and time are the indispensible tools of any physical notion, it stands just before Physics.

It is an inaccessible ground of space-time.

Such a picture finds a natural geometrical support in the point-singularity which arises in Friedmann's theory. The radius of space can start from zero. Such singular event which arises when space has a zero-volume is a bottom of space-time which terminates every line of space-time.

I do not pretend that such a singularity is inescapable in Friedmann's theory, but I simply point out how it fits with the quantum outlook as a natural beginning of multiplicity and of space-time.

This is the philosophical background of the Primaeval Atom hypothesis.

As far as I can see, such a theory remains entirely outside any metaphysical or religious question.

It leaves the materialist free to deny any transcendental Being. He may keep, for the bottom of space-time, the same attitude of mind he has been able to adopt for events occurring in non-singular places in space-time.

For the believer, it removes any attempt to familiarity with God, as were Laplace's chiquenaude or Jeans' finger. It is consonant with the wording of Isaias speaking of the « Hidden God » hidden even in the beginning of creature.

It does not mean that cosmology has no meaning for philosophy.

The view we have proposed may be contrasted with that of Pascal in his *Pensées*.

We may reverse Pascal's wording and say that the Universe not being infinite neither in size nor in duration, has some proportion to mankind.

Science has not to surrender in face of the Universe and when Pascal tries to infer the existence of God from the supposed infinitude of Nature, we may think that he is looking in the wrong direction.

There is no natural limitation to the power of mind. The Universe does not make an exception, it is not outside of its grip.

The Primaeval Atom.

A detailed description of the beginning of multiplicity cannot be inferred from the present state of physics, but is not excluded that, when the nature of atomic processes will be really understood, when the reason which has determined the actual value of the mass of the universe, whether it will turn out to be Eddington's $3/2\ 136 \times 2^{256}$ protons or some other figure, shall be known, it will be possible to say something definite about the state of matter, when it was a unique quantum unit.

In the meantime, the best we can do is to call it an Atom, rather in the Greek sense of the world than of this very complicated thing which is a modern atomic nucleus.

Even if Physics can determine, from laws to be discovered, what has been the state of the Primaeval Atom, this will not make possible to present cosmology as a completely deductive theory.

In fact, the splitting of the Atom, can have occurred in many ways and there would be little interest to know their relative probabilities. The one which really occurred might have been very improbable.

Deductive cosmology cannot begin before the splitting has proceeded far enough to reach practical macroscopic determinism.

The beginning of cosmology is therefore expanding space starting from zero and filled up with the pieces of the Primaeval Atom, presumably small, more or less stable, atoms, such as these which are observed today in actual physics.

Any information on the state of matter at this moment must be inferred from the condition that the actual universe has been able to evolve from it.

A quite different situation should have arisen if the universe starting from a zero value of the radius of space, would have been the result of a rebounding universe. I mean, if the universe had once contracted into a very small volume and had just started to rebound.

This has been called a « Phenix » universe. All details of the structure of matter in the contracting period should have been completely burned out and obliterated. The minimum volume needs not to be strictly zero, but it must be small enough to produce in the material tensor the effective negative pressure required. This is incompatible with an astronomical state of matter formed of distinct stars and surely requires that all matter would have been agglomerated into one continuous fluid and even probably that all atomic nuclei would be fused together in what has been called a nucleonic gas.

Strict spherical symmetry would be unlikely and the vanishing volume may have a more or less flattened structure.

With this understanding the Phenix universe theory is quite conceivable.

As we have said, any detail of the contraction period should have been destroyed. Nevertheless the fact that this expansion is made of matter which has already been used would result in its distribution being in a state of maximum entropy, in a state of statistical equilibrium.

The expansion of the Phenix Universe would produce a continuous mass of gas at great temperature where the most probable distribution of velocities known as the Maxwellian distribution would be accurately realised.

On the contrary, the distribution coming out from fresh matter would be a distribution of minimum entropy, i.e. a very unprobable distribution, very far from the thermodynamic equilibrium. This distribution of the velocities will have very little in common with a Maxwellian distribution. The velocities will not be strongly concentrated around a mean value, which could be considered as the velocity of the gas. In fact, it cannot be called a gas. The only feature it has in common with a gas is that it is formed of a great number of individual « molecules », but they have not the Maxwellian distribution which is the real characteristic of a gas. It is better to describe such a state of matter as being corpuscular radiation travelling along in every direction. It is an assembly of corpuscular rays.

I do not see how a useful cosmology can be built by starting from the Phenix nucleon gas.

We shall take as a starting point of our theory the assembly of corpuscular rays and our first problem will be to understand how some gas may arise in this assembly : how, locally, the nuclei can arrange themselves so that they are strongly concentrated around some mean velocity.

2. COSMOLOGICAL PROCESSES

Now that we have made clear what are our fundamental assumptions, we have to try to describe the evolutionary steps which have lead from the assembly of corpuscular rays to the present structure of the universe.

The points we have to discuss are : First, how far it is likely that something of the primitive rays are present in the actual cosmic rays; then, how distinct gaseous clouds may have arisen from the assembly of rays; third, how these gaseous clouds have arranged themselves in proto-galaxies and finally how the galaxies are themselves arranged in clusters with large individual velocities of their individual galaxies.

Cosmic rays.

The effect of the expansion on the corpuscular rays is to reduce their kinetic energy in proportion to the expansion, i.e. as I/R.

This is true for anything which has some peculiar velocity and the same phenomenon would apply as well for the gaseous clouds and later on for the galaxies.

This fundamental property of expanding space may be explained in an intuitive way as follows.

The peculiar velocity we are speaking about, is the difference between the velocity of the corpuscule and the normal velocity at the place where it happens to be (the velocity of the « substratum »). Let us suppose, for definiteness, that the two velocities are in the same direction. Then the corpuscule has a velocity abnormally large for the place where it is. As a consequence of this large velocity, it will reach places farther away where the normal velocity is greater.

Then, with the same velocity as before, the corpuscule will be found to have a velocity less abnormal. The peculiar velocity will be reduced and some elementary arithmetic will show that the reduction works as I/R.

The computation may be done in a simpler way by assuming that the moving particle describes a geodesic of :

$$ds^2 = -R^2 d\sigma^2 + dt^2$$

(ds element of length of an elliptic space of radius I; $R = R(t)$).

The equation of the geodesic is clearly :

$$R^2 \frac{d\sigma}{ds} = C$$

or :

$$R \frac{d\sigma}{ds} = \frac{C}{R}$$

$R \dfrac{d\sigma}{ds}$ and not $R \dfrac{d\sigma}{dt}$ is the momentum, taking account of the variation $\dfrac{dt}{ds}$ of the mass with velocity. Therefore this formula applies to particules with relativistic velocities.

The formula would be correct even for light. In that case it gives the red-shift due to the expression. The result is deduced better from variation of the formula :

$$\sigma = \int_{t_1}^{t_2} \frac{dt}{R}$$

It may be noted, that, mathematically, the problem we have dealt with is identical with the deduction of Clairaut's theorem on the geodesics of a surface of revolution.

Let us apply this result to the problem of the cosmic rays.

Our assumption is that stellar matter arises in some way from the primaeval assembly of corpuscular radiation. We cannot say how far this transformation will be complete but it is safe to assume that the amount of corpuscular rays which has been transformed into gas is not of a very different order of magnitude than what has been left unaffected by the process.

Now, the mass of gas produced is the matter of actual stars and its density can be estimated. Current estimates give in c.g.s units 10^{-30}.

On the other hand, the kinetic energy of the cosmic rays can be estimated from the total of ionisation produced in the various secondary processes due to their absorption in the atmosphere. It is given as pairs of ions per square centimeter. This can be transformed in ergs per cubic centimeter and then in grams per cc. The figure turns out to be 10^{-34}.

If the radius of space did increase by a factor of 10^4, this gives equality at the start. This estimate would lead to a density of matter of 10^{-18} at the time of formation of the gas and this seems to be plausible.

The situation is less satisfactory, if we consider comparison, not in energy, but in proper mass.

In the cosmic rays, proper mass is smaller than kinetic energy, and furthermore proper mass is not reduced by the expansion of space.

The argument in energy is more fundamental, as conservation of energy is much more stringent than conservation of proper mass.

This would lead to an hypothesis of creation of hydrogen, not from nothing, but from transformation into proper mass of the kinetic energy of rays with individual energy some 10^4 times that of the actual cosmic rays.

Such an hypothesis would explain the great abundance of hydrogen in the gaseous matter with which stars have been built.

It can be tested by observation, because it would result in the primary cosmic rays having much less abundance in hydrogen than in stellar matter. Observation is difficult because protons are formed in a secondary way by « stars » atomic processes.

There are already some indications in published observation in favour of this hypothesis and, indirectly, computation of the total intensity from the observed heavy rays, seems to give the total known amount. This would leave little room for a notable contribution from protons.

There may be other significant differences in the abundances of elements in stars and in cosmic rays.

The very small abundance in stars of some light elements is explained because these elements should have been burned out at lower temperature than the one which prevails in stars.

This argument would not work for cosmic rays if they are remnant of the primaeval assembly of corpuscular rays.

This can be tested by observation.

The formation of gaseous clouds.

The formation of gaseous clouds in an assembly of fast moving electrical particles may be explained in the following way.

In such an assembly, one would expect the occurrence of large magnetic fields with very varied character, according to chance fluctuations in the distribution of the rays.

In such a case, it might happen that, occasionally, some magnetic field would have a more or less permanent character. It would form a magnetic region, travelling with some definite velocity.

Such a moving field would be able to select and retain every incoming corpuscule which has nearly the same velocity as the moving field. These particles would be turned into spiral orbits and retained within the field, while faster ones would go through with small deviations.

This mechanism would give the germs of the gaseous clouds. The nearly equal velocities of the particles gathered together would make possible gentle elastic collisions between them and will bring about the Maxwellian distribution.

Then, when gaseous germs will be thick enough, they will be able to absorb even fast moving particles and the cloud will grow up by accretion from these incoming particles. It is in this process of absorption that the hypothetical creation of hydrogen from kinetic energy may take place.

The resulting velocity of the gaseous clouds will be the initial velocity of the magnetic perturbation. Presumably it will be a very large velocity. Therefore one must expect that the gaseous clouds formed out of the corpuscular rays will have very great peculiar velocities. These velocities will be somewhat reduced, later on, by the same process we have explained at the occasion of the cosmic rays, as I/R.

The formation of the galaxies.

The variation of the radius of space R occurs according to Friedmann's equation :

$$\left(\frac{dR}{dt}\right)^2 = -1 + \frac{2M}{R} + \frac{R^2}{T^2}$$

where $M = 4\pi G\rho R^3$ is (in natural units) the total mass, ρ being the density, T is related to the cosmical constant λ by the relation :

$$\frac{\lambda}{3} = \frac{1}{T^2}.$$

The equilibrium value $R = R_E$ is given by :

$$\frac{d^2R}{dt^2} = -\frac{M}{R^2} + \frac{R}{T^2} = 0$$

A universe which would remain permanently with the constant value $R = R_E$ is called an Einstein's universe. It would be in equilibrium but in unstable equilibrium. Then $M = (T/3)^{3/2}$.

There exist solutions of Friedmann's equation where M is somewhat greater than the equilibrium value $(T/3)^{3/2}$. Then the acceleration $\frac{d^2R}{dt^2}$ is negative for small values of R. Velocity started assymptotically with infinite value, slowed down without being completely reduced to zero, or reversed in sign, when approaching the equilibrium $R = R_E$. The radius remained sometimes in approximate equilibrium, but, finally, overtook it and the expansion was resumed with accelerated velocity.

We shall show how the formation of proto-galaxies from the gaseous clouds might arise during the slow motion through Einstein's unstable equilibrium.

The Cosmical Constant.

This theory essentially relies on the cosmological term in the relativistic equations. The legitimacy of the introduction of the cosmological term has been challenged by outstanding authorities, Einstein and also de Sitter. It has been customary to discuss and condemn relativistic cosmology by arguments which would have no meaning if the cosmological term should not have been dropped.

Using the strong expression framed by Eddington in other occasions, we may say that this opinion « continues to work devastation in astronomy », or using more sober language, that it has sterilised many attempts to make use of relativity in astronomy.

I took occasion of being called to contribute to the book offered to Einstein at the occasion of his 70 year birthday to try to vindicate the cosmological constant and to provoke some new consideration of the question by Einstein.

I have not succeeded to convince him, although he agreed, of course, that the question could not be considered as definitely settled and that, in the meantime, the use of the cosmical constant in cosmology was legitimate.

The cosmical term arises naturally from any presentation of the relativity theory and to forget about it, by arbitrarily putting it equal to zero, is not a real solution of the difficulty.

There is one apparently superfluous constant in the theory.

A superfluous constant is a blame if a theory has any right to be considered as complete. For instance, the existence of some definite parameter such as the radius of space in elliptical geometry was a logical inconsistency if geometry has to be considered as a theory complete in itself.

On the contrary, this « superfluous » constant was essential to make possible to relate geometry to gravitation as was achieved by relativity.

Eddington has written : « I would rather reverse to Newtonian theory than to drop the cosmical constant. » Eddington's point of view was that the superfluous cosmological constant would provide the way to connect relativity with quantum theory. It is not the place to discuss how far Eddington succeeded in this endeavour, but it remains that if some extension of relativity towards a broader field, such as quantum theory, has to be achieved the superfluous λ term shall be very much welcomed.

In the meantime, there is nothing to do than to use the cosmical term in astronomical applications. Pending a theoretical determination of λ, comparison with observation may provide empirical determination of the debated constant.

Coming back, for a moment, to the first stages of the expansion, when R was very small, we may deduce from Friedmann's equation the angular distance σ or fraction of the radius of space that light has been able to cover :

$$\sigma = \int_0^t \frac{dt}{R} \simeq \sqrt{\frac{6R}{R_E}}$$

This illustrates, from the gravitational side, what we have discussed earlier from the quantum point of view, how essentially separated one from the other are the different parts of earlier expanding space and how impossible it would be that some statistical equilibrium would have time to prevail in it, if it is not imposed, beforehand, as in the Phenix universe hypothesis.

Proto-galaxies.

In discussing the effect of the unstability of Einstein's Universe on the moving clouds of gases which reach it, it is not necessary to consider the whole space as in Friedmann's equation. It will be sufficient to study in detail what would happen in some region not too great as regards the radius of space.

In such a case, euclidean geometry is a very good approximation, and, for small velocities of expansion as we have to postulate, even Newtonian approximation is completely legitimate.

Of course, we should have to take into account the λ term. This would mean that Poisson's equation would have to be modified by the introduction of a constant ρ_o simply related to the cosmical constant. We should have to write :

$$\Delta V = 4\lambda G (\rho - \rho_o)$$

Such a formula has a much wider significance than Friedmann's equation. It is true that it is valid only in a region of moderate size; but, in its domain of validity, it is free from the simplifying assumption of perfect symmetry which demands Friedmann's equation.

The true density is supposed to be nearly equal to ρ_0. It is not supposed to be a constant, but may have fluctuations from place to place. This will have the effect of dividing space into attractive and repulsive regions.

When these fluctuations are due to fast moving gaseous clouds, it must be expected that such fluctuations will move with large velocities.

Let us consider an attractive region moving with some velocity. Such a region will have a selection effect on incoming gaseous clouds. Clouds which meet this region with a velocity smaller than the velocity of escape will be retained, while faster ones will escape.

The result will be an assembly of clouds which move in the same way, and this will be a proto-galaxy.

Notice that the velocity of the proto-galaxy will be the velocity of the moving attractive region and not, as would occur in other theories, a kind of mean fluctuation in average collisions. With the same velocities of the primitive gaseous clouds, the selection process will give a much greater velocity for the resulting proto-galaxy.

A very important feature of such a proto-galaxy is that it would contain gaseous clouds with as large angular momentum as it is compatible with the value of the velocity of escape.

We do not intend to try to describe the subsequent evolution of the proto-galaxy, but the fact that large angular momenta prevail is a well known requirement of the known structure of galaxies. Presumably, the first effect of the obvious gravitational instability will be a radial collapse with large density in the central region and therefore collisions between clouds of gas leading to star formation.

With large angular momenta, such a process can only be very partial and slower processes leading to collisions in transverse motion must be important. It is obvious that such process may be in relation with the formation of a disc.

We may notice that distinct gaseous clouds are known to exist in the galaxy and that they may be remnants of the primitive ones.

In what follows we shall not make any distinction between proto-galaxies and galaxies. Whether stars are already formed or gaseous clouds still prevail, we shall simply speak of galaxies. With this understanding we shall now deal with the problem of the clusters of galaxies.

3. THE PROBLEM OF THE CLUSTERS OF GALAXIES

A theory of the clusters of galaxies has been proposed by the author some twenty five years ago. It has been reconsidered recently by Bonnor.

It is based on a model with vanishing pressure :

$$ds^2 = -\left(\frac{\partial r}{\cos\chi\,\partial\chi}d\chi\right)^2 - r^2\,(d\theta^2 + \sin^2\theta\,d\varphi^2) + dt^2$$

Matter is formed of a number of independent concentric shells ; each of them is characterised by some value of X. There is no pressure (also $T^{14} = 0$) so that there is no interaction between the neighbouring shells. The density ρ may arbitrarily vary, the mass inside of some shell being an arbitrary function m of X. One has, of course :

$$\frac{dm}{d\chi} = 4\pi\rho r^2\frac{\partial r}{\partial\chi}$$

Then the motion depends on an equation :

$$\left(\frac{\partial r}{\partial t}\right)^2 = -\sin^2\chi + \frac{2m}{r} + \frac{r^2}{T^2}$$

very similar to Friedmann's equation. In fact Friedmann's equation is obtained if $r = R\sin\chi$ and $m = M\sin^3\chi$ (M a constant). But, if for instance $m = M\sin^4\chi$, each shell moves according to some solution of Friedmann; but the solution is different for different values of χ and may even be of a different type.

For large values of χ, the motion may be of the ever-expanding type, as we have supposed to be the case for the universe at large. For some value of χ, it may be the elementary solution, with two coincident roots of the quartic, which describes the assymptotic approach towards Einstein's equilibrium, while for smaller values, it is of the collapsing type.

It was tempting to identify the collapsing regions with galaxies and the assymptotic regions, which remain in equilibrium while the general expansion is resumed, should be identified with the clusters of galaxies.

There is some observational evidence in support of this theory. The density of the clusters is the cosmical density ρ_0 and it can be inferred from observations of the red-shift. The test was satisfactory.

Nevertheless, I think now that this theory was very incomplete and that its fundamental assumption that pressure could be neglected is not justified.

It is true that the p term, as far as it expresses the change of mass with velocity, may be safely neglected because the astronomical velocities are small with regard to that of light.

But the peculiar velocities of the galaxies (or of the clouds which form them) realises a mixture of matter which makes impossible to consider matter as formed of independent concentric shells as was done in the former theory. We have seen already how the large velocities of the gaseous clouds modify and regularise the collapsing process which gives the proto-galaxies with their great velocities.

A more convenient theory of the clusters of galaxies may be obtained by considering fluctuations in an Einstein's universe with pressure.

The galaxies (or proto-galaxies) move (without encounters) in a universe in equilibrium. We simplify the problem by the assumption that, at each place, the distribution of velocities is uniform within a sphere of some radius, taken as unit of velocities.

As we shall consider phenomena on a small scale, we shall adopt Newtonian approximation, i. e. Poisson's equation with the modifying ρ_0.

In equilibrium ρ is equal to ρ_0.

We shall study infinitesimal perturbations of this model. We assume spherical symmetry. In this way, it is easy to work out the equations of the problem. It is found that static solutions exist. They may be described as stationary waves occurring in the assembly of the moving galaxies. They may form a permanent cluster of galaxies in the Einstein's universe. It is even possible to work out perturbations of a finite amplitude.

These condensations have a definite size which is related to the maximum velocity, which has been taken as a unit.

It is convenient to choose as unit of length, the length travelled with the maximum velocity during a time $T/\sqrt{3}$ i.e. some 2×10^9 years. Then one finds a radius of the condensation, or wave length, of the standing wave, equal to $\pi/\sqrt{3}$.

Infinitesimal condensations which vary with time may be analysed in condensations described by products of a time factor by a space factor. Then, the time factor is found to be an exponential $e^{\theta t}$. Units are such that $\theta = 1$ refers to the uniform expansion or contraction.

For real θ such condensations are larger than in the static case ($\theta = 0$) the larger the condensation, the larger is θ ; $\theta = 1$ means condensations (or rarefactions) of infinite size.

We see how the mixing effect due to the peculiar velocities, i.e. the pressure effect, introduces some regularity in the results of the instability.

In the long run, terms with the greatest exponent must finally prevail, therefore instability will be disrupted in the same sense in the whole space.

If it occurs in the sense of expansion, any collapsing tendency of finite size will have a time factor $e^{\theta t}$ which increases more slowly than the e^t of the general expansion.

This can be studied rather easily, when condensations are supposed to be infinitesimal. It would be much more difficult to extend the results to condensations of a finite size.

The problem of the formation of the clusters of nebulae may be studied in the following way :

Let us consider, in the Einstein's equilibrium with pressure, some condensation increasing exponentially as say $e^{t/2}$; when the expansion will start again with time factor e^t, the general expansion will obviously counteract the development of the local condensation.

The problem will be : How the condensation will finally dissipate?

The chief point of interest is, if the condensation is smoothed out uniformly, or, on the contrary, if it fades out from the edge, leaving a central part of large density.

In the second case, such a model may lead to a reasonable interpretation of the clusters of galaxies.

In such a problem, boundary conditions are essential.

The boundary condition must express that the clusters of galaxies are not isolated, but that each of them is surrounded by similar clusters.

As spherical symmetry must be assumed, for mathematical convenience, it is necessary to introduce some kind of averaging in the representation of the surrounding clusters.

One may imagine the cluster under study as surrounded by a sphere of radius varying as the radius of space.

Individual galaxies cross this spherical boundary, the same amount passing inside as outside.

Mathematically, it is more convenient to imagine that the same galaxy rebounds inside the boundary. The boundary itself must have the acceleration given by the gravitation field of the cluster.

In that way, the boundary condition can be strictly formulated.

Some years ago, I had thought convenient to introduce in the problem a very efficient simplifying assumption which appeared to be reasonable and that I had called quasi-isotropic solutions.

When I realised more clearly what were the boundary conditions, I found out that they are not compatible with the quasi-isotropic condition.

That means that it is inevitable to consider motions in every direction and that it is not possible to separate the radial motion from the rest of the problem. That was the point of the quasi-isotropic solution.

In fact the velocity distribution may be described by two variables, V, mean of radial velocities with the same transversal motion, W, the square of the total common velocity. These two variables depend, not only on r and t, but on a third variable k, which it is convenient to take as the constant of angular momenta of each trajectory.

It is, of course, sufficient to work out the motion of the galaxies at the boundary of the velocity distribution, as Liouville's theorem guaranties that the occupied volume in velocity-space will be occupied with uniform density.

It is possible to write down Taylor's series expressing V and W by powers of the variables r and k. (t is taken in the exponential factor, as explained earlier).

It has been found that these series practically converge up to the boundary and, with some labour, the computation has been carried out for $\theta = 1/2$.

As fast means of computation become to be available in Belgium, it is hoped that these results may be used as a starting point for the solution of the problem of clusters of galaxies.

The method must be to compute a family of trajectories moving in the gravitational field. From the trajectories the occupied volume of velocity should be determined in a number of points. This would give the density and then again the gravitational field, whose value could be ameliorated and produced further on.

An essential difficulty would be the convenient choice of an available parameter which will describe how far the condensation has been able to proceed, before the general expansion is well started.

According to this theory, the cluster of galaxies would be a more or less stable standing wave. While the condensation would remain more or less permanent, individual galaxies will escape from the cluster and be replaced by other ones coming from the surrounding field.

It is, essentially, the degree of permanency of such an exchange of galaxies between the cluster and the field, exchange which has begun during the equilibrium, that the theory must work out.

One may expect, as in the static case, that some relation would arise between the size of the cluster and the velocity of its members.

Rough estimates indicate that the relation found in the statical case would give clusters of a somewhat larger size than the real ones.

If, as said earlier, clusters dissipate from the edge and not in a uniform way, a reduction of the size should be expected and may give the observed size.

It is a general feature in stationary mechanical systems that strong condensations are connected with large velocities. This is the essence of Liouville's theorem and, equivalently, it appears in the virial theorem or in Bernouilli's theorem.

The point of the theory is that it works even when gravitation is balanced by the cosmic repulsion.

I think that this was not quite evident and it is now established, at least in the infinitesimal case.

CONCLUSION

In concluding this report, we must emphasise some points connected with astronomical observation, which are connected with the theory, supposed to be accepted.

First, as had already been pointed out for cosmic rays, one must expect that each evolutionary step has been incompletely fulfilled, so that primitive matter engaged in this process still exists and in quantity of essentially the same order of magnitude as what has evolved from it.

From this, we may deduce that remnants of the original gaseous clouds must be in the same amount as matter condensed into stars. This is found from observation. We may also conclude that invisible gaseous matter between the galaxies is spread with a density nearly the same, 10^{-30}, as that of the galaxies themselves.

This is contrary to the opinion of many cosmologists and is not decided by observation.

It is essential for our theory, because, otherwise, one should have to expect that cosmic rays would have suffered an appreciable absorption during their long journey through expanding space.

It is also in connection with the fundamental point that, actually, the cosmological term is the main term in Friedmann's equation, and therefore that Hubble's $T_H = \frac{r}{v} = 4 \times 10^9$ years is a good approximation for the cosmological T, the subtangent of the curve R versus t. This gives also the value $\rho_0 = 10^{-27}$ of the cosmic density and therefore the present ratio R/R_E of about 10.

The main uncertainty, in interpreting Friedmann's equation, is the unknown relation between M and T. This is not a very serious matter.

The uncertainty relates to the precise age of the Universe or, otherwise speaking, of the magnitude of the delay during equilibrium.

Theoretically this delay might be infinite if the theoretical relation $T = M \times 3^{3/2}$ would be exactly fulfilled. But this has not to be taken too seriously, a logarithmic infinity is the weakest form of infinities. Physically, it means some indetermination of the exact epoch.

If the critical relation is fulfilled up to one per cent the age t/T ($T = 4 \times 10^9$ years) turns out to be about 5. It is increased by 3 for any added decimal in the accuracy with which the relation is fulfilled. For an absurd approximation of one in a million, the relative age would be increased only by 15.

Notwithstanding this fundamental uncertainty, one may confidently put the age of the Universe somewhere between 20 and say 60 times 10^9 years.

The recent passage of the Universe through a period of gravitational instability, besides being shown by the clusters of galaxies, is manifested also by the large fluctuations which are found in the repartition of the galaxies and in the distribution of their velocities.

This would make very uncertain any attempt to determine the type of expansion and the character of space from counts of galaxies and velocity determinations.

In discussing statistical data related to the clusters, the effect of the cosmical repulsion may alter essentially the conclusions which would otherwise be substantiated.

The interpretation of the clusters of galaxies as permanent condensations of passing by galaxies outweighs the value of the arguments which had lead some cosmologists to question the general exactness of the determination of the masses of the galaxies from their observed rotation or, alternatively, to fill space with a large amount of invisible matter.

If the solution of the problem of the clusters of galaxies turns out to be in the expected direction, the objection against our interpretation of Friedmann's equation shall be removed. These large velocities will be found to have an essential cosmologic significance.

They cannot be reconciled with some type of chance formation such as was considered by Jeans. Such a process, besides being too slow, will provide only condensations with small peculiar velocities.

On the contrary, they would be a sign of the large velocities of the clouds of gas from which proto-galaxies have been formed and these velocities themselves would be a sign of the origin of these clouds from the primitive radiation.

Discussion of Lemaitre's Report

Dr. Pauli. — I wish to emphasize the independence of the question whether space is finite and closed from the other question whether the cosmological constant is zero. Particularly, the latter is compatible with the former according to Friedman's solution of the expanding universe. I would like to hear Dr. Lemaitre's view on it.

Dr. Lemaitre. — I agree on this point. Nevertheless this would not introduce anything like the instability on which my theory is based.

Dr. Heckmann. — The only known way of designing the fundamental equations of Einstein's theory of gravitation leads necessarily to the Λ-term. If one wants to drop Λ, one has to state that a new axiom, either theorical or observational in origin, is used.

If one stipulates that space is finite, the sign of curvature of space is still open.

Dr. Lemaitre. — It is true that finiteness of space can be reconciled with a negative curvature, but that can be done only in a very artificial way, while finiteness directly follows in the case of positive curvature.

Dr. Wheeler. — Regarding the question of a cosmological constant, it is appropriate to recall why Einstein always regretted his introduction of this new and invented term into relativity theory. The analogy between gravitation theory and electric theory brings to attention Poisson's equation :

$$\nabla^2\Phi = -4\,\pi\rho\,.$$

There is no objection in principle to substituting the equation :

$$\nabla^2\Phi + \varkappa^2\Phi = -4\,\pi\rho$$

but experiment speaks against the added term. The analogy in structure between gravitation and electric theory therefore suggests — Einstein pointed out — that there should also be no added term in the field equations of general relativity. This principle of analogy can be regarded as the supplementary principle referred to by Professor Heckmann — a principle that demands $\Lambda = 0$.

Historically, the cosmological term was introduced to give a universe of unchanging size. It lost its justification in Einstein's eyes when it was discovered that the universe is *not* constant in size.

Dr. Lemaitre. — Of course that was Einstein's point of view. It is appropriate to recall that other attitudes have been held and that the cosmical constant has been considered as fundamental in other theories. We can quote Weyl : « The cosmological factor which Einstein added to his theory later is part of ours from the very beginning. » And Eddington, after quoting this sentence added : « Not only does it unify the gravitational and electromagnetic fields, but it renders the theory of gravitation and its relation to space-time measurement so much more illuminating, and indeed self-evident, that return to the earlier view is unthinkable. I would as soon think of reverting to Newtonian theory as of dropping the cosmical constant. » (The expanding universe — page 24) and later, page 104 : « Being in this way based on a fundamental necessity of physical space, the position of the cosmical constant seems impregnable; and if ever the theory of relativity falls into disrepute the cosmical constant will be the last stronghold to collapse. To drop the cosmical constant would knock the bottom out of space. »

Dr. Klein. — If we take the standpoint that general relativity theory, based on the invariance of the laws of physics against general coordinate transformations, and quantum theory are both fundamental parts of our knowledge of the laws of nature, then we are led to consider a length of the order of magnitude : $\sqrt{\frac{\gamma h}{c^3}}$, γ = gravitational constant, h = Planck's quantum of action, (c = vacuum velocity of light) as the natural unit of length. In fact, with unit of lenght : $= \sqrt{\frac{8\gamma h}{c^3}} \approx 10^{-32}$ cm the action integral divided by $\hbar c$ (which plays, as well known, an important rôle in the present formalism of quantum field theory) is particularly simple in that the coefficient of the curvature scalar appearing in the gravitational part of this quantity is just unity. Now, with this unit of length an additional term with the cosmological constant would have a coefficient $\sim 10^{-118}$, which would be very strange in a fundamental theory. This would make it plausible that the cosmological constant vanishes. To get such a constant as a secondary consequence of fundamental equations without it would probably be excluded.

Dr. Lemaitre. — I do not see how purely dimensional argument can dispose of the cosmical constant. I suppose that any Fundamental Theory must introduce, as does the original one proposed by Eddington, some large non-dimensional number like the one that he describes as the total number of particles in the universe. This may essentially alter the picture resulting from the adoption of one or another natural unit of length.

Dr. Bondi. — How can a process of capture operate when, in a static state, you cannot get rid of energy by radiation on balance? Does this not affect your picture of the formation of proto-galaxies in the Einstein state?

Dr. Lemaitre. — I agree that the term « capture » that I have just used is misleading. As I did in the written report, I prefer to describe the process as a « selection effect », by which the moving collapsing regions retain in proto-galaxies the matter coming with convenient energy.

Dr. Heckmann. — As I understand it, Lemaitre simply considers those parts of the whole distribution where, in a certain volume of space, energy is negative; these are the « collapsing » regions.

Dr. Bondi. — I do not object to the existence of contracting regions of negative energy, only to the capture by such regions of clouds moving with positive energy.

Dr. Schatzman. — In your theory of clusters of galaxies, you obtain a distribution much flatter than the actual distribution. The observed distribution is ten to hundred times more concentrated than your theoretical distribution.

Dr. Lemaitre. — The theoretical distributions that I have computed for static solutions essentially depends on a condition of isotropy in the velocities. It would be of interest to inquire if an higher degree of concentration can be obtained if this rather artificial assumption is discarded.

Dr. Ambartsumian. — There is now evidence in favour of the existence of sources of cosmic rays in the Universe (Crab Nebula, stars and even the sun). Is Dr. Lemaitre of the opinion that there are two different kinds of cosmic rays?

1. The remnants; 2. The cosmic rays of recent origin?

Dr. Lemaitre. — Yes, that is my opinion.

Dr. Bondi. — The more we can account for the odd features of our world (cosmic rays, heavy elements, etc.) in terms of processes that can take place in the existing state of the universe, the less are we driven to considering any different state of the universe in the past.

Dr. Hoyle. — It is true that given the three constants c, h and Λ, one can construct a natural unit of time, of length, and of mass, just as in the steady state theory. But it seems to me a noteworthy distinction between the two cases that, while in the steady state theory the natural units have an immediate relation to the *present* condition of the universe, in Lemaitre's case the natural units only become dominant in the ultimates asymptotic condition of the Universe when the mass density tends to zero.

Dr. Lemaitre. — I believe that, in the present condition of the universe, the density is small enough for the cosmic term being already dominant. In other terms, the cosmical T is not significantly different from Hubble's T_H. In any way, the necessary correction can be computed.

Dr. Mac Crea. — Apart from any particular theory such as general relativity, physical laws deal with local peculiarities against the general background of the universe. It is possible that the meaning of physical law becomes less definite, or at any rate different, as we attempt to deal with larger and larger parts of the universe.

As regards the theory of general relativity, it is possible to regard it as tentative, i.e. we explore the consequences of using Riemannian geometry. There is nothing compulsory in this or in the particular principles stated by Heckmann. Thus it is our good fortune that we have to consider only two tensors, the Einstein tensor and the fundamental tensor, or rather a linear combination of these. The coefficient of the fundamental tensor is proportional to the cosmical constant. We still have to consider whether this is zero or non-zero, according to the principles stated by other speakers. From a general point of view, however, this need not be regarded as a fundamental physical problem. It is more an accident of the mathematical approach; were we to start with a more general geometry, we should have more such « constants » to deal with in the same way.

Dr. Gold. — The laws of physics are devised by us as convenience to describe by means of generalisations the large variety of circumstances that may arise. If there were many universes that showed a variety of features we should like to use the same methods. But with only one universe to look at the justification for such an approach disappears. The laws of physics ought on a large scale then to become just a description of what there is.

Dr. Mac Crea. — I wish to say how thoroughly I agree with what Gold has said about the « degeneration » of physical laws.

Dr. Bondi. — A theory is a brief description of experimental results. If one extrapolates, as one is bound to do in cosmology, one must be careful to distinguish between the formulation of the theory (usually quite rigid) and its experimental foundations with their concomitant inaccuracies. There is accordingly no unique way of getting from the small to the large and the guidance given by a theory is uncertain to this extent. Thus any choice is arbitrary, in some measure, and if a modification of the theory can be made that leads to a simpler large scale picture without conflicting with the local evidence, then this is a legitimate procedure which is likely to be best suited to the circumstances.

Dr. Heckmann. — The limit in the freedom of extrapolation of laws of nature is defined by our aim to understand events in large distances.

« Understanding » means interpretation by known local laws.

Dr. Hoyle. — I would say in relation to your comment that any change in the laws of physics is justified so long as predictions of new observations can be made.

Dr. Gold. — In extrapolating the laws of physica to a very large scale one frequently has to admit ignorance; it is then not a question of one « straightforward » and a « modified » extrapolation, but just of having to find a reason for choosing one of several possible and equally consistent ones.

Some considerations regarding the earlier development of the system of galaxies

by O. KLEIN

INTRODUCTION

Since the days of Galilei the trend of physics has been the search for fundamental laws which could be tested by means of observations and experiments referring to here and now and which do not imply any knowledge of the whole universe. Cosmology in its two main lines — the expanding universe and the steady state world — would seem to mean a radical departure from this once new and revolutionary but now and so far useful habit of physical investigation.

As wellknown, the revival of cosmology came with Einstein's theory of the closed universe which as an essential background had the so-called Mach's principle acco ding to which inertia had to be explained as gravitation-like effects from the distant masses of the universe. In fact, although Einstein had succeeded to incorporate what would seem the really fruitful part of Mach's principle in his general theory of relativity — based on the equivalence principle and the viewpoint of covariance under general transformations of the space-time coordinates — without any reference to the world at large he himself was dissatisfied with the need of boundary conditions. Indeed, boundary conditions are necessary when given physical phenomena are described without any reference to the world at large and characteristic of what may be termed Galilean physics even if the greater part of it — such as Newtonian mechanics, electromagnetism, quantum theory and the whole edifice of relativity theory its cosmological applications excepted — has been developed long after Galilei's own time. In relativity theory the essential boundary condition is that outside of the region — as well in time as in space — of the phenomenon under consideration we should be able to choose a flat space-time frame corresponding to the situation in special relativity theory. This general boundary condition is, of course, closely related to the observational philosophy of Einstein and Bohr. In this connection the implication of this boundary condition is important in that it allows only of such solutions of the Einstein gravitational equations where the total mass within the region under consideration is limited by means of the wellknown Schwarzschild singularity, while the opposite is true for the solutions used in cosmological considerations. Since there seems to be no natural way of passing the Schwarzschild singularity the application of such solutions to problems of natural science would seem to put an unexpected end to our questioning not motivated in the laws of Nature themselves.

Without denying the possibility that observations may once force us to accept one or other of the cosmological solutions, I want to stress — since general relativity is usually involved in discussions about cosmology — the essential difference between the Mach's principle and the so-called cosmological postulate on one side and the principles on which relativity theory is based on the other side. Thus, from the point of view of these latter principles the Mach's principle would belong to the class of statements which can neither be disproved nor proved. The same may possibly hold with respect to the cosmological postulate. Thus the hypothesis that our system of galaxies is just an enormous stellar system limited in the ordinary sense of the word governed by the ordinary laws of physics including Einstein's gravitational theory would not in itself contradict the cosmological postulate but would make it inapplicable as a heuristic principle. In fact, there may be any number of similar systems in the world in different stages of evolution — perhaps for ever outside the reach of our observations.

1. INTERPRETATION OF EDDINGTON'S RELATIONS

As an indication in favour of the just mentioned hypothesis we may perhaps consider the following simple interpretation of the famous Eddington relations between cosmological and atomistic quantities. They are, as wellknown :

$$\frac{\text{Radius of universe}}{\text{Radius of electron}} \sim (\text{number of atoms in universe})^{1/2}$$

$$\sim \frac{\text{electric attraction}}{\text{gravitational attraction in hydrogen atom}} \sim 10^{39}$$

Let e be the elementary electric charge, m_e, m_p the masses of the electron and the proton, respectively, γ the gravitational constant,

$$d = \frac{e^2}{m_e c^2}$$

the « radius of the electron » and R the « radius of the universe », a quantity which may be taken as of the order of magnitude of the distance to the remotest observable galaxies, which again is of the order of magnitude cT, where c is the vacuum velocity of light and T the reciprocal Hubble constant. Then the Eddington relations read :

$$\mathrm{R} \sim \frac{e^2}{\gamma m_e m_p} d, \quad \mathrm{M} \sim \left(\frac{e^2}{\gamma m_e m_p}\right)^2 m_p \qquad (1)$$

where M is the « mass of the universe ».

Let us now assume that our metagalactic system at some decisive stage of its evolution consisted of a huge hydrogen gas cloud of a mass value approaching the upper limit compatible with its dimensions according to general relativity theory, the limit being set by the singularity of the outward Schwarzschild solution. And let us further assume that it contains radiation exerting pressure by means of Thomson scattering on the electrons of the cloud — scattering cross section $= 8\pi/3\ d^2$ — whether they be free or the radiation of sufficiently high frequency for this assumption to hold. For the radiation pressure to play a role it is necessary that the mean free path of a photon

$$\frac{1}{\frac{8\pi}{3}\ \frac{\rho}{m_p} d^2},$$

where ρ is the average mass density, is smaller than the radius R of the cloud. Thus we have the following two inequalities :

$$\frac{2\gamma M}{c^2 R} = \lambda < 1, \quad \frac{2d^2 M}{m_p R^2} = \frac{1}{\mu} > 1 \qquad (2).$$

Rearranging the relations so as to express R, M and ρ in terms of λ and μ we get :

$$R = \lambda\mu \frac{c^2 d^2}{\gamma m_p} = \lambda\mu \frac{e^2}{\gamma m_p m_e} d = \lambda\mu \times 6.3 \cdot 10^{26} \text{ cm.}$$

$$M = \frac{\lambda^2\mu}{2} \frac{c^4 d^2}{\gamma^2 m_p} = \frac{\lambda^2\mu}{2} \left(\frac{e^2}{\gamma m_p m_e}\right)^2 m_p$$

$$= \lambda^2\mu \times 4 \cdot 10^{54} g = \lambda^2\mu \times 2 \cdot 10^{21}\, M_{\odot} \qquad (3)$$

$$\rho = \frac{1}{\lambda\mu^2} \frac{3}{8\pi} \left(\frac{e^2}{\gamma m_p m_e}\right)^{-1} \frac{m_p}{d^3} = \frac{1}{\gamma\mu^2} \times 4 \cdot 10^{-27} \text{ g/cm}^3$$

We see that with λ and μ approaching unity, M is of the order of magnitude of the observed part of the metagalactic system while the values of R and ρ point to a somewhat more condensed state of the system than the present one.

Let us now make the further assumption that the state considered is the result of a foregoing contraction followed by an expansion of the cloud under the action of its proper gravitation and then of radiation pressure from a still earlier state of extreme dilution, the contraction having been brought to a standstill and later reversed by means of the pressure of radiation created in atomic collisions at the expense of gravitational potential energy. Then, assuming that, due to the comparative opacity during the relevant evolutionary stages and the immense dimensions of the cloud, but little of the radiation has been able to escape, the total energy of the system at the present stage — i.e. the sum of its potential gravitational energy and the Hubble motion kinetic energy — would have a negative value not far from zero. With the Hubble relation :

$$v = \frac{r}{T} \qquad (4),$$

where v is the velocity of a galaxy and r its distance from the center of the system, this assumption means that :

$$\frac{8\pi}{3} \gamma\rho' T^2 \gtrsim 1 \qquad (5)$$

where ρ' is the present average density in the metagalactic system. With the recent value of T ($13 \cdot 10^9$ years), this gives :

$$\rho' \sim 10^{-29} \text{ g/cm}^3.$$

2. EQUATIONS GOVERNING THE LATER PHASE OF THE EVOLUTION OF THE SYSTEM

It is clear that processes so far only roughly indicated for the evolution of the hypothetical gas cloud must be closely studied before one can at all be certain of the conclusions tentatively drawn above. Especially a study of the creation of radiation and its frequency distribution is needed in order to judge the permissibility of the assumption suggested by the Eddington relations that the Thomson scattering cross section $\frac{8\pi}{3} d^2$ gives the order of magnitude of the relevant scattering of radiation within the cloud.

Leaving this question aside for the moment some preparations will be outlined here for the study of the later phase of evolution of the system where under the influence of radiation pressure the expansion gradually develops. Thus, we shall write down the equations governing this phase assuming without further justification that the effect of radiation on the matter of the cloud is simply due to Thomson scattering. Further we assume central symmetry and that the velocities in question are small enough compared to the vacuum velocity of light for the neglect of second order quantities. This neglect is certainly justified for the earlier stages of the acceleration process, where, however, the general relativity corrections are probably important. For the later stages both kinds of corrections may probably be neglected up to very great distances from the center, maybe for the entire system if no high precision is demanded. Anyway the special relativity corrections are easily introduced when needed.

Let the line element ds be given by :

$$ds^2 = \alpha dr^2 + r^2 (d\theta^2 + \sin^2 \theta \, d\varphi^2) - c^2 \beta dt^2 \qquad (6),$$

where α and β are functions of r and t. Let T_i^k be the components of the momentum-energy density tensor of matter and radiation together, whereby the indices i, k take the numbers 1, 2, 3, 4 corresponding in that order to the four coordinates r, θ, φ, t. Then the central symmetry assumption implies, as wellknown, that apart from the diagonal components T_i^i only T_1^4 and T_4^1 are different from zero. Putting :

$$M(r) = \frac{c^2}{2\gamma} r \left(1 - \frac{1}{\alpha}\right), \quad \alpha = \frac{1}{1 - \frac{2\gamma M(r)}{c^2 r}} \qquad (7),$$

where M (r) like α is also, in general, a function of t, we have for the determination of α and β the following equations :

$$\frac{\partial M(r)}{\partial r} = 4\pi r^2 \left(- \frac{T_4^4}{c^2}\right), \quad \frac{c^2}{2\alpha\beta} \frac{\partial \beta}{\partial r} = \gamma \left(\frac{M(r)}{r^2} + 4\gamma r \, T_1^1\right) \qquad (8).$$

The assumption that the system has originally condensed from very large dimensions to a state approaching the Schwarzschild limit implies that the gravitational energy lost during contraction is comparable with the total mass energy Mc^2. As shown by a simple estimate this would entail that gas pressure is entirely negligible in comparison with radiation pressure if anything approaching, however distantly, a temperature equilibrium between matter and radiation is reached. In fact, the gravitational energy available would in a temperature equilibrium correspond to a temperature $\sim 100^\circ$ K, where with the matter density $\sim 10^{-27}$ g/cm^3 almost all thermal energy belongs to radiation. Thus, for the matter part of the momentum-energy tensor we can take :

$$\theta^{ik} = \rho_m \, v^i \, v^k \qquad (9),$$

where v^i are the components of the velocity four-vector and where ρ_m is the mass density in a locally inertial rest frame. For these we have : $\left(\text{with } v = \frac{dr}{dt}\right)$

$$v^1 = v^4 \, v, \quad v^2 = v^3 = 0, \, v^4 ,$$

whereby :

$$\beta c^2 (v^4)^2 - \alpha (v^1)^2 = c^2 ,$$

or :

$$v^4 = \frac{1}{\sqrt{\beta - \alpha \frac{v^2}{c^2}}} \approx \frac{1}{\sqrt{\beta}}$$

Thus :

$$\theta_i^k = \begin{pmatrix} \rho_m \frac{\alpha}{\beta} v^2 , & 0, & 0, & \rho_m \frac{\alpha}{\beta} v \\ 0 , & 0, & 0, & 0 \\ 0 , & 0, & 0, & 0 \\ -\rho_m c^2 v , & 0, & 0, & -\rho_m\left(c^2 + \frac{\alpha}{\beta} v^2\right) \end{pmatrix} \qquad (10).$$

For the momentum energy tensor of radiation we have with corresponding approximation :

$$S_i^k = \begin{pmatrix} p , & 0, & 0, & \overset{4}{S}_1 \\ 0 , & p, & 0, & 0 \\ 0 , & 0, & p, & 0 \\ \overset{1}{S}_4 , & 0, & 0, & -3p \end{pmatrix} \qquad (11),$$

p being the radiation pressure and :

$$\overset{1}{S}_4 = -\sqrt{\frac{\beta}{\alpha}} \frac{L(r)}{4\pi r^2}, \qquad \overset{4}{S}_1 = \frac{1}{c^2}\sqrt{\frac{\alpha}{\beta}} \frac{L(r)}{4\pi r^2} \qquad (12),$$

where $L(r)$ is the luminosity in the local rest frame at the distance from the center determined by r. Like $M(r)$ the quantity $L(r)$ is, of course, in general a function of t also. For any value of r and t we have thus to determine the following six functions $M(r)$, β, ρ_m, v, p, $L(r)$. For these we have the following equations, namely :

(1) The continuity equation for the conservation of material particles, which reads ($v \ll c$) :

$$\frac{\partial}{\partial t} (\sqrt{\alpha}\rho_m) + \frac{1}{r^2} \frac{\partial}{\partial r} (\sqrt{\alpha} r^2 \rho_m v) = 0 \qquad (13)$$

(2) The conservation law for matter momentum (equation of motion for matter) :

$$\frac{1}{\sqrt{\beta}}\left(\frac{\partial}{\partial t}+v\frac{\partial}{\partial r}\right)\left(\frac{\alpha}{\sqrt{\beta}}v\right)+\frac{1}{2\beta}\frac{\partial\beta}{\partial r}=f_1 \qquad (14),$$

where f_1 is the component in the r-direction of the radiation force pro unit volume four-vector.

(3) The conservation law for radiation momentum :

$$\frac{1}{\sqrt{\alpha\beta}}\frac{\partial}{\partial t}\left(\frac{\alpha L(r)}{4\pi r^2}\right)+\frac{\partial p}{\partial r}+\frac{2p}{\beta}\frac{\partial\beta}{\partial r}=-f_1 \qquad (15).$$

(4) The conservation law for radiation energy :

$$\frac{1}{4\pi\sqrt{\alpha\beta}\quad r^2}\frac{\partial\beta L(r)}{\partial r}+3\frac{\partial p}{\partial t}+2\frac{\dot{\alpha}}{\alpha}p=f_4 \qquad (16)$$

where f_4 is the time component of the just mentioned radiation pressure four-vector.

Assuming Thomson scattering we have for this four-vector the following covariant expression :

$$f_i=-\frac{8\pi}{3}d^2\frac{\rho_m}{m_pc}\left(S_{ik}v^k+S_{kl}\frac{v^kv^l}{c^2}v_i\right) \qquad (17),$$

where S_{ik} are the covariant components of the radiation momentum-energy tensor*. As is easily seen the f_i fulfill the condition necessary for a force four-vector :

$$v^i f_i = 0 \qquad (18),$$

which makes the energy conservation law for matter a consequence of the equations of motion.

In the approximation in question we have simply :

$$f_1=\frac{2}{3}\frac{d^2\rho_m}{m_pc}\frac{\sqrt{\alpha}\;L(r)}{r^2},\quad f_4=-f_1v \qquad (19).$$

We need one more relation, namely for $\dot{\alpha}$, the time derivative of α, which with proper boundary conditions is a consequence of

* As kindly pointed out to me this expression for the radiation force was already given by Landau and Lifschitz in their admirable book on Field theory, p. 287.

the equations already given but which is most easily derived from the field equation :

$$R_4^1 = -\frac{8\pi\gamma}{c^4} T_4^1 \qquad (20),$$

where with the line element (6) the component R_4^1 of the curvature tensor is equal to : $-\frac{\dot{\alpha}}{\alpha^2 r} = -\frac{2\gamma \dot{M}(r)}{c^2 r^2}$, so that we obtain :

$$\dot{M}(r) = \frac{4\pi}{c^2} r^2 T_4^1 \qquad (21).$$

The equation (20) is seen to be the integrated form of the divergence relation :

$$\frac{\partial T_4^4}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 \; T_4^1) = 0 \qquad (22)$$

and contains the wellknown result that as soon as T_4^1 vanishes faster than $1/r^2$ for $r \to \infty$ the expression for M (r) — if convergent — goes towards a constant value, the total mass of the system, α and β tending simultaneously towards the outward Schwarzschild solution. A similar theorem holds according to (13) for the quantity $\int_0^r 4\pi \; \sqrt{\alpha} \; r^2 \; \rho_m \; dr$, which represent the sum of the rest masses of the atoms in a sphere of radius r around the center. If, as we have assumed, our system has condensed from a state of very large dimensions and low density where relativistic effects were negligible, and if very little energy has escaped during the evolution, then the two masses just mentioned, which coincided in the beginning, will continue to be almost equal, so that :

$$\int_O^R 4\pi r^2 \left(\rho_m + \frac{3p}{c^2}\right) dr \approx \int_O^R 4\pi\sqrt{\alpha} \;\; r^2 \rho_m \; dr \qquad (24),$$

R being the radius of the system appropriately defined.

In summarising the foregoing considerations we shall use the following convenient system of units :

length : $\frac{c^2 \, d^2}{\gamma m_p} = 6.3 \times 10^{26}$ cm

time : $\frac{cd^2}{\gamma m_p} = 6.7 \times 10^8$ year mass : $$\frac{c^4 d^2}{\gamma^2 m_p} = 8 \times 10^{54}\text{g} = 4 \times 10^{21}\ M_{\odot}$$

mass density : $$\frac{\gamma m_p^2}{c^2 d^4} = 3.2 \times 10^{-26}\ \text{g/cm}^3$$

luminosity : $$\frac{c^5}{\gamma} = 3.6 \times 10^{59}\ \text{erg/sec}$$

With these units our equations become :

$$\frac{\partial M(r)}{\partial r} = 4\pi r^2 (\rho_m + 3p), \qquad \frac{1}{2\alpha\beta}\frac{\partial\beta}{\partial r} = \frac{M(r)}{r^2} + 4\pi r p,$$

$$\alpha = \frac{1}{1 - \frac{2M(r)}{r}} \tag{A}$$

$$\frac{\partial}{\partial t}(\sqrt{\alpha}\rho_m) + \frac{1}{r^2}\frac{\partial}{\partial r}(\sqrt{\alpha}\, r^2 \rho_m \mathrm{v}) = 0 \tag{B}$$

$$\frac{1}{\sqrt{\beta}}\left(\frac{\partial}{\partial t} + \mathrm{v}\frac{\partial}{\partial r}\right)\left(\frac{\alpha}{\sqrt{\beta}}\mathrm{v}\right) + \frac{1}{2\beta}\frac{\partial\beta}{\partial r} = \sqrt{\alpha}\,\frac{2}{3}\frac{L(r)}{r^2} \tag{C}$$

$$\frac{1}{\sqrt{\alpha\beta}}\frac{\partial}{\partial t}\left(\alpha\frac{L(r)}{4\pi r^2}\right) + \sqrt{\alpha}\,\frac{2}{3}\rho_m\frac{L(r)}{r^2} + \frac{\partial p}{\partial r} + \frac{2p}{\beta}\frac{\partial\beta}{\partial r} = 0 \tag{D}$$

$$\frac{1}{4\pi\sqrt{\alpha\beta}\,r^2}\frac{\partial\beta L(r)}{\partial r} + \sqrt{\alpha}\,\frac{2}{3}\rho_m\frac{L(r)}{r^2}\mathrm{v} + 3\frac{\partial p}{\partial t} + 2\frac{\dot\alpha}{\alpha}p = 0 \tag{E}$$

$$\frac{\dot\alpha}{\alpha} = -8\pi\alpha r\left(\rho_m \mathrm{v} + \frac{1}{4\pi r^2}\sqrt{\frac{\beta}{\alpha}}L(r)\right) \tag{F}$$

$$\int_O^R 4\pi r^2 (\rho_m + 3p)\, dr \approx \int_O^R 4\pi\sqrt{\alpha}\, r^2 \rho_m dr \tag{G}$$

CONCLUDING REMARKS

Without entering on the difficult problems posed by these equations I shall only add a few further remarks. There are two interesting limiting cases, of which the one is wellknown, namely the classical solution (valid also for a limited system) of an expanding system of pressure-free matter without radiation corresponding to zero total energy, namely :

$$v = \frac{r}{T}, \quad T = \frac{3}{2}t, \quad \rho_m = \frac{1}{6\pi t^2} \qquad (25),$$

which satisfies the condition :

$$\frac{8\pi}{3}\rho_m T^2 = 1 \qquad (26),$$

mentioned earlier. If a solution of equations (A) to (G) could be proved to exist, which starts with $v = 0$ and for large t-values goes over into the solution just mentioned, then we would have a good beginning for the further study of our problem. On the other hand, this solution, or rather its contracting counterpart, will probably be useful as a starting point for the study of the evolution of the system during the contraction phase.

Moreover, it would be interesting to know, whether the other limiting case, a stationary solution of radiation and matter (with an extra point source at the center to replace the disappearing radiation) of finite mass does exist. If so, it may be used as an auxiliary for the study of the intermediate phase. We know that no such solution exists for pure radiation, the gravitational field of which is not strong enough to keep the mass within finite bounds. In fact, at large distances from the center the density is inversely proportional to the square of the distance, β increasing as the square root of the distance, so that the outward Schwarzschild solution is never approached*. On the other hand, matter alone would condense up to the point where pressure begins to play a role. In a large cloud of thin matter the Schwarzschild limit may be approached before this stage is reached, in which case doubtlessly inner condensations would take place, the solution (25) being only valid as long as relativistic effects as well as pressure may be neglected.

Therefore, it seems not unreasonable to expect a finite mass solution for matter and radiation together. As a preparation to the investigation of this problem I have looked at the simplified case, where all non-classical terms are neglected exept the gravitational interaction between matter and radiation and of radiation with itself. For these simplified equations a stationary solution

* O. Klein, *Arkiv f. mat., astr. o. fys. Bd.*, **34 A**, No. 19, 1947.

does, indeed, exist. This solution is very curious, however, and contradicts the simplifying assumptions on which it is based. Thus, between $r = \frac{2}{3} r_o$ and r_o ($r_o = \frac{2}{3}$ L, L being the constant value of L (r)) there is both matter and radiation corresponding to the formulae :

$$\rho_m = 24p_1 \frac{(r_o - r)(r - \frac{2}{3} r_o)}{r^2}, \qquad p = 4p_1 \left(\frac{r_o - r}{r}\right)^2 \qquad (27),$$

p_1 being a constant uniquely determined by L (pressure at $r = \frac{2}{3} r_o$). But the sphere within r_o is completely free from matter and filled with radiation, the density of which decreases slowly form a value p_o at the center to the value p_1 at $r = \frac{2}{3} r_o$, there being a numerical relation between p_1 and p_o. The total mass of the system is equal to $r_o = \frac{2}{3}$ L, which means that the neglect of relativistic effects is far from being justified, the Schwarzschild singularity inequality being $M < \frac{r_o}{2}$. On the other hand, the solution has much similarity with the kind of system considered in the introductory elementary remarks about the Eddington relations and would yield these, when the only integration constant L is so chosen that the opacity condition is just narrowly fulfilled. I am not able to tell whether a somewhat similar solution of the rigorous equations does exist.

I wish to give my best thanks to Dr. Bertel Laurent for much helpful discussion and careful reading of the manuscript.

ADDENDUM

As to the solution of equations (A) to (F) we may, perhaps somewhat schematically, start from a state, say at $t = 0$, just in between the original contraction and the subsequent expansion, in which state $\mathrm{v} = 0$ and $\mathrm{L} = 0$. Further we shall tentatively assume that in this state the radiation energy is more strongly concentrated towards the centre $r = 0$ than corresponding to gravitational equilibrium. Moreover, we shall use a series expansion in terms of powers of r of the quantities in question hereby limiting ourselves to the lowest power appearing in the equations.

Thus, beginning with equation (D) we see from (A) that the term with the lowest power of r of $\dfrac{2p}{\beta}\dfrac{\partial\beta}{\partial r}$

$$\text{is } \frac{16\pi}{3}\, p_0\,(\rho_0 + 6p_0)\, r\,,$$

where ρ_0 and β_0 are the values at $r = 0$ of ρ_m and β respectively. Likewise the lowest order term of :

$$\frac{\partial p}{\partial r} \text{ is } -\,kr, \text{ where } k = -\,2\left(\frac{\partial p}{\partial r^2}\right)_0.$$

Thus, we see that the expansion of $\mathrm{L}(r)$ will begin with a third power term, which we shall denote by $4\pi\lambda r^3$, and we get :

$$\frac{1}{\sqrt{\beta_0}}\frac{\partial\lambda}{\partial t} + \frac{8\pi}{3}\rho\lambda - k + \frac{16\pi}{3}\, p\,(\rho + 6p) = 0 \qquad (1),$$

where β_0 is the value of β at $r = 0$ determined by (A) and the boundary condition $\beta \rightarrow 1$ for $r \rightarrow \infty$, and where we have omitted the index $_0$ of ρ_0 and p_0.

It follows now from (C) that the expansion of v starts with a linear term, which we write : $\sqrt{\beta_0}\,\eta\, r$, and we obtain :

$$\frac{1}{\sqrt{\beta_0}}\frac{\partial\eta}{\partial t} + \eta^2 = \frac{4\pi}{3}\,[2\lambda - (\rho + 6p)] \qquad (2)$$

From (B) we get, further, in the same way :

$$\frac{1}{\sqrt{\beta_0}}\frac{\partial\rho}{\partial t} + 3\eta\rho = 0 \qquad (3),$$

and from (E), using (F) similarly :

$$\frac{1}{\sqrt{\beta_0}} \frac{\partial p}{\partial t} + \lambda = 0 \tag{4}$$

Considering (E) and (F) we see that from the second order terms in r, which did not contribute to the derivation of (4), we may derive an equation for :

$$\frac{\partial k}{dt},$$

which will, however, as a new variable contain the quantity :

$$\left(\frac{1}{r} \frac{\partial \lambda}{\partial r}\right)_{r=0}$$

In taking further terms of the power series expansion into account we may derive an equation for the time derivative of this quantity and so on. In order to get a general idea of the time behaviour of the five quantities ρ, p, λ, η, k we shall, however, neglect the r-derivative of λ. Using the expansion following from (A) :

$$\sqrt{\frac{\beta}{\alpha}} = \sqrt{\beta_0}\,(1 - \frac{2\pi}{3}\,\rho r^2) + 0(r^4) \tag{5}$$

we get, thus :

$$\frac{1}{\sqrt{\beta_0}} \frac{\partial k}{\partial t} = \frac{16\pi}{9}\,\rho\eta\,(\lambda - 6p) + \frac{4\pi}{9}\,\rho\lambda \tag{6}$$

Introducing the independent variable :

$$\tau = \int_0^t \sqrt{\beta_0}\, dt \tag{7}$$

the proper time at $r = 0$, instead of t, we obtain finally the following five equations for the five quantities : $\rho, p, \lambda, \eta, k$ as functions of τ :

$$\begin{aligned}
\frac{d\lambda}{d\tau} &= k - \frac{16\pi}{3}\left(p\,(\rho + 6p) + \frac{1}{2}\rho\lambda\right)\\
\frac{d\eta}{d\tau} &= \frac{4\pi}{3}(2\lambda - \rho - 6p) - \eta^2\\
\frac{d\rho}{d\tau} &= -\,3\eta\rho\\
\frac{dp}{d\tau} &= -\,\lambda\\
\frac{dk}{d\tau} &= \frac{4\pi}{9}\,\rho\,[\lambda + 4\,(\lambda - 6p)\,\eta]
\end{aligned} \tag{8}$$

In the units defined above and used in these equations we should expect ρ to be of the order of magnitude 1 for τ equal to zero. An estimate of the corresponding value of p may be obtained from (G), whereby the classical approximation together with the assumption of a uniform distribution of p and ρ throughout the volume would give a rough estimate of the order of magnitude. In this way we obtain $p \approx 0.1\ \lambda\rho$ where λ is the fraction used in equation (2). If we call the initial value — assumed to be positive — of $k - \frac{16\pi}{3} p(\rho + 6p)$, γ, then for sufficiently small τ-values we have :

$$\lambda = \gamma\tau\,,\ \eta = \frac{4\pi}{3}\,\tau\,(\gamma\tau - \rho - 6p) \qquad (9).$$

We shall not at present attempt a closer discussion of the solution of (8) but it should be noted that according to recent estimates the present value of η (corresponding to $T \sim 13 \cdot 10^9$ years) is about 0.05 in the units used here.

Discussion of Klein's Report

Dr Oppenheimer. — Are there many universes? I believe that you think there are.

Dr Klein. — Yes, I think so.

Dr Oppenheimer. — Can one see out of ours? I think you hope that we shall be able to do it.

Dr Klein. — Yes, although it must be very difficult, of course.

Dr Bondi. — Do you mean that the red shifts do not go on increasing with distance?

Dr Klein. — If there is a limit to the system, I would guess that the Hubble law is an approximation which would gradually cease to be applicable at very large distances.

Dr Mac Crea. — The Schwarzschild singularity has been mentioned. I believe that physical considerations show that a singularity must exist. For, if we bring a test particle towards a massive particle, there must be a distance at which the gravitational energy given up is equal to the proper-mass of the test particle. Hence we cannot add mass at this critical distance. This is a property that cannot be transformed away. It applies, of course, only to a massive particle of *positive* mass.

Dr Fierz. — May I remark that Mach's Principle is much older than Mach? As I learn from Dr Gold, it is due to Bishop Berkeley, and later it was restated by L. Euler. Euler rejects it as absurd. The argument was always this : the centrifugal force which after Newton proves the existence of absolute space — Newton uses the example of the rotating bucket — could be due to the fixed stars, against which a relative rotation takes place.

The principle obviously is not fulfilled in general relativity, as there the metric g_{ik} is logically independent of matter, and so there is absolute rotation. On the other hand, as we know now space to be never really empty – the quantized fields cannot be taken out of space — the problem has changed. So it happens that everybody has a slightly different idea of what Mach's Principle means. We should no longer use this notion, as it only helps to produce confusion.

Dr Heckmann. — The Gödel solution shows that there exists absolute rotations in Einstein's theory. Mach's principle therefore is not incorporated in the theory right from the beginning but has to be stated separately.

Dr. Wheeler. — The principle of Mach and Einstein has in the past often been regarded as an idea to be either affirmed or denied. Perhaps this reaction arises because the principle has been often worded so definitely : « The distribution of matter and energy uniquely determines the metric ». It would seem more productive to regard the Einstein-Mach principle as a principle so to be formulated that it shall have a meaning and shall constitute an organic part of Einstein's general relativity. More specifically relativity constitutes a system of differential equations which is incomplete until it is supplemented by a boundary condition. It appears not unnatural to think of the Mach principle as a principle from which one may some day learn how to formulate boundary conditions that would make physical sense. Consider by way of analogy Laplace's equation (Tabel 1). This familiar differential equation has to be supplemented by a boundary condition, such as the asymptotic vanishing of the potential, in order to make the connection between charge and potential unique. The similar demand of Einstein's equations that they should give a unique connection between the stress energy tensor and the metric will of course require an entirely different kind of boundary condition because of the non-linear character of the metric. We are far from knowing enough today about the equations of general relativity to know what kind of boundary condition is needed to insure this unique connection. Is it enough to demand that the space be closed? We do not know.

It would seem out of place to take a final stand on the Mach principle until this mathematical question shall have been elucidated. However, two points seem relevant to such a future discussion.

TABLE 1.

Analog between electrostatics and general relativity as suggesting in what direction it may someday be possible to formulate Mach's principle so that it shall have a meaning which is consistent with, and an organic part of, general relativity, as giving a reasonable choice of boundary conditions.

Differential Equation	Mach principle or its electrical analogue	Boundary condition that results from this principle	Integral representation equivalent to differential equation plus boundary condition
$\nabla^2\Phi = -4\pi\rho$	Potential uniquely determined by charge	Φ asymptotically goes to zero	$\Phi = \int \frac{\rho dr}{r}$
$R_{ik} = T_{ik}$	Metric uniquely determined by distribution of mass-energy	Closed space?	Approximate representation of metric analogous to above only good at small distances (Einstein, Thirring)

First, the metric at any future time is determined by the metric on any initial space-like surface; in other words, by the initial value data. Now the recent work of Lichnerowicz and Faures has shown that this data is not freely disposable but is subject to important conditions which in mathematical terms have somewhat the character of a Poisson equation in a three-dimensional curved space. In other words, if the principle of Mach and Einstein is to have any meaning at all in the sense considered here it must be possible to formulate it and bring into evidence entirely within the mathematical framework of the three-dimensional initial space-like surface. Second, the Schwarzschild singularity has traditionally furnished an example of a metric which is *not* consistent with Mach's principle. A test particle indefinitely far from the source of attraction has its inertial properties completely determined. This is only to say that the inertial properties of the space, and the metric, are *not* in this instance uniquely determined by the distribution of mass. There is only one center of mass and it is obviously insufficient, because of its rotational symmetry, to fix the inertial properties of test particles — and therefore the properties of the metric — for motions that are not purely radial. The Einstein-Mach principle in the formulation suggested here would therefore *exclude* the simple Schwarzschild solution and say it can not represent a physically acceptable space. However, it would not exclude a closed space built up as in figure 6 from the union of a number of Schwarzschild spaces of limited extent. In this example there are obviously sufficiently many mass to determine a reasonable frame of reference and to be consistent with the view « that the distribution of mass and energy determines the metric ». This illustration gives one additional incentive to ask whether *closure of space* is the kind of boundary condition that might reformulate the physical language of Mach's principle into well defined mathematical terms.

The steady state theory

by **F. HOYLE**

Formalism.

The simplest way to introduce the equations of the steady state theory is through a generalization of well-known Newtonian theory. In the latter, the hydrodynamic equations of a fluid may be derived from divergence equations :

$$\frac{\partial T^{\mu\nu}}{\partial x^{\nu}} = 0, \quad \mu, \nu = 1, 2, 3, 4 \tag{1}$$

where x^1, x^2, x^3 are space coordinates, and $x^4 = ct$. When :

$$T^{\mu\nu} = \rho v^{\mu} v^{\nu}, \quad v^{\mu} = \frac{dx^{\mu}}{dt}, \quad \rho = \text{mass density}, \tag{2}$$

equations (1) are equivalent to the Eulerian equations of a freely moving fluid in the absence of internal stress and thermal effects. Stress terms may be included by adding a three dimensional tensor p^{ij} to the spatial part of $T^{\mu\nu}$ $(\mu, \nu \equiv i, j = 1, 2, 3)$, equations (1) being still retained. For an isotropic medium the stress tensor p^{ij} is expressible in terms of the spatial derivatives of the velocity components, viz. :

$$p^{ij} = \alpha\left(\frac{\partial v^i}{\partial x^j} + \frac{\partial v^j}{\partial x^i}\right) + \beta\delta^{ij}\frac{\partial v^k}{\partial x^k}, \tag{3}$$

where α, β are constants depending on the nature of the fluid, and δ^{ij} is the Kronecker symbol.

The effect of thermal pressure may likewise be included by adding $P\delta^{ij}$ to the spatial part of $T^{\mu\nu}$.

The equations of motion in relativistic cosmology are formally very similar to the above Newtonian scheme.

Equations (1) must be replaced by the appropriate Riemannian form :

$$T^{\mu\nu}{}_{,\nu} = 0, \tag{4}$$

and the velocity v^{μ} in (2) must be defined by :

$$v^{\mu} = \frac{dx^{\mu}}{ds}. \tag{5}$$

Once again ordinary stress and pressure effects can be included by appropriate additions to $T^{\mu\nu}$. These additions are, however, essentially three dimensional in character.

The steady state theory is obtained by adding to $T^{\mu\nu}$ a four dimensional generalization of (3).

$$T^{\mu\nu} = \rho v^{\mu} v^{\nu} + \alpha\,(v^{\mu\nu} + v^{\nu\mu}). \tag{6}$$

The term $v^{\mu\nu}$ is the contravariant derivative of v^{μ}, and α is a universal constant. A term involving a second constant β might also have been introduced into (6) in analogy with (3). This is not done here, because such a term is redundant, at any rate in the smooth cosmological approximation. That is to say, it does not lead to any effects that are not already given by the α term.

An objection to the form of (6).

It is reasonable to object to the new term in (6) on the ground that it has the undesirable property of depending on the local velocity of matter but not on the local density. The best answer to this objection lies not by way of introducing the local density, but in removing the dependence on the local velocity. This can be done by a procedure described briefly below in which a vector field u^{μ} is defined. The new term in $T^{\mu\nu}$ then takes the form $\alpha(u^{\mu\nu} + u^{\nu\mu})$ instead of that given in (6).

The procedure consists in first defining the mean motion of an observer relative to the universe. To do this, integrals must be taken over the observer's past light cone, the integrals being constructed from the momentum of matter in such a way that the contributions from different parts of the light cone are added together in a covariant manner. It is also important that nearby matter be not so heavily weighted that it contributes to the integrals more strongly than distant matter.

Now at any point of space-time choose an observer relative to whom the mean motion of the universe is zero. The vector u^{μ} is then taken as a vector of unit length tangent to the world line of this observer.

On any reasonable basis for defining u^{μ} it turns out that $u^{\mu} = v^{\mu}$ in the smooth cosmological approximation. Indeed we may employ the latter equation as an excellent approximation for u^{μ} even when account is taken of deviations from the smooth approximation. The effect of condensation into galaxies would cause small terms to appear in $T^{\mu\nu}$ when the v^{μ} field was used that would not appear with the u^{μ} field. These terms would be smaller than the main terms in $T^{\mu\nu}$ by factors of the order of the ratios of the random velocities of the galaxies to the velocity of light. Such ratios are of order 10^{-3}. Hence the difference between the use of v^{μ} and u^{μ} would be so small that the extra complexity required in the definition of u^{μ} hardly seems worthwhile from a practical point of view, although it may be highly desirable from a logical standpoint. The use of $u^{\mu} = v^{\mu}$ is an approximation of very much the same character as the Weyl postulate.

The dynamical equations of the steady state theory.

The equation of the theory now take the usual form :

$$R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R = - \varkappa T^{\mu\nu}, \ \varkappa = 8 \pi G/c^4, \qquad (7)$$

where $R^{\mu\nu}$ is the Ricci tensor, and G is the gravitational constant. For an isotropic homogeneous large scale distribution of matter the Robertson-Walker form of the line-element may be used, namely :

$$ds^2 = c^2 dt^2 - S^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 \left(d\theta^2 + \sin^2\theta \, d\varphi^2\right)\right] \qquad (8)$$

With this line-element, and after some reduction, equations (7) give :

$$\varkappa \rho c^4 S^2 = 3 (\varkappa c^2 + \dot{S}^2), \qquad (9)$$

$$2 \alpha c^2 \varkappa S\dot{S} = 2 S\ddot{S} + (kc^2 + S^2), \qquad (10)$$

as differential equations that must be satisfied by the mass density ρ and by the expansion function $S(t)$.

Equation (10) is much simplified by the following substitutions :

$$\begin{aligned} 3\,H &= 2\,\alpha\varkappa c^2 = 8\,\pi\alpha G/c^2, \text{ using the definition of } \varkappa \text{ in (7)}, \\ \tau &= \frac{1}{H}\frac{\dot{S}}{S}, \\ k' &= \frac{kc^2}{H^2}, \\ \xi &= \log S, \end{aligned} \tag{11}$$

which lead to :

$$\frac{d\tau^2}{d\xi} = -\,3\,\tau^2 + 3\,\tau - k'e^{-2\xi}. \tag{12}$$

In an continually expanding universe ξ increases with time t, and the term in k' eventually becomes unimportant. Neglecting this term we have :

$$\frac{d\tau^2}{d\xi} = 3\,\tau\,(1 - \tau), \tag{13}$$

which shows that τ tends asymptotically to unity. From (11) we see that the function $S(t)$ tends asymptotically to $\exp\,(Ht)$. With $k = 0$ this gives the de Sitter line-element, viz. :

$$ds^2 = c^2dt^2 - \exp\,(2\,Ht)\,[dr^2 + r^2\,(d\theta^2 + \sin^2\theta\,d\varphi^2)]. \tag{14}$$

The density of matter is not zero, as it is in the usual discussion of the de Sitter line-element, however, for (9) with $\dot{S}/S = H$, $k = 0$, gives :

$$\rho = 3\,H^2/8\,\pi\,G. \tag{15}$$

This density is maintained by the creation of matter. Thus differentiating (9) with respect to t we have ($k = 0$).

$$\varkappa\,c^4\,\frac{\partial\rho}{\partial t} = 6\,\frac{\dot{S}}{S}\left(\frac{\ddot{S}}{S} - \frac{\dot{S}^2}{S^2}\right),$$

which together with (10) (again with $k = 0$) gives :

$$\frac{\partial\rho}{\partial t} = \rho\,\left(2\,\alpha\varkappa c^2 - \frac{3\,\dot{S}}{S}\right) = 3\,\rho\,\left(H - \frac{\dot{S}}{S}\right). \tag{16}$$

The first term in the brackets on the right hand side of (16) represents the creation of matter, and the second term represents the effect of universal expansion, the two terms just coming into balance as $S \rightarrow \exp\,(Ht)$.

A remark on creation.

Since the above equations refer to a smooth fluid it is clear that the details of creation cannot be discussed within the terms of reference of such a classical model. Yet the inherent plausibility of the creation of matter can be demonstrated by a very simple argument.

It is well known that there exists a cut-off distance when the line element is of the de Sitter form. Defining distance at time t by $r \exp(\mathrm{H}t)$, no signal emitted at time t at a distance greater than c/H from an observer can ever be received by the observer, no matter how long he waits for the signal to arrive, and no matter how sensitive his detection equipment. Signals emitted at distances less than c/H can be so observed, however.

The mass lying within this « observable horizon » is :

$$\frac{4\,\pi\rho}{3}\left(\frac{c}{\mathrm{H}}\right)^3 = \frac{c^3}{2\,\mathrm{GH}}, \text{ using (15) for } \rho. \qquad (17)$$

Multiplying (17) by G and dividing by c/H gives $\frac{1}{2}c^2$. In a crude Newtonian sense this result may be interpreted by saying that every particle in the universe exists in a gravitational potential well whose depth in such that the gravitational energy of the particle is comparable with its rest mass (the factor 1/2 in $1/2c^2$ is obviously not of significance here since the present calculation clearly cannot be interpreted in an exact fashion). Continuing the argument, the process of creation can accordingly be thought of as involving no energy expenditure — a particle is created at a negative potential that compensates for its rest mass. Accordingly to quantum theory, particle creation might well be expected under these circumstances.

A remark on the natural constants.

It is possible to obtain the current value of $\dot{\mathrm{S}}/\mathrm{S}$ by observation. The best estimate available, due to Sandage, gives :

$$\frac{\dot{\mathrm{S}}}{\mathrm{S}} \simeq (4 \times 10^{17}\ \text{sec.})^{-1} = \mathrm{H}. \qquad (18)$$

In the steady state theory the same value of $\dot{\mathrm{S}}/\mathrm{S}$ would be obtained at any other epoch. This is not the case in relativistic cosmology, for in relativistic cosmology it is impossible to build a natural unit of time out of the constants that appear in the theory, namely c and G. The situation is otherwise in the steady-state theory because of the creation constant α introduced into expression (6) for $T^{\mu\nu}$. Equipped with α, c, G the theory yields a natural unit of time H^{-1}, a natural unit of length cH^{-1} — the « cut-off » radius, and a unit of mass. The latter shows itself in the constant universal density given by (15).

If it were not for the existence of certain well-known numerical agreements between dimensionless quantities of order 10^{39} that can be constructed from a combination of cosmological theory and microscopic physics, it would largely be a matter of taste which of these two points of view is to be preferred. These agreements raise doubt, however, as to whether relativistic cosmology can be accepted in conjunction with the usual assumption that the constants of physics are unchanging with time. For unless the agreements are sheer coincidence we must surely conclude that if the current value of $\dot{S}/S$ is ephemeral ($\dot{S}/S$ appears in constructing certain of the dimensionless quantities) then so must be the basic constants of microscopic physics — a view that indeed was urged long ago by Dirac.

The alternative lies in the steady state theory which gives constant values to the cosmological quantities. A sophisticated theory of the creation process would evidently seek to connect the constant α with the constants of microscopic physics. If this could be done the agreements in the quantities of order 10^{39} would no longer be attributable to coincidence, but would appear as an expression of the theory of the creation process itself.

It may be added that some restriction can be placed on the extent to which the natural constants of microscopic physics can have changed during the past. Thus a rather complicated quantity built out of the gravitational constant and the constants of microscopic physics appears to approximately the seventh power in the expression for the luminosities of main sequence stars. In view of this sensitive dependence it would not be possible to change the quantity in question by any appreciable factor without running into serious conflict with the astrophysics of stellar evolution.

Other restrictions arise in connection with the study of geophysics and of meteorites, particularly concerning the half lives of naturally radioactive elements. If the natural constants have varied appreciably over the last five billion years, it is necessary for instance that they should have varied in such a way as to prevent serious discrepancies between widely different methods of estimating the ages of meteorites, e.g. the strontium-rubidium method depending on β-decay, and the uranium-lead method depending on α-decay.

Irregularities.

To date it has not been possible to observe intergalactic matter. Observation in cosmology therefore is concerned with irregularities, of which the galaxies are the most important. Much of the remainder of this report will accordingly be concerned with the formation of galaxies.

The problem of irregularities takes a somewhat curious form in the steady state theory. Since there is no finite origin of time it is not necessary to provide for a finite origin of any particular type of irregularity. What is necessary, however, is that any observable irregularity must be reproducible from generation to generation — a generation being defined by $(3H)^{-1}$. An irregularity in its nature occupies only a finite volume. Consequently the expansion of the universe tends to separate irregularities, thereby tending to reduce their mean density of occurrence in space. Thus if N is the mean density of some particular type of irregularity, expansion gives a contribution $-3N/H$ to dN/dt. This must be compensated by a term $+3N/H$ representing the mean reproduction rate of the irregularity in question. Unless this compensation occurs, the spatial density will either fall steadily to zero (in which case no such irregularity should be observed), or the spatial density will increase until some form of balance does set in.

This may be summarized by saying that the irregularities we now observe must be just those for which there is an exact compensation between reproduction and expansion.

Magnetic fields.

Essentially the same considerations apply to magnetic fields. Evidence concerning cosmic rays and radio-astronomy suggests that a magnetic field probably exists not only within the galaxies but also on a universal scale. Expansion must tend to weaken such a field. Consequently if a universal field is to be maintained in the steady state theory some dynamo process is required to offset the effect of expansion. It is of interest to consider the following very simple form of dynamo process.

Suppose that the material in a volume V condenses to form N galaxies (how this may happen is discussed later). After condensation, each line of force of any pre-existing, approximately uniform, magnetic field that crossed V will pass on the average through $\sim N^{1/3}$ of the galaxies. Suppose further that the galaxies are initially condensed into a comparatively tight cluster with volume small compared to V, but that the cluster ultimately disperses back again into V with the galaxies taking up more or less random positions (how this may occur is also discussed later). The effect on the magnetic field is to produce a « turbulence » in which the length of each line of force is increased by $\sim N^{1/3}$, thereby increasing the field intensity also by $\sim N^{1/3}$. Thus provided N is a number of appreciable magnitude, say $\sim 10^2$, the increase in magnetic intensity is adequate to offset the dispersive effect of expansion over as time $\sim H^{-1}$.

By itself, this process would of course increase the local complexity of the field. Expansion has the opposite effect, however, so that a balance in which the intensity of the field was just maintained could also be a balance in which the local complexity of the field was just maintained.

It is of interest to compare the above remarks with the status of magnetic fields in relativistic cosmology. In the latter it must be argued that the magnetic field was a feature of the universe at its origin unless it can be shown that the magnetic field has been built up by electromagnetic processes that have occurred during the lifetime of the universe, i.e. during a lifetime of order 4×10^{17} sec.

The difficulty of providing for an electromagnetic origin of the magnetic field can be appreciated from the following very simple considerations. Let a battery be inserted in a thick wire of small resistance bent into the form of a circle of radius a. The self-inductance $\sim 4\pi a/c^2$, while the current required to maintain a field of magnitude H is of order $ac\mathrm{H}/4\pi$.

If this field is built in time T the back voltage due to the growing field is :

$$\sim \frac{Hac}{4\pi T}\frac{4\pi a}{c^2} = \frac{a^2H}{cT}$$

Setting this equal to the EMF of the battery, we obtain an estimate for the driving field that is necessary to establish H in time T. Reasonable values are $H \sim 10^{-7}$ gauss, $a \sim 10^{23}$ cm, $T \sim 4 \times 10^{17}$ sec. The required EMF $\sim 10^{11}$ c.g.s. units, or about 3×10^{13} volts. So far no plausible suggestion has been made as to the origin of such an enormous impressed voltage. It is emphasised that this driving field is not one that can be removed by a Lorentz transformation, but must genuinely be « seen » by the intergalactic material. Because of the very low resistance of ionized hydrogen only a small electric field is required to maintain a magnetic field once established, but the voltage required to build the field is not small.

Yet unless an explicit origin for the field can be established, relativistic cosmology leads to the somewhat unpalatable conclusion that in the early history of the universe the dominating form of energy was magnetic. The magnetic energy density changes proportionately to S^{-4}, whereas the energy density of matter changes proportionately to S^{-3}. It follows that, since $S \to 0$ at the origin of the universe, the magnetic energy density supplies the dominant divergence at the zero of time.

Preliminary remarks on the formation of galaxies.

The formation of galaxies presents a curious problem, for the universe combines both expansion and condensation. This apparent contradiction is overcome in Lemaître's cosmology by arranging for the formation of galaxies to have occurred at an epoch when the universe was quasi-stationary. No such provision is made in other forms of relativistic cosmology, the origin of the galaxies being by-passed with the rather vague hypothesis that islands of higher density were present within the expanding cosmological material. At a certain stage these islands are supposed to have resisted the general expansion and to have condensed into stars. How and why this condensation took place is left in an equally vague condition.

The question now arises as to how the condensation of galaxies can be understood in the steady state theory. As an approach to the problem it is of interest to note that in the steady state theory the intergalactic matter may be at a very high kinetic temperature (in relativistic cosmology, on the other hand, adiabatic expansion from an initial very high density lowers both the radiation and the kinetic temperatures to small values). For example, if the created particles are taken to be neutrons, the thermal kinetic energy arising from neutron decay is of order $2 \cdot 5 \times 10^{17}$ ergs per gm., which is sufficient to raise the kinetic temperature to a value of order 10^9 dg. K.

The thermal energy density, W say, may be estimated in the following way, when the energy is derived from neutron decay. Write :

$$\frac{dW}{dt} = 3\,\rho H \varepsilon - 3\,WH - 2\,WH, \tag{19}$$

where ρ is the steady state cosmological density given by (15), and ε is the thermal energy supplied per unit mass of neutrons (the energy taken by neutrinos being regarded as lost). The first term on the right of (19) is the rate of creation of thermal energy, the second takes account of the decline of the energy density due to geometrical expansion, and the third term arises because each particle loses random energy in the expansion, the random velocity falling proportionately to S^{-1} (the factor 2 in this term being present only when the particle energies are non-relativistic).

In a steady-state $dW/dt = 0$, so that (19) gives :

$$W = \frac{3}{5}\rho\varepsilon = \frac{9}{40\,\pi}\,\frac{H^2\,\varepsilon}{G}, \text{ using (15) for } \rho. \tag{20}$$

With (18) for H, $\varepsilon \simeq 2 \cdot 5 \times 10^{17}$ ergs per gm., and $G = 6{\cdot}67 \times 10^{-8}$, we obtain :

$$W \simeq 1 \cdot 7 \times 10^{-12} \text{ ergs per cm}^3. \tag{21}$$

for the thermal energy density.

Condensation will take place in the hot gas if local cooling occurs in a sufficiently small volume, the cooled gas being then compressed by surrounding hotter gas. The necessary condition on the volume of the cooled region is that the differential velocity of expansion of the universe taken across it shall not exceed the velocity of sound in the outer, hot, compressing gas. An estimate can be obtained from this condition of the largest irregularities that can arise from cooling.

The best observational estimate for the differential expansion velocity is 80 km per sec. per megaparsec. For an irregularity of diameter D mpc. the differential velocity is therefore $\sim 80\,\mathrm{D}$ km. per sec. If this to be less than the speed of sound in hydrogen at temperature $\sim 10^9$ dg.K. we have $\mathrm{D} < \sim 50$ mpc. This estimate refers of course to the irregularity that develops in the hot cosmological material, not to the size of the cool condensation, which after compression may well be an order of magnitude smaller, say 10 mpc.

Preliminary remarks on the cooling problem.

Quite a strong case can be made to show that the galaxies were indeed formed by radiative cooling in a hot gas at temperature $\sim 10^7$ dg.K., and density $\sim 10^{-27}$ gm. per cm^3. As will be seen later, this radiative cooling cannot be the primary cooling process referred to above. Nevertheless, it is of interest to consider this question first, before coming to the more difficult problem of primary cooling, since the results obtained add a considerable measure of plausibility to the general concept that pressure gradients play an important role in the condensation process.

At the low densities under consideration in cosmology a hot gas is essentially completely transparent to its own radiation. The rate of radiation by ionized hydrogen at temperature T and particle density n atoms per cm^3 is :

$$8 \cdot 7 \times 10^{-4}\,\mathrm{T}^{1/2}\left(1 + \frac{3 \cdot 85 \times 10^5}{\mathrm{T}}\right) n \text{ ergs per gm. per sec.,} \quad (22)$$

the first term in brackets being due to Bremmstrahlung and the second to recombination. The thermal energy possessed by the hydrogen is :

$$\sim (1 \cdot 3 \times 10^{13} + 2 \cdot 5 \times 10^8\,\mathrm{T}) \text{ ergs per gm.,} \quad (23)$$

where the first term is the energy of ionization and the second term is thermal. For the hydrogen to be able to radiate an appreciable fraction of its energy in a time $\sim \mathrm{H}^{-1}$ we accordingly require that the product of H^{-1} and (22) shall be comparable with (23). Using (18) for H^{-1} and neglecting terms that are small at high temperature this requirement is expressed by :

$$n \cong 7 \times 10^{-7} T^{1/2}. \tag{24}$$

It is emphasised that we are not concerned here with primary cooling in material at density given by (15) (using the value of H in (18), (15) gives $\rho \cong 10^{-29}$ gm. per cm^3), but with secondary cooling in material of higher density. It will be supposed that an approximate equality of pressure is maintained at all stages of cooling, however, so that the pressure is determined by that which exists within the very hot primary material. The latter can immediately be estimated from (21). Thus we require a pressure everywhere of order 10^{-12} dynes per cm^3. This requires :

$$nT \cong 5 \times 10^3. \tag{25}$$

Combining (24) and (25) yields $n \cong 10^{-3}$ atoms per cm^3, and $T \cong 10^7$ dg.K.

The implication of the present argument is that radiative cooling on a cosmological time scale is important at densities $> \sim 10^{-27}$ gm. per cm^3, but not at significantly lower densities.

At still higher densities radiation rapidly becomes a dominating process. Always using pressure equilibrium, as expressed in (25), the above formulae show that the time scale for radiative cooling declines closely as $n^{-3/2}$ when $T > \sim 3 \times 10^5$ dg.K., and as $n^{-5/2}$ for still lower temperatures. Remembering that n increases as the inverse cube of the radius R of a condensation, it follows that the time scale for cooling declines as $R^{4.5}$ for $T > \sim 3 \times 10^5$, and as $R^{7.5}$ for still lower temperatures. Viewed on a cosmological time scale, cooling rapidly becomes catastrophic as R decreases.

This catastrophic cooling is precisely what is required to ensure that an initially slowly cooling cloud will condense into blobs rather than as a whole.

Cooling ceases when T declines below 10^4 dg.K. for the reason that the hydrogen no longer remains ionized. Thus cooling ceases (and hence condensation) when $n \cong 1$ atom per cm^3, as can be seen from (25). Since this is just the mean density found in the galaxies, it is natural to associate the condensed blobs with the galaxies. It appears remarkable that the densities within the galaxies can be related in such a simple fashion to the creation rate and to the energy of neutron decay (the latter quantities determining the right hand side of (25)).

Other remarkable results can also be derived by equally elementary arguments. The picture is of a cooled blob forming within a gas of temperature $\sim 10^7$ dg.K. The speed of sound in the gas ~ 500 km per sec. $=$ V say. Evidently no appreciable compression of a condensation of radius R can take place in a time less than $\sim$ R/V. Since the time of cooling declines with R much more rapidly than this, the question evidently arises as to whether cooling might be so rapid that compression cannot keep step with it.

For a condensation of density 10^{-27} gm. per cm^3, and mass of galactic order, R $\cong 3 \cdot 10^{23}$ cm, so that R/V $\cong 10^{16}$ sec. This refers to the situation before the cooling begins to accelerate, namely to the stage at which the time scale for cooling is of order $H^{-1} \cong 4 \times 10^{17}$ sec. Evidently then, compression has no difficulty to begin with in keeping step with cooling. But the situation is otherwise when R has decreased to about a third of its initial value. Marked pressure differences must develop between the interior of the condensation and the hot surrounding gas, for the surrounding gas cannot then compress the condensation fast enough to maintain approximate pressure equality.

Pressure equality can ultimately be restored, however, when n rises to ~ 1 atom per cm^3, since T cannot decline appreciably below 10^4 dg.K., so that (25) is again satisfied. But during the rise of density to 1 atom per cm^3 the condensation is driven inwards by gross pressure differences. Appreciable dynamical motions must consequently develop. These dynamical motions will have a speed of the same order as, but less than, V. The speeds may be set at $\sim \frac{1}{2}$ V, the factor $\frac{1}{2}$ being chosen somewhat arbitrarily.

Two possibilities now arise. The mass of the condensed blob may be insufficient for its gravitational field to control the dynamical motions, in which case we may expect the blob to fly into a number of smaller pieces. Or the opposite situation may occur. It is evidently of interest to determine the minimum mass required in this second case.

With $\frac{1}{2}V \cong 250$ km per sec., the dynamical energy amounts to $\sim 3 \times 10^{14}$ ergs per gm. Accordingly, we require the gravitational potential energy, GM/R, where R, M are the radius and mass respectively, to be order 3×10^{14} ergs per gm. Assuming that the distribution of mass within the blob is approximately uniform, we can also write :

$$M = \frac{4}{3} \pi R^3 m_H n, \tag{26}$$

where m_H is the mass of the hydrogen atom. Using (26), and taking $n \cong 1$ atom per cm³, at the end of condensation, we obtain $M \cong 10^{11}{}_{\odot}$, in very good agreement with the masses of the larger cass of galaxy.

Not only this, but it is to be expected that in the epoch of pressure disequilibrium rotary motion and random motion (i.e. motion relative to the cosmological background) will be developed. It again appears significant that the larger class of galaxy is observed to possess average random motions and motions of rotation that are both remarkably close to $\frac{1}{2}V = 250$ km per sec.

A difficulty in relativistic cosmology.

It seems scarcely possible that galaxies could condense at all if their random motions were large compared with the motions that could be controlled by their gravitation. Yet this condition is grossly flouted in many forms of relativistic cosmology.

Random motions decline with time in an expanding universe proportionately to S^{-1}. In relativistic cosmology the density declines proportionately to S^{-3}, so that :

$$\frac{u_1}{u_2} = \left(\frac{\rho_1}{\rho_2}\right)^{1/3}, \tag{27}$$

where ρ_1, u_1 ; ρ_2, u_2 are the mean cosmological density, and the mean random velocity of the galaxies, at times t_1 and t_2 respectively.

Now the argument that the galaxies are islands of high density that managed to avoid the general expansion implies that the galaxies formed when the universal density was of order 10^{-24} gm. per cm³.

Let time t_1 refer to this epoch, and let t_2 refer to the present. Then with $\rho_1 \cong 10^{-24}$ gm. per cm^3, $\rho_2 \cong 10^{-29}$ gm. per cm^3, and $u_2 \cong 250$ km per sec., we obtain :

$$u_1 \cong 10{,}000 \text{ km per sec.},$$

a wholly implausible result.

The fact that the random motion of a galaxy is usually comparable with the velocity of escape from its periphery seems a significant datum, explicable in the steady state theory in the terms described above, but not in relativistic cosmology, not unless it be admitted that the galaxies condensed at a time when the mean density was not significantly different from the present value. And this supposition leads immediately to severe time-scale difficulties.

The mean spatial density of galactic material.

In the steady state theory, material of age T has a lower average density than the universal value given by (15), the extent of the reduction being determined by a factor exp(— 3HT). This means that if we allow a time H^{-1} for the radiative cooling described above to take place, the mean density of the galaxies (i.e. the density given by imagining their material to be spread uniformly throughout space) should be less than the universal density by at least the factor e^{-3}. If we also allow an additional time H^{-1} for the operation of a primary cooling process (still to be discussed) the reduction is by a total factor e^{-6}. Hence on this basis the mean density of galactic material should be about 1 per cent of the density given by (15); i.e. about 10^{-31} gm. per cm^3, in excellent agreement with observation.

Primary cooling.

(The considerations of the remainder of the present report were arrived at in collaboration with Prof. T. Gold.)

In earlier sections radiative cooling by a gas of density $\sim 10^{-27}$ gm. per cm^3, and temperature $\sim 10^7$ dg.K. was considered. Such a gas has the same internal pressure as the general cosmological substratum, but has considerably lower temperature and higher density.

The question now arises as to how the cosmological gas with temperature $\sim 10^9$ dg.K., and density $\sim 10^{-29}$ gm. per cm^3 might be cooled to a significant degree.

Reference to the radiation formula (22) shows that at this low density less than 1 per cent of the thermal energy can be radiated directly in a time of order H^{-1}. The question therefore arises as to whether the energy can be lost through conduction into higher density gas at lower temperature. Galaxies are very effective in this respect, on account of the comparatively high value of n and the low value of T. Thus material with $n \simeq 1$ atom per cm^3, and $T \simeq 10^4$ dg.K., is able to radiate at about 3 ergs per gm. per sec. One gram of galactic gas can therefore dispose in a time H^{-1} of about 10^{18} ergs. For comparison, the heat content of the cosmological gas $\sim 2 \cdot 5 \times 10^{17}$ ergs per gm. It follows therefore that each gram of galactic gas is in principle able to dispose of the energy contained in two or three grams of cosmological gas.

Thus if the total mass of high density galactic gas were comparable with the mass of low density cosmological gas it might well be possible to cool the latter by a thermal energy transfer to the cool galactic gas. This is not so, however. The mean density of the galactic material, as indicated in the previous section, is no more than about 1 per cent of the mean density of the hot gas.

A 1 per cent energy loss by radiation may nevertheless be sufficient to give an adequate *indirect* form of cooling. Since we have to do with cool gas embedded in hot gas it is not at all implausible to suppose that some form of heat engine operates between the two. If so, the efficiency may be very high on account of the large temperature ratio between the hot and cold regions —10^5 for the cool matter of the galaxies and 10^2 for cooled matter at temperature $\sim 10^7$ dg.K. Hence an efficiency greater than $0 \cdot 99$ is theoretically possible in such an engine. This would mean that less than 1 per cent of the energy of the hot cosmological gas need be dissipated by radiation. The remaining 99 per cent may well be converted into dynamical motion.

At this stage it is necessary to remember the presence of a magnetic field, for we now have a possible dynamical source for electromagnetic processes that is vastly more powerful than any process in which galaxies are involved. On empirical grounds such a source is urgently needed, since attempts to explain the acceleration of cosmic rays and the origin of radio emission in terms of processes occurring within the galaxies have almost invariably run into serious limitations of energy. This was true of the acceleration of cosmic rays even when it was supposed that cosmic rays might be local to the galaxies. Now that the failure to find an upper limit to the cosmic ray spectrum suggests that the cosmic ray energy density may well be maintained throughout intergalactic space no process internal to the galaxies could possibly meet the necessary energy requirement.

Yet the present theory supplies exactly the right energy density for the cosmic rays. Thus if the cosmical gas is cooled by its energy first being converted to dynamical energy (heat engine) and then into cosmic ray energy, perhaps by a Fermi process, the ultimate energy density of the cosmic rays must be comparable with the value of W given by (21); i.e. must be of order 10^{-12} ergs per cm^3. This is precisely the case.

The energy spectrum of the cosmic rays.

Suppose that each cosmic ray particle is accelerated by a Fermi process in accordance with the equation :

$$\frac{dE}{dt} = \gamma E, \tag{28}$$

where E is the particle energy, and γ is a mean constant determined by averaging over many magnetic collisions. Let Θ be the energy density of cosmic rays. Then, neglecting for the moment the injection of new cosmic rays, we have :

$$\frac{d\Theta}{dt} = \gamma\Theta - 3\,H\Theta, \tag{29}$$

the first term on the right representing the rate of gain of energy, and the second arising from the geometrical effect of expansion. No term is required in (29) to take account of the effect of expansion on the velocities of the particles (red-shift effect), since this term is regarded as being absorbed into the constant γ in (28).

Now in a steady state $\frac{d\Theta}{dt} = 0$, so that $\gamma = 3H$, and each particle gains energy proportionately to $\exp(3Ht)$. On the other hand the particle density declines proportionately to $\exp(-3Ht)$. Thus by a slight variant of Fermi's well known argument it follows that the number of particles between E and $E + dE$ is proportional to dE/E^2.

Let $N(E, t)$ be the number of particles per unit volume with energy greater than E at time t. Then, if injection is neglected :

$$N(E + \gamma E\delta t, t + \delta t) = (1 - 3\delta t \cdot H)\, N(E, t).$$

In a steady state N is independent of t so that :

$$\gamma E \frac{dN}{dE} = -3\,HN ; N \sim E^{-3\,H/\gamma}.$$

The differential spectrum is accordingly $dE/E^{(1 + 3H/\gamma)}$. It has already been pointed out by Farley (*Varenna Conference on Cosmic Rays*, 1957) that the combination of a Fermi gain mechanism and of the loss rate due to the expansion of the universe leads to a power law spectrum. Farley was in some difficulty, however, to secure an adequate rate of gain, since he worked in terms of acceleration within the galaxies, instead of in terms of the far more powerful source considered above. Although explicit use was made of the formulae of the steady state theory, it was implied that similar considerations could be given for other forms of cosmology. This appears to be an oversight, for it is essential to have a volume expansion that is exponential with respect to time, and this is given only by the steady state theory.

It is not at all surprising that this argument yields too flat a spectrum since it ignores the energy that is fed into Θ by injection. This must appear as a positive term on the right of (29), thereby reducing γ (required to give $d\Theta/dt = 0$), and hence steepening the spectrum.

Suppose now that the additional term on the right of (29) is equal to $\gamma\Theta$, so that the modified equation takes the form :

$$\frac{d\Theta}{dt} = 2\,\gamma\Theta - 3\,H\Theta, \tag{30}$$

implying that half the increment to Θ in a small time interval comes from newly injected cosmic rays and half from the acceleration of old cosmic rays. Then for $d\Theta/dt = 0$, we now require $\gamma = 1 \cdot 5$ H. This leads, again by the slight variant of Fermi's argument, to the differential spectrum dE/E^3, in excellent agreement with the observed form at high energy.

If tentatively we regard the newly injected particles as having energies less than 10 Bev, a cosmic ray particle would require some 50 generations of acceleration to achieve an energy of 10^{20} ev. Very high energy particles are relics of much earlier epochs of the universe. There should be no upper limit to the cosmic ray spectrum.

Injection of cosmic rays.

In an earlier section it was seen that a dynamo action on an intergalactic magnetic field could be achieved through the formation and disruption of condensations. The way in which disruption might take place was passed over without discussion, however. The possibility now arises that clusters are disrupted through the increase of cosmic ray pressure within them. If escape of cosmic rays from the cluster is sufficiently slow, and if the external hot gas continues to supply mechanical work which is ultimately converted to cosmic ray energy, the necessary pressure will eventually be achieved.

Before expansion takes place the internal cosmic rays are subject to no loss. With a Fermi gain mechanism, and with particles injected at the bottom (possibly from stars) at a constant rate, the form of the spectrum is dE/E. Thus the main energy in this case lies at the top of the spectrum, E_0 say. The value of E_0 is determined by the gain rate and by the length of time for which the first injected particles are able to gain energy before the pressure becomes adequate to expand the cluster.

It is attractive to suggest that $E_0 \sim 5$ Bev, and that new cosmic rays come to be injected into the general intergalactic distribution in this way. The intergalactic distribution is then made up of two parts, a lower distribution of new cosmic rays with spectrum dE/E and an upper distribution of old cosmic rays with spectrum dE/E^3. With $E_0 \sim 5$ Bev as the approximate energy of separation of the two distributions, this suggestion is in satisfactory accord with the facts. It is noteworthy that there is no divergence of the integrated spectrum at its base (say at $0 \cdot 1$ Bev).

One further point deserves comment. Since the gain process is regarded as taking place at densities much lower than galactic densities, a serious obstacle to the acceleration of heavy nuclei is removed. The proportion of heavy nuclei and of protons will depend only on the injection proportions at the base of the spectrum.

Predictions of the theory.

Predictions can be classified under three headings, geometrical, astrophysical, and electromagnetic. The geometrical predictions are derived from the line element, taken together with the assumption that the irregularities are of the same kind in all parts of space-time. They are typified by the apparent magnitude — red shift relation for galaxies. To this may be added an angular diameter — red shift relation for objects of standard size, and also a number — apparent magnitude relation for galaxies and for cosmic radio sources.

The equations for these relations are well-known and will not be given here.

The astrophysical predictions of the theory are very different from those of relativistic cosmology. Galaxies are being formed all the time, so they should not be of uniform age. The material of the universe has not experienced a phase of primaeval high density. Hence any processes that require high density must take place locally within the stars. This is the case for the synthesis of elements from hydrogen. The mean density in space $\sim 10^{-29}$ gm. per cm^3. The material consists of very hot ionized hydrogen and hence should not be detectable by a 21 cm radio-wave technique.

Hydrogen radiation at temperature 10^7 dg.K. is in the form of quanta with energies of about 1 kev. A rough estimate can be made of the flux of these quanta at a typical point of space, and amounts to $\sim 10^3$ per cm^2 per sec. Unfortunately such quanta are strongly absorbed by the K electrons of oxygen, nitrogen, and carbon present in the interstellar gas, and hence cannot be directly observed from within our own galaxy. Secondary emissions from these atoms might possibly be observed in other galaxies, however.

Tentative electromagnetic predictions of the theory have already been largely described in the foregoing — that cosmic rays are a primary phenomenon, that there is no upper limit to the cosmic ray spectrum. In some cases a magnetic field may be self-contained within a galaxy or group of galaxies, but in general it is to be expected that the lines of force will not be closed within a cluster of galaxies.

Final remarks.

The status of the steady state theory can be improved in two ways, one by obtaining a better understanding of the creation process, the other by the successful comparison of prediction and observation. Of these, the latter appears currently the most profitable, since attempts to improve the basic equations of the theory (e.g. the introduction of the u^{μ} field, mentioned at the outset) will remain largely arid so long as the theory continues to be entirely classical in form.

BIBLIOGRAPHY

H. Bondi and T. Gold, *M.N.R.A.S.*, **108**, 252, 1948.
H. Bondi, *Cosmology*, Cambridge University Press, 1952.
F. Hoyle, *M.N.R.A.S.*, **108**, 372, 1948.
F. Hoyle, *M.N.R.A.S.*, **109**, 365, 1949.
W.H. McCrea, *Proc. Roy. Soc. A*, **206**, 562, 1951.
W.H. McCrea, *Proc. Roy. Soc. A*, **235**, 11, 1956.
F.A.E. Pirani, *Proc. Roy. Soc. A*, **228**, 455, 1955.
D.W. Sciama, *M.N.R.A.S.*, **155**, 15, 1955.

It will be found that the accounts given by these authors often differ quite substantially in emphasis from the present report. In particular, the reader is referred to Sciama's paper for a quite different and independent account of the formation of galaxies.

Discussion of Hoyle's Report

Dr. Møller. — I would like to ask Professor Hoyle a question regarding the physical meaning of the expression (6) for the matter tensor $T^{\mu\nu}$. Surely we do not need to worry about energy conservation, in the form of the steady state theory given in Hoyle's report, because this is a consequence of the field equations (7) independently of the particular form of the right hand side of these equations. Contrary to usual beliefs it is even possible to define a consistent expression for the total energy density consisting of a matter part and a gravitational part. If this expression for the energy density, which is given in a paper appearing in the next number of *Annals of Physics*, is applied to the case of the metric (8) for a homogeneous and isotropic universe one finds that the energy density is zero everywhere and at all times. This means that the positive matter energy is constantly counterbalanced by a corresponding amount of negative gravitational energy. So also with the expression (6) for $T^{\mu\nu}$ which implies a continuous creation of matter we will certainly have conservation of the sum of matter energy and gravitational energy for any finite region in space. But my question is now the following. Can the expression (6) be regarded as a kind of phenomenological expression for $T^{\mu\nu}$ which ultimately should be derivable by considering suitable elementary processes in which matter is created ? Do you have definite ideas about the kind of elementary processes which could explain the appearance of the α term in your expression for $T^{\mu\nu}$?

Dr. Hoyle. — The answer to the first part of Dr. Møller's question is that my expression (6) is certainly phenomenological.

The answer to the second part is that one ought to be able to derive the α term from the theory of elementary processes, but that so long as the whole gravitational theory remains outside modern particle physics, I do not see how this can be done. After all, the whole gravitational theory is really phenomenological !

Dr. Schatzman. — The paper of Hoyle is centered on the hypothesis of continuous creation, and I want especially to discuss that hypothesis. In equation (3), the tensions p^{ij} are derived classically from the derivatives $\frac{\partial v^j}{\partial x^i}$. The coefficients α and β depend on the fluid properties. But later on (equation 6), the quantity becomes a universal constant, not depending on the properties of the fluid, but attributed to the properties of space. Such an hypothesis is completely arbitrary.

On the other hand, it seems extremely difficult to grasp the creation process which Hoyle has imagined. In order to show how the process implies conservation of energy, Hoyle shows that in the domain of influence, of radius c/H, of a single particle, its potential energy is negative and equal to its rest energy. One should imagine a kind of flux of sub-particles in vacuum, conveying from everywhere towards the center of a sphere of radius c/H the collection of these sub-particles leading to the production of an elementary particle. But naturally, such a process, if it were possible, should be treated in a quantum way and should give all properties of the created particles, including spin.

I would like, on the contrary to show how, acording to the most modern ideas of quantum mechanics, one can reach a annihilation vague picture of what might be creation and ambulation of elementary particles.

The starting point of all recent work in quantum mechanics lies in the fact that elementary particles (leptons, mesons, hyperons) transform into each other, according to some selection rules. The fact that these particles transform into each other shows that they are different aspects of the same substratum.

We have thus supposed that the vacuum is not empty, but is the seat of an intense agitation.

Interaction between two particles appears like a coupling between two real waves of ether. With such ideas in mind, one could imagine that, from time to time, and rarely, as a consequence of a weak coupling between an elementary particle and the chaotic ether, a particle can be annihilated, and can come back to the chaotic ether. Conversely, the chaotic ether could, from time to time, as a consequence of fluctuations, lead to the creation of an elementary particle. In such a theory, creation and annihilation of particles can be imagined. However, one has the impression that the two processes should be equally probable, the number of creations being equal to the number of annihilations. But the possibility remains that curvature of space favors one process compared to the other. Anyhow, the difficulty would remain of knowing the spin and the symmetry properties of the created particles.

If such a theory could be elaborated, nothing says that the rate of creation which will be found, would have any relation with the rate of creation elaborated by Dr. Hoyle.

On general grounds, it seems difficult to reject the possibility that photons can interact with vacuum, and that such a weak interaction leads to the change of wave length of photons during their propagation.

The main point is that we observe the red shift; but we can imagine three phenomena which can determine the cosmological model, without, for the time being, any possibility of distinghuishing between these phenomena : namely recession, age effect on the photons, continuous creation. Dr. Hoyle has supposed that the whole red shift is related to creation and recession; but it is quite possible that a deeper physical theory, yet to come, would give the real magnitude of these phenomena.

Dr. Hoyle. — I would like to emphasize my previous point about the gravitational theory still being phenomenological. Newton's inverse square law was a purely phenomenological guess, but it was a very valuable guess. In a similar way, I think it is possible to see from looking at the cosmological problem what sort of a form the creation of matter might take. Particularly, I think it is clear that the creation, and its reverse process of annihilation, cannot be understood unless the curvature of space is included in the problem, as Dr. Schatzman mentioned. It must be much easier to guess the form of the expression for creation by working from the macroscopic cosmological side rather than by working from microscopical field theory.

Dr. Heckmann. — I feel the difficulty that an infinite number of similar arbitrary alterations of the field equations is possible, each of equal or even larger basis in normal physics. Do you want us to prove or to disprove each time such theories ? I feel they direct research in a wrong direction.

Dr. Hoyle. — There is no question of an infinite number of possibilities if the restriction is made that the resulting theory must be in adequate agreement with observation.

Exactly the same objection could be made to Newton's inverse square law. It was only one of an infinite number of possibilities. And it was not — and still is not — explicable in terms of microscopic physics.

Dr. Mac Crea. — Dr. Hoyle described the compression of cool material by surrounding hot material. This is very important. For, from the mean density of matter in the universe and the gravitational constant, we can form only one characteristic time. Apart from factors of the order of unity, this is the Hubble time and it is also the time for formations of condensations by gravitation alone. Thus gravitation alone can scarcely produce condensations in the time required; so we must have Hoyle's pressure-effect.

However, there is a difficulty. Neutron decay will in the first instance give the electrons a temperature of the order 10^9 degrees, as stated by Hoyle, but not the protons. The time of relaxation, required to give the protons the same kinetic temperature as the electrons, is, I think, many times the Hubble time. Consequently I doubt whether the " hot " material can move fast enough to compress the " cool " material in the way suggested by Hoyle. The general problem is : How does a mixture of electrons and protons with different kinetic temperatures expand into a vacuum ?

Dr. Hoyle. — In the latter problem, the electrons and protons come to share their energies. The only difference from the usual problem is that the initial acceleration into the vacuum is slower.

You can see this by considering the simpler problem of an infinite slab of gas. The gas expands normal to the slab, equally in the two possible directions. The ultimate kinetic of expansion is given by the total thermal energy of the gas, irrespective of how this energy happens to be distributed in the first place.

Dr. Wheeler. — I understand that your picture assumes that hot gas with a density of the order of 10^{-29} g/cm^3 squeezes cooler gas — originally of about the same density — to raise the density of that cooler gas to a value of the order of 10^{-24} g/cm^3. This increase in density by a factor of 10^5 will be accompanied by a decrease in linear dimensions of the order of 50. However, we know of course that whenever a gas of lower density is pushing on a gas of higher density, Rayleigh-Taylor instability occurs and prongs and spikes will be formed. The mechanism is seen most easily by thinking of the accelerative force replaced by an equivalent gravitational field. A pool of liquid lying on the floor is stable but placed on the ceiling will pour down against the supporting force of the air in spikes and sprongs. Would not this instability phenomenon in the case of a gas compressed fifty fold seem sufficient completely to destroy the integrity of the gas volume and to disperse it into a multitude of oddly shaped pieces ?

Dr. Hoyle. — Some degree of instability would be desirable to produce a number of galaxies rather than one simple condensation.

I would not accept Dr. Wheeler's point unreservedly, however. We know that prominences form by cooling in the solar corona, the rise of density being very considerable, although not quite as great as the factor I have considered. Also dense clouds form in the interstellar medium, the rise of density being in some cases by a factor 10^3, and possibly even more than this. In the latter case, pressure effects probably play an important role.

Dr. Bondi. — It has been suggested that in the steady state theory one departs from the ordinary procedures of physics. This is not so, as the procedure adopted is closely analogous to that in thermodynamics where equilibrium is postulated and observable consequences deduced without detailed consideration of the processes establishing equilibrium.

It may be worth pointing out here, that Gold and I have consistently taken a rather different attitude from Hoyle's. We feel that, as the assumption that the universe is in a steady state leads to observable consequences without any field theory formulation, no advantage is gained by tackling now the obscure and highly ambiguous problem such a formulation presents.

The observational tests of the steady state should be made and they are not mixed up with the problem of the formulation of a field theory.

Dr. Gold. — The condensation of hot gas to make galaxies may occur mainly in the cold gas that remains in the wake of existing galaxies. The long filaments that have been seen associated with galaxies have been a major puzzle as no mechanism can be thought of that would draw out stars into such lanes or hold them together. These may be the wakes of cold gas, and the energy emitted may arise from the energy available as the hot gas condenses on the cold.

Dr. Heckmann. — Am I right in understanding that the present discussion concern points which are not really directly connected with Hoyle's idea of a steady universe ?

Dr. Fierz. — May I ask Dr. Hoyle, if his process of creating matter is not more specifically one of creating nucleons only and not anti-nucleons. Charge must be conserved, but not nucleon-number. If this would be so, time reversal also would be destroyed.

If matter and antimatter would be created, but these two would repel one another, then one would give up the equivalence-principle and so get out of general relativity completely.

Dr. Hoyle. — Since the mathematical part of the theory is macroscopic, it only gives the rate of increase of proper density, and doesn't identify the created particles. This has to be done by a separate postulate. In my paper I considered the possibility that the created particles are neutrons, since this hypothesis leads to very interesting consequences. But there are clearly other possibilities that might be considered — e.g. that the created particles are hyperons.

Dr. Mac Crea. — The energy situation described by Hoyle may be imagined by the help of a rough physical analogy. Suppose we stretch a strip of elastic and consider a part of given length. We pull matter out of this part; but we also do work on this part and so put the mass-equivalent of this work into it. It is possible to imagine the properties of the material to be such that the mass pulled out is equal to the mass put in. Thus we obtain a model of system that is expanding but the amount of mass (or energy) in any portion of constant size remains constant.

Dr. Perrin. — Do you expect antinucleons placed at the same point to have opposite gravitational potential energies ?

Dr. Hoyle. — From the cosmological point of view, the situation would be extremely interesting if this were so, but of course, as Fierz has already remarked, the relativity theory would then be destroyed. At the present moment, I prefer to accept the relativity theory and to suppose the gravitational energies are the same, not opposite.

But I think the other possibility must not be overlooked.

In addition to this, I would like to answer Dr. Lemaître's question about conservation : I would not agree that the steady state theory violates conservation. It changes the nature of the quantity that is conserved, but the whole history of the conservation laws of physics shows repeated changes of the conserved quantities.

Regarding this question of ghosts : the theory being considered within the framework of a smooth fluid naturally will not deal with atomicity, so that it cannot be said whether there is a compensating appearance and disappearance of individual particles. But it can be said that matter (in the sense of a fluid) appears. This shows itself clearly in the equations of the theory.

The Arrow of Time

by **T. GOLD**

Subjectively, we are very clear about the sense of the arrow of time. There is no doubt in our minds which way time runs, what is future and what is past. The fact that introspection gives us a clear answer to the question whether there is a sense in which time runs makes it all the harder to discuss the question objectively as a problem in fundamental physical theory.

At first one might think that there is no real problem there. Why should not the time coordinate be equipped, as it were, with an arrow at each point which singles out for any process the positive time direction? Why should the world not be quite unsymmetrical with respect to past and future?

We have no doubt that the world is, in fact, unsymmetrical in this way. But it is a remarkable fact that the laws of physics, one by one as they have been discovered, have been found to be quite symmetrical with respect to the sense of time. Newton's laws of gravitation and dynamics single out no sense of the time coordinate. If somebody recorded the motion of the planets and reversed the record of the time coordinate, this would leave it an example of a dynamical system that is as much in accord with Newton's laws as the actual. The change from Newton's laws to Einstein's did not affect this symmetry.

The laws of electrodynamics, the Maxwell-Lorentz theory, similarly are quite symmetrical, and so are those of quantum theory. Could we argue that all this is accidental and that we will discover some other physical law which clearly specifies the sense of time and which is responsible for giving us our ideas on the subject? This, I think, is not a plausible explanation, since systems we understand in detail seem to show time's arrow. But yet there must be some influence that serves to determine the arrow of time.

Usually at this point in the argument statistical or thermodynamical ideas are presented and the case is made out that it is through the investigation of these fields that the elusive arrow is found. « The entropy of any isolated system will always increase and never decrease »; so, it is said, you must merely look at a system at two instants of time and determine at which the entropy is greater. That will then be the later instant.

As there is no doubt about the correctness of this, the argument is usually not pursued any further. We and everything around us are simply taken to be aware of the arrow of time by the operation of the statistical processes which, after all, we understand very clearly. Why should there be anything else to think about?

One has to pursue this reasoning a little further, though. Why does the arrow of time appear when we are dealing with the statistical superposition of effects, each of which is determined by laws which have no arrow? Surely the fact that we had to deal with the problem in statistical terms rather than compute in detail the behavior of all the constituent parts of our system, that constituted merely a lack of precision; surely it is not by rejecting information about our system that we can make it reveal to us the sense of time which it would otherwise not show. So, let us see whether we can find the arrow without the statistics.

To see whether the system that we examine does or does not reveal the arrow of time, let us suppose that we are given a number of snapshots of it only, and we are asked whether we can be sure to sort them into the correct temporal sequence. We are told that the system was interfered with before the first and after the last of the snapshots, but was left quite undisturbed in the intervening period. Now if, for example, the system were a box full of gas and we found on our snapshots that one contained all the gas in one half of the box, another one showed 70% in that half of the box and 30% in the other half, and a third showed just 50% in each half of the box, then we should surely order the snapshots in that way. We would say that somehow the gas must have been put into that half of the box where it was found, and the first snapshot must have been taken very shortly after it was put there, because we know that when left to itself, it will quickly expand through the rest of the box and fill it uniformly. In whatever way the box is interfered with after the last snapshot was taken does not enter into our considerations; but the interference the system received before the first is of importance in deciding the arrow of time. Of course, this argument is not absolutely certain; we might have ordered the snapshots the wrong way round, for the gas might have been uniformly distributed in the box to start with, but all the motions of its molecules might have been so contrived that by the operation of the ordinary laws that apply to the collision of molecules they will all have migrated at one time into half the box. But this, we think, is highly improbable. We can thus take on a bet and offer very high odds that we will be able to order the snapshots in the right sequence, but we cannot actually prove it. If we did a similar experiment with fewer particles, the same would apply, but we could offer only lower odds. Essentially, looking at a system of many particles is not very different from having many systems of a few particles on the same snapshots. There we would order the snapshots according to the appearance of one of the systems on the successive pictures, and we would then decide that our probability of being right increased when we saw each one of the other systems agreed with the sequence we had decided on. The certainty about the sense of the arrow of time then arises just from having many checks, and for this reason complicated systems reveal the effect most clearly; but this does not explain why each of our simple systems displayed an arrow at all.

The interference from outside clearly had something to do with it. If we take any system and isolate it from external influence completely and for a very long time and then take a series of snapshots, there will no longer be any way of deciding on the sequence from a subsequent examination of the pictures. We might still be able to recognize clearly the operation of fundamental physical laws in the changes that had taken place from one snapshot to the next. If all the snapshots were taken in sufficiently rapid sequence, we might be able to arrange them in order *l* to *n*, but we would not know whether it was *l* or *n* that was taken first. The physical laws that determine the motions that we can see all being time-symmetrical, there might be plenty of clues to demonstrate the laws and to find the neighbors in the sequence to any picture, but no clue at all about the arrow of time.

When the system was not isolated, there was, as usual, no doubt about the sense of time. After it became isolated, the arrow of time evidently persisted for a while, not definable with certainty, but only with a probability that decreased from a high value initially to zero. The time scale of this decrease of probability depended upon the details of the system.

It is this rule that isolated systems initially, after their isolations, retain and then gradually lose the arrow of time that makes its appearance in the statistical and thermodynamical definitions. « Entropy of an isolated system always increases », is a way of saying that after the system was isolated, it still showed changes from which the sense of time could be deduced. In some systems the effect is best described in thermodynamic terms, and entropy is then the relevant quantity. But in other systems other statistical descriptions may be more convenient. It is inconsequential from this point of view whether the system is deterministic or not; that case can be argued equally well with a number of billiard balls assumed to behave accurately according to Newton's laws of motion, as with a system of photons and atoms in a box where we do not know of any way of specifying the laws of motion except through probabilities.

Some simple mechanical systems seem to give a more clear-cut answer than others, but on closer examination are really not different in principle. For example, a ratchet with a tooth may be known to be an isolated system during the time that it changed from one state to another. There would seem to be no question that it must initially have had momentum in the direction in which it does not jam. One might think that any system only has to be equipped with such a ratchet mechanism in order for the arrow of time to be defined there. But, of course, this only works through the dissipative mechanism of the claw, and one has to allow that the process could happen in reverse if all the thermal motions of all the atoms in the claw and in the ratchet were all just right, namely, just the reverse of those that would be set up by the ratchet moving in the allowed direction. The claw would then spontaneously bounce open and the bar would recede by one tooth.

This would be in no conflict with the laws of motion, but, because of the great number of atoms whose motions would have to be just right, it is an effect whose chance of occurrence is negligibly small. But if we had such a ratchet in a system that had been isolated for a very long time, then the probability of it moving by Brownian motion by one tooth in one direction is exactly equal to that of moving by one tooth in the other; and then again no arrow of time would be in evidence.

So we can be confident that the same rule applies to all systems: interference from without enables them to show time's arrow, and this may persist for some time after the interference has ceased. All completely internal effects merely reflect the physical laws that apply, and all those are then strictly time-symmetrical. On whatever scale we choose our system, we have to go to a larger scale to understand how it contrived to know the arrow of time.

Up to what scale, can we pursue this argument? On which scale do we find a law whose operation in fact serves to determine time's arrow for all the smaller scales? Let us take, for example, a star, and suppose we could put it inside an insulating box. It would still be true, then, that when the star has been in the box for long enough (which in this case will perhaps be rather long), time's arrow will have vanished. There is no reason to expect anything to be different in principle from the laboratory scale. But now if we were to open for a moment a small window in our box, then what would happen? Time's arrow would again be defined inside the box for some time, until the statistical equilibrium had been reestablished. But what had happened when we opened the hole? Some radiation had, no doubt, escaped from the box and the amount of radiation that found its way into the box from the outside was incomparably smaller. Some influence from outside had got in — though the only physical effect was that photons from inside got out.

The escape of radiation away from the system is, in fact, characteristic of the type of « influence » which is exerted from outside. Any outside influence to a system that gives it time's arrow can be traced to be associated with that process. The thermodynamic approach would be to explain that free energy was required for the interference, and that free energy can only be generated from the heat sources in the world by means of heat engines working between a source and a sink. There may be a variety of sources, but the sink is always eventually the depth of space, although there may be a number of intermediate cold bodies.

It is this facility of the universe to soak up any amount of radiation that makes it different from any closed box, and it is just this that enables it to define the arrow of time in any system that is in contact with this sink. But why is it that the universe is a non-reflecting sink for radiation? Different explanations are offered for this in the various cosmological theories and in some schemes, indeed, this would only be a temporary property. In the steady state universe it is entirely attributed to the state of expansion. The red shift operates to diminish the contribution to the radiation field of distant matter; even though the density does not diminish at great distances, the sky is dark because in most directions the material on a line of sight is receding very fast, and its radiation, therefore, shifted very far to the red.

The large scale motion of the universe thus appears to be responsible for time's arrow. A picture of the world lines of galaxies would clearly reveal the sense of time, namely, the sense in which the world lines are diverging. As we go to a smaller scale, this type of divergence is, for most purposes, quite negligible, and it is thus clearly not the local effects of the universal expansion law that make themselves felt. But it is the electromagnetic radiation that brings the effects of expansion down to a small scale. Radiation in the world is almost everywhere almost all the time violently expanding. This expansion of the radiation is, however, only made possible by the expansion of the material between which the radiation makes its way, that is, the expansion of the universe.

If we examine the pattern of world lines of systems that are open to the universe there will be much branching apart and much less convergence when looked at in the sense in which we think of time. As an example, the average hydrogen atom in the universe will suffer a conversion to helium in a time of the order of 10^{11} years at the present rate. This corresponds to an emission into the universe of some 10^{6} photons in the visible spectrum. This photon expansion going on around most material is the most striking type of asymmetry, and it appears to give rise to all other time assymetries that are in evidence. The preferential divergence, rather than convergence, of the world lines of a system ceases when that system has been isolated in a box which prevents the expansion of the photons out into space. Time's arrow is then lost; entropy remains constant.

The motion of the universe is thus most intimately connected with all processes down to the smallest scale. A more profound understanding of physics than we now have might, in fact, allow one to deduce the expansion of the universe from an observation of the small scale effects only.

We see the universe expanding and not contracting. Does this mean that of the two possible senses of motion, nature chose one? Surely not. This would be the case if the laws of physics were not time-symmetrical : then an expanding universe would be a system that is distinguishable from a contracting one. The laws could describe two types, and ours would be one of them. In such a case the laws of physics would be capable of defining more schemes of the world than we have to look at. The laws would be too wide to fit the case, and we would suppose this due to some misunderstanding we have made. But just this is avoided by the time-symmetry of the laws. In a universe where no arrow of time exists except that defined by the motion, there is only a single possibility. We would need an independent clock to say whether the universe is expanding or contracting, and we have none. All the clocks we do have are themselves run by the motion.

It follows that if all the laws of physics are time-symmetrical, they would not be able to describe a contracting universe. If, naively, we think that, after all, we might have seen blue shifts, in the spectra of distant galaxies instead of the red shifts we must be making an error in pursuing the detailed consequences of the motions of the galaxies. If, in calculating radiation effects, we took the particular solution (of the intrinsically time-symmetrical electrodynamic theory) given by the retarded potentials, then, of course, there would appear the second possibility. But that is rather like supposing that an independent clock exists and that the laws are not time-symmetrical. Wheeler and Feynman, and Hogarth have considered the question of a time-symmetrical electromagnetic theory and the way in which the choice of retarded potentials appears appropriate depending upon the cosmological boundary conditions.

There is nothing new in the idea that the physical laws are more symmetrical than the universe to which they apply. For example, the principe of Galilean relativity states that the physical laws are the same in all inertial systems. But, on the other hand, one particular such frame can be singled out through the observation of the universe, namely, that particular frame where the observer would see the expansion of the universe occur symmetrically around him. A cosmological observation, therefore, specifies one out of an infinite set of frames that would all be equivalent from all other points of view. With the arrow of time, it is not really dissimilar. A cosmological observation, such as, for example, opening a window in a box and letting some radiation escape, is the means of distinguishing between the two otherwise quite equivalent senses of time.

At this point one should think, perhaps, why it is that we are subjectively so sure that time « really goes » in one sense and not the other. With symmetrical physical laws we can, after all, construct the present equally well from a sufficient knowledge of the future as from such a knowledge of the past. Why, then, do we give the past a status quite different from that given to the future? We do not generally think of predicting the past, of constructing it from the present and the knowledge of the physical laws, yet that is what we do with respect to the future. We think of some evidence about the past as entirely definite, and we think it a rule that the future can not be known with certainty. Why is this so in a system operating with time-symmetrical laws? Why do we believe in the cause and effect relationship between events when, after all, there is no strictly logical way with time-symmetrical laws of specifying which is the cause and which the effect? Today's position and momenta of the planets are the cause of their position and momenta tomorrow. But this could equally well be stated the other way around — that tomorrow's configuration causes today's. In more complicated systems, as we have seen, there is a general overwhelming tendency for branching of world lines in the forward direction of time. If a lot of information is lacking for precise prediction, as for example all the photons out in space that have escaped, then the configuration of the system at one instant will serve to define much better its configuration in the sense in which its world lines are generally converging than in the sense in which they are generally diverging. The past will be better known than the future. It is difficult to reconcile oneself to this explanation that the assymmetry between past and future which seems so profound to us should be no more than an asymmetry connected with the probability of « predicting » correctly into the two senses of time. The qualitative difference arises from a statistical quantitative one. But, of course, one must appreciate that we are systems of a high order of complexity, and that the statistics are concerned, therefore, with a very large number of possible states. In such systems there may be, in effect, complete certainty attached to the consequences of the laws of chance. And these consequences appear, then, as the laws of physics.

The symmetry of the laws with regard to the time axis is then just what was required to prevent them from being too wide, from being able to describe more than just our universe. What is the situation with respect to other symmetries of physical laws?

Symmetry with respect to the sign of the electric charge and with respect to mirror reflection was thought until recently to be separately obeyed by all the laws of physics. Now the discovery of the non-conservation of parity in weak interactions implies that the laws are not invariant to a mirror reflection alone. A certain « handedness » is shown to be resident in elementary particles. Since a right-handed screw becomes a left-handed screw when viewed in the mirror, such a particle will be transformed into something different by reflection. But this does not force us to believe that nature is not mirror-symmetrical, for there may be complete symmetry between matter and anti-matter. For every right-handed particle of matter there may be a left-handed one of anti-matter having all the same properties, but possessing the opposite sign of electric charge or magnetic dipole moment. The symmetry may be complete, but only for the combined operation of mirror reflection and a change in the sign of the charge. I suppose that most physicists now would regard this as the most likely situation.

If symmetry is preserved only with respect to the combined operation, then this could really be understood best if charge had a geometrical representation, possessing a « handedness .»

Should we now think that the universe is in actual construction, symmetrical between matter and anti-matter? Or have we here a another case of a symmetry in the laws not represented by symmetry of construction?

It is, of course, possible that amongst the various galaxies there are as many made of matter as of anti-matter. Not enough meeting ground exists between them for the annihilation process to be plainly demonstrable. From observation we can not yet tell. If, in fact, all galaxies were constructed from one type of matter only and anti-matter appeared everwhere, as it does here, only as a rare freak, would this, then, imply that the laws of physics are too wide? One might think that, after all, if our type of matter is the right-handed sort, the same laws of physics would have allowed the construction of a universe entirely similar to ours except made of matter of the left-handed type. Two universes would be specified by the laws and only one of them arbitrarily selected as ours.

But, do these laws really specify two different universes? The difference between a right-handed system and a left-handed system can be defined only when they can be compared. There is no absolute definition of either. If two systems cannot be compared, either directly or via some intermediary systems, then there is no way of defining whether they are of the same or the opposite handedness. But this is just the situation with respect to the two universes that would seem to be defined. If they are different universes, they cannot be brought together to be compared, and unless they can be so compared, there is no sense to be attached to the statement that they are of different handedness. The two universes, if they cannot be compared, are thus identical. We have then, again, the situation that a precise symmetry in the physical laws, namely that between matter and anti-matter, is just what is required to assure that only one type of universe is specified. If any law of physics had not been accurately time-symmetrical, then an expanding and a contracting universe would be different possibilities. If the symmetry between matter and anti-matter were not complete, then two different universes, one composed of matter and one composed of anti-matter, would be possibilities. These symmetries are, then, in each case, just what is required to allow the laws of physics to describe as the only possibility the type of universe we have.

REFERENCES

(1) A. Grünbaum, *Das Zeitproblem Archiv für Philosophie*, **7**, 165 (1957).

(2) J.A. Wheeler and R.P. Feynmann, *Rev. Mod. Phys.*, **21**, 425 (1949).

(3) J. Hogarth, « Advanced and Retarded Potentials in an Expanding Universe » (unpublished thesis).

Discussion of Gold's Report

Dr. Rosenfeld. — Firstly, from the historical point of view, I might recall that the situation regarding the sense of time in thermodynamics was already made quite clear by Clausius, in the paper in which he formulated the second law. The axiomatic analysis was performed in a more physical way than Caratheodory's by Ehrenfest, who emphasized the logical possibility of the two opposite senses of irreversibility and the empirical origin of the axiom which fixes this sense for the actual world.

The point I wish to make, however, is that the question of the sense of time arises at a deeper epistemological level than the statistical one at which it is usually considered. In order to give a meaning to the fixation of the sense of time, it would seem necessary to go back to the process of observation in its most primitive aspect : in order to be able to make any statement about any object — not only complex systems, but individual atomic systems as well — one must get some signal from this object, and the transmission of such a signal (a photon for instance) from object to observer defines the sense of the time; we call the time at which the signal reaches the detector «later» than that at which it has left the object. This fits in with the general conception of the necessity, for the definition and use of the quantities entering into the formulation of the fundamental laws, to make a sharp distinction between the system about which we are talking and the system we use as observing apparatus. The need for such a separation was first explicitly recognized in quantum theory, but it is just as imperative from the epistemological point of view in classical physics, although there its quantitative implications are negligible.

The part of the system used as observing device is characterized by a property which might be called « memory », although it need not have any immediate connection whith the physiological memory of a human observer : it has to be in a general sense some permanent recording of the interaction of the object-system with the apparatus (such as a spot on a photographic plate).

It is essential for the unique definition of the quantities concerned that the observing system be macroscopic.

Insofar one may say that the irreversibility connected with the process of observation has a statistical character, one must realize, however, that in discussing the mechanism of the observation process, one deprives this process of its function and regards it itself as an object of observation. It is therefore unavoidable to maintain the irreversibility arising from the separation of the system into object and observing device as a primary one. The statistical irreversibility exhibited by complex systems must thus be analysed on its own right, in terms of a time concept involving a sense of time already defined in the way indicated by reference to the process of observation.

The last remark also applies to the question of the meaning of terms « expansion » or « contraction » applied to the evolution of the universe, and would seem to remove the basis for Gold's conclusions in this respect.

Dr. Prigogine. — I would like to point out that boundary conditions are not sufficient to build a flux of irreversible processes and to give a sense to the arrow of time.

Indeed the simplest irreversible processes, like heat conduction, viscosity etc..., occur in isolated systems.

In recent years great advances have been made in the quantitative description of irreversible processes.

It appears that, to express it in a somewhat broad and imprecise form, irreversibility is a property of N body systems for N large and for special sets of initial conditions.

Under such conditions one can define total information and macroscopical information and prove that for all times macroscopical information is flowing out to the molecular information and is in a sense lost.

One may however note that as a consequence of Poincaré's recurrence time the motion of a finite system is always reversible for infinitely long times. Therefore the irreversibility to which I referred is a relative one. But as is well known the recurrence times for large systems are extraordinarily long.

Nevertheless it is true that « absolute » irreversibility is a property of the universe as a whole.

Dr. Zanstra. — Some speakers have remarked that the arrow of time is imposed on us. I should like to specify this a little further.

The physical processes which take place in human bodies are all irreversible processes and as such have the arrow of time in them. So thoughts are impressed on the brain by a physical process, say, in a rather materialistic way, like a gramophone record is recorded. For this reason memory gives the same arrow of time as irreversible processes. (At any rate as far as the human body goes, apart from or as part of the universe as a physical system.)

Dr. Mac Crea. — Consider two observers moving independently in any manner in free space. Whatever may be their initial motions, after a sufficiently long time each will see the other as *receding* from himself. In other words, the observers can see each other approaching for at most a finite time, while they *must* see each other receding for an infinite time.

Thus we appear to obtain an arrow of time from the simplest possible kinematic situation.

This property was described by the late E.A. Milne. I do not properly understand its significance.

Dr. Schatzman. — It seems to me that we have to be more clear about the philosophical foundations of the concepts of space and time.

Matter is in motion in space and time; motion is the way of existence of matter; we cannot conceive matter without motion and motion without matter. In these conditions the exact reversing of all velocities in an isolated system would not mean the reversing of time. For example, it seems impossible to consider that among two identical double stars one revolving in one direction, the other in the opposite direction, one is going down the sense of time the other is going up the sense of time.

Dr. Oppenheimer. — If the weak interactions of atomic physics, would — contrary to expectation — not be invariant for time inversion, would this have any consequences for cosmological or cosmogonic questions?

Dr. Gold. — I cannot see that the intervention of a human observer or of a photographic process are very different from this point of view. In each case, the statistical behaviour of a system may show an asymmetry in the two senses of time, and in each case this can be attributed to the system having been in an « unusual » condition near one end of the period of observation.

It is true that for very long periods of time systems may exhibit a clearly defined sense of time, and so may subsystems of this that may have become isolated from them. It is with those that thermodynamic discussions have usually been concerned. Looking far enough back for any such system we see however that there occurred some process which established a connection with the large scale system of the Universe, such as the emission of radiation into the depth of space. It is through this interaction that the large scale motion of the Universe is of importance, and it defines with overwhelming probability the sense that we call the forward one for the system, as the one in which the Universe is seen as expanding.

It is not clear to me what the consequences would be if the boundary conditions given by the Universe had been different at different epochs; but this would seem to undermine the arguments, and the reason for the uniformity in the behaviour of systems would then be lost.

There is no difference between a contracting and an expanding universe — the sense of time we adopt as the forward one in such a moving universe is the one in which it is seen as expanding.

As to Professor Mac Crea's remark, attributed to Milne, I must say that I can not see any time asymmetry implied at all. If the points can be thought of as becoming indefinitely removed from each other they can also be thought of as starting indefinitely far apart.

If any fundamental process were found to be not time invariant, this would of course destroy all these arguments. There would then be two possibilities for the appearance of a moving universe, and we would have only one realised in fact, namely the expanding one.

Some Implications of General Relativity for the Structure and Evolution of the Universe *

J. Barclay Adams, B. Kent Harrison, Louis T. Klauder, Jr., Raymond Mjolsness, Masami Wakano, John A. Wheeler, and Raymond Willey.

Palmer Physical Laboratory
Princeton University, Princeton, New Jersey

SUMMARY

For the Solvay Congress a review was requested of some of the conceivable observational tests of Einstein's general relativity that relate to the structure and evolution of the universe. Five tests are taken up here. Most have been considered before. However, some of the work to be reported is new. Therefore, it may be appropriate to summarize all five points before going into detail :

(1) In the idealization of a spherical isotropic uniform universe the kinematics is completely determined by the maximum radius and by the equation of state of the medium that fills the space. Details appear in figures 1-4;

(2) The maximum radius and the equation of state also determine an expected value for the total density of energy as a function of time. Observations to date are far from being sufficiently complete to check these predictions (fig. 5);

(3) Asymmetric departures from sphericity are magnified in early and late stages of the dynamics of the universe by an instability

* Reported in abbreviated form by J. A. Wheeler at the Solvay Conference. Brussels, 10 June 1958.

phenomenon akin to the well known Taylor instability in hydrodynamics;

(4) Gravitational contraction of a sufficiently massive star provides a mechanism to convert a significant fraction of the mass of a nucleon into energy;

(5) A system which contains as many nucleons as the sun, or more, arrives at the end of its thermonuclear evolution at a condition which lies outside the scope of existing theory. The further investigation of this condition may well make important contributions both to elementary particle physics and to astrophysics.

MEANING OF THE TERM, " EINSTEIN'S THEORY ", (J.A.W.)

In speaking about Einstein's theory in this connection we mean not only the system of differential equations associated with his name, but also two further points, the present tentative arguments for which he gives in his book ([1]) :

(1) The universe is closed;

(2) No " cosmological " term is to be added to the field equations.

It is probably best at this time to accept these points as part of what we mean by the phrase " Einstein's version of general relativity ". To inquire further about the foundation of these assumptions would lead us to an analysis of Mach's principle. This principle requires a much deeper theoretical investigation than it has received so far, but this is not the occasion to start the study!

To regard Einstein's theory in proper perspective, we can recall the revolutionary concept of Riemann in 1854 that space is not an abstract mathematical construction that stands unmoved above the battles of matter and energy. In his view it not only affects physics but also is affected by physics and even owes its existence to physics. Einstein, when he gave this view a definitive mathematical formulation, arrived for the first time at a master theory of physics, on a higher level than any theory before. This theory (1916), as well known, not only derives the field produced by arbitrary distributions of mass-energy, but also (1938) deduces the equations of motion of concentrations of mass energy. The equations of motion followed for the first time as a consequence of field theory, not as a separate postulate.

However far Einstein went in formulating gravitation theory and in tracing out its consequences, it is appropriate to recall that he was guided by a larger vision. Let me describe his dream in these words : Matter and fields are not independent actors striding about an indifferent arena of space and time. Instead, curved empty space is the magic medium of which everything is made. Curvature of one kind here represents a gravitational field. Curvature of another kind there manifests itself as an electromagnetic field. A semi-stable condition of curvature shows itself to the observer as a mass that holds together for a substantial period of time. All of matter, force and motion come down to pure geometry and its change with time; reduce, in a word, to " geometrodynamics ". Behind the mathematics of general relativity this is Einstein's new vision of physics.

Whether Einstein was fundamentally right in regarding matter as pure geometry, we are very far from being able to decide today. Our judgements are completely at sea. We are only now in process of discovering what are the new implications of Einstein's geometrodynamics when it is combined with quantum theory. Here lie deep issues of theoretical physics.

Fortunately these deep issues have little bearing on most of the large scale phenomena in the universe. Of such phenomena general relativity provides a simple straight-forward treatment. Moreover, Einstein's geometrodynamics as we know it today is unique, in the sense that we know of no acceptable alternate theory of the same all-inclusiveness. The simplicity and uniqueness of general relativity give great importance to its consequences, especially to those which today or in the future may be subject to observational test.

KINEMATICS OF A SPHERICAL ISOTROPIC UNIFORM UNIVERSE, (L.T.K., Jr.; M.W.; J.A.W.; R.W.)

Robertson and Walker have shown that the conditions of sphericity, isotropy and uniformity demand that the metric be transformable to the form :

$$dl^2 = -dT^2 + a^2(T)\,[d\chi^2 + \sin^2\chi\,(d\theta^2 + \sin^2\theta\,d\varphi^2)]; \qquad (1a)$$

or equivalently :

$$dl^2 = a^2(\eta)\,[-d\eta^2 + d\chi^2 + \sin^2\chi\,(d\theta^2 + \sin^2\theta\,d\varphi^2)]. \qquad (1b)$$

Here $dT = cdt = a(\eta)\,d\eta$ measures elapsed cotime in a comoving frame of reference, $a(T)$ is the radius of curvature of the spherical space, θ and φ are the familiar spherical polar coordinates, and χ is a third angle that gives distance from the origin on being multiplied by $a(T)$. The time parameter η has the following simple significance : it measures the angle around the universe from which light has reached the observer by the time in question.

Einstein's assumptions of closed space and zero " cosmological constant " exclude many of the numerous cosmological models that have been considered in the past. The equation of state *uniquely* determines the connection between a and T (2). The actual case will lie between two extremes of zero pressure and maximum pressure (pure radiation) :

(1) Friedmann's dust filled universe. The pressure is zero. The radius of curvature and the cotime elapsed from the start of expansion are connected by the parametric formulas of a cycloid (fig. 1) :

$$\begin{aligned} T &= \tfrac{1}{2}a_0(\eta - \sin\eta)\ (\doteq a_0\eta^3/12 \text{ for small } \eta); \\ a(T) &= \tfrac{1}{2}a_0(1-\cos\eta)\ (\eta \text{ from } 0 \text{ to } 2\pi). \end{aligned} \qquad (2)$$

Here a_0 is the maximum radius of curvature. Hubble's fractional rate of expansion, H, is given by the formula :

$$\begin{aligned} 1/H &= a/(da/dt) = (a_0/2c)(1-\cos\eta)^2/\sin\eta; \\ &= (T/c)(1-\cos\eta)^2/[\sin\eta\,(\eta - \sin\eta)]\ (> 3T/2c). \end{aligned} \qquad (3)$$

One can use equation (3) to deduce from values of H and T the present value of the parameter η and hence the present and maximum

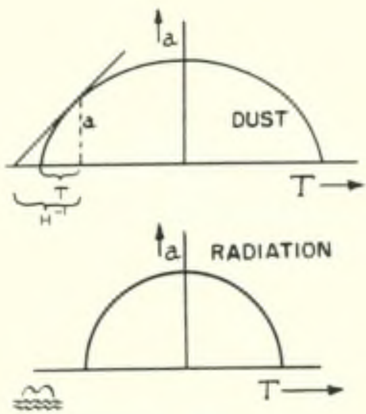

a
a
DUST
T
T
H^{-1}
a
RADIATION
T

Fig. 1.

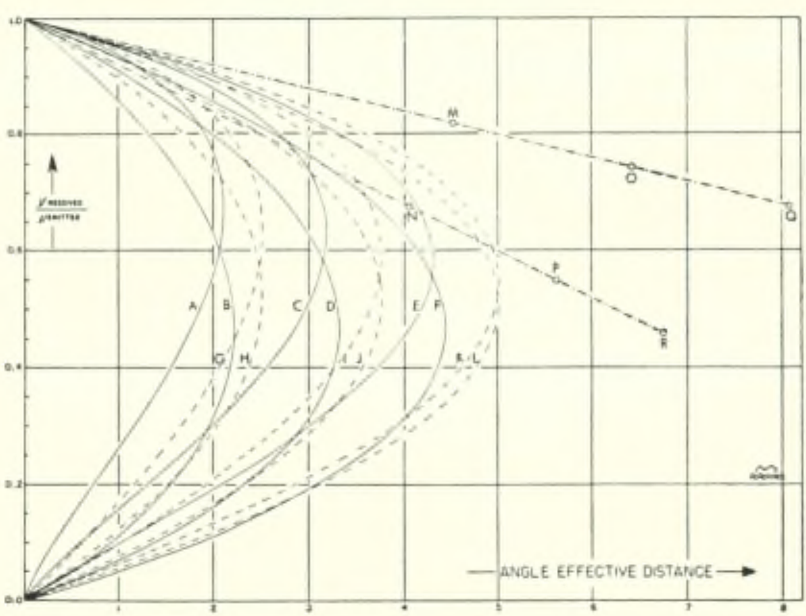

ANGLE EFFECTIVE DISTANCE
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R

Fig. 2.

radii, a and a_0. The total mass, M, remains constant. Also constant — but zero — is a quantity analogous to the sum of the rest energy, kinetic energy and potential energy of the system :

$$\mathrm{M}c^2 + \mathrm{M}\,(da/dt)^2 - 4\,\mathrm{GM}^2/3\,\pi\,a = 0. \tag{4}$$

The mass has the value :

$$\mathrm{M} = 3\,\pi\,c^2 a_0/4\,\mathrm{G}. \tag{5}$$

and the density of matter is :

$$\begin{aligned}\rho &= \text{mass/volume} = \mathrm{M}/2\,\pi^2 a^3 \\ &= (3\,c^2/8\,\pi\mathrm{G})\,(a_0/a^3) = (3\,c^2/8\,\pi\mathrm{G}a_0{}^2)\,(1/\sin^6\ \tfrac{1}{2}\,\eta) \\ &= (3\,c^2/8\,\pi\mathrm{G})\,(a^{-2} + \mathrm{H}^2/c^2) > (3\,c^2/8\,\pi\mathrm{G})\,(\mathrm{H}^2/c^2)\end{aligned} \tag{6}$$

In these formulas the multiplying constant $(3\,c^2/8\,\pi\mathrm{G})$ has the value $(1796 \times 10^{-30}\ \mathrm{gm/cm^3}) \times (10^9\ \text{light year})^2$.

During the expansion a planet continues Newtonian motion about its sun, stars continue Newtonian motion in the gravitational field of their galaxy, and likewise the diameter of the galaxy remains unchanged except insofar as altered by galactic evolution or — in early days — by contact between galaxies. Absolute diameters of typical galaxies are reasonably well known. The apparent or angular diameter of the galaxy decreases owing to its recession from the observer. The ratio of the two diameters defines the " angle effective distance " of the source :

$$\begin{aligned}&(\text{`` angle effective distance '' of source in light years}) \\ &= \frac{(\text{true diameter of source in light years})}{(\text{apparent or angular diameter of source in radians})} \\ &= \frac{a\ (\mathrm{T}_{source})\ \sin\chi\ \delta\theta}{\delta\theta}.\end{aligned} \tag{7}$$

Let the typical galaxy under consideration be parametrized by its angular coordinate χ upon the expanding sphere, the observer being at $\chi = 0$. Assume known H and T_{obs}. Then equation (3) gives η_{obs} and a_0. Equations (1) and (2) give (Landau and Lifshitz)

$$\begin{aligned}\chi = \chi_{source} - \chi_{observer} &= \int_{obs}^{source} d\chi = \int_{\substack{\text{ray of light}\\ \text{from source}\\ \text{to observer}}} d\mathrm{T}/a\,(\mathrm{T}) \\ &= \int d\eta = \eta_{obs} - \eta_{source}\ ;\end{aligned}$$

or : $$\eta_{source} = \eta_{obs} - \chi. \tag{8}$$

The quantity η_{source} parametrically determines the time of emmission and the radius of space at that time. This information fixes the " angle effective distance " $a\,(T_{source}) \sin \chi$.

Example : $H^{-1} = 12 \times 10^9$ yrs, $T_{obs}/c = 7.5 \times 10^9$ yrs; $\eta_{obs} = 1.083$, $a_{obs}/c = 19.96 \times 10^9$ yrs, $a_o/c = 75.14 \times 10^9$ yrs; a galaxy separated from us by an angle as great as $\chi = \pi/2$ has not had time enough since the beginning to send a signal to us; our " horizon " is limited to an angle of $1.083 \times 180^0/\pi = 62^0$; consider then a galaxy at $\chi = \pi/4$; $\eta_{source} = 1.083 - 0.785 = 0.298$, $T_{source}/c = 0.17 \times 10^9$ yrs, $a_{source}/c = 1.65 \times 10^9$ yrs; and the " angle effective distance " of the galaxy is $1.65 \times 10^9 \sin 45^o$ 1.17×10^9 light years — even though its distance is $19.96 \times 10^9 \sin 45^o$ $= 14.1 \times 10^9$ light years at the time the observer receives the light.

It is simple to analyze the red shift by considering a standing wave whose scale grows with the scale of the expanding universe :

$$\frac{\nu_{received}}{\nu_{emitted}} = \frac{\lambda_{emitted}}{\lambda_{received}} = \frac{a_{source}}{a_{obs}} \,. \tag{9}$$

In the example this ratio is $1.65 \times 10^9/19.96 \times 10^9 = 0.0826$.

The fraction of the spherical space included within the angle χ of the observer is :

$$\begin{aligned} f(\chi) &= (1/2\pi^2)\int_o^{\chi}\int_o^{\pi}\int_o^{2\pi} (\sin\chi \sin\theta \, d\varphi)(\sin\chi \, d\theta)(d\chi) \\ &= (1/2\pi)(2\,\chi - \sin 2\,\chi)\,(\chi \text{ from } 0 \text{ to } \pi)\,. \end{aligned} \tag{10}$$

To the extent that galaxies of a given type are distributed approximately uniformly, f will also represent the fraction of all galaxies of that type which are located within the angle χ. In the example this fraction is $f\,(45^o) = 0.0909$. Figures 2, 3 and 4 show the connection between red shift, " angle effective distance ", and fractional count f for selected values of the present Hubble constant and present epoch. These values are listed in the following table. The example already discussed appears as case D in the table.

TABLE 1

The closed spherical isotropic universe of general relativity as affected in size and density by the values adopted for (1) the present epoch (or time elapsed since the start of expansion), T_{obs}/c, and

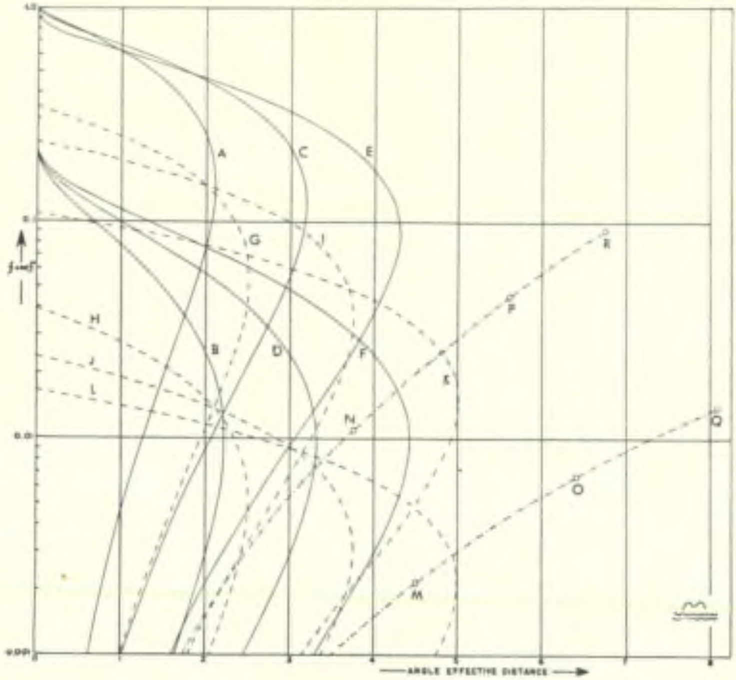

A
C
E
G
I
H
J
L
B
D
F
K
N
R
P
Q
O
M
ANGLE EFFECTIVE DISTANCE

Fig. 3.

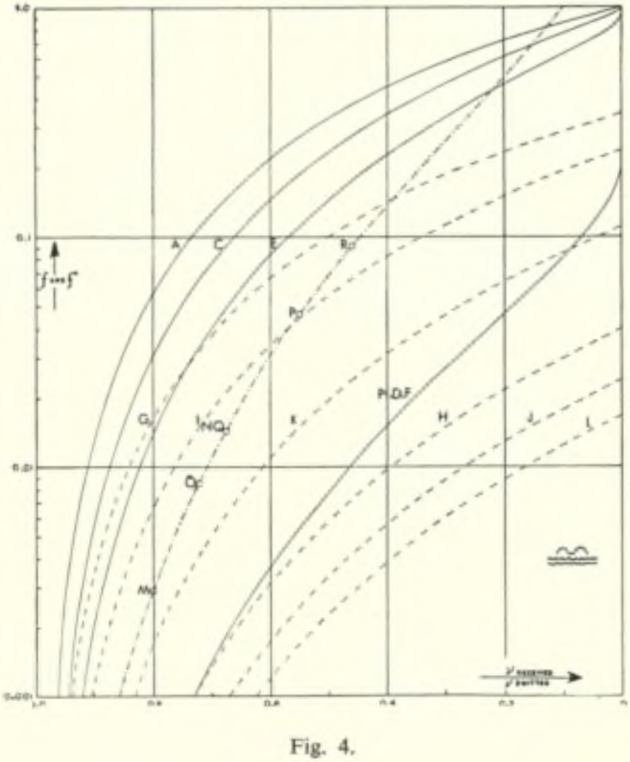

A
C
E
R
P
P.D.F
H
J
L
G
I NQ
K
O
M

Fig. 4.

(2) the Hubble expansion coefficient, H; and by the equation of state — pure dust, with no pressure or pure radiation, with maximum pressure. Other quantities appearing in the table are the maximum radius, a_0; the present radius, a; the parameter η which runs from 0 at the start of expansion to 2π for the dust filled universe (or to π for the radiation filled universe) at the end of recontraction; the present density, ρ, of matter plus mass equivalent

TABLE I.

State	Curve	T_{obs}	H^{-1}	η_{obs}	a_{obs}	a_0	ρH	ρ_{obs}
Dust	A	5	25	2.835	3.86	3.95	2.87	124.32
	B	5	8	1.083	13.31	50.09	28.1	38.2
	C	7.5	25	2.627	6.57	7.03	2.87	44.3
	D	7.5	12	1.083	19.96	75.14	12.5	17.0
	E	10	25	2.354	10.37	12.16	2.87	19.7
	F	10	16	1.083	26.62	100.2	7.0	9.5
	—	6	13	2.137	7.12	9.27	10.6	46.0
Radiation	G	5	25	1.319	6.45	6.66	2.87	46.0
	H	5	11	0.583	16.66	30.27	14.8	21.4
	I	7.5	25	1.128	11.86	13.12	2.87	15.6
	J	7.5	16	0.490	30.00	63.74	7.0	9.01
	K	10	25	0.840	22.39	30.07	2.87	6.49
	L	10	21	0.433	45.47	108.4	4.07	4.97
	—	6	13	0.535	21.89	42.90	10.6	14.32
Bondi	M	5 *	25	—	4.78**	—	2.87	2.87
Gold	N	5 *	12.5	—	4.13**	—	11.5	11.5
Hoyle	O	7.5*	25	—	6.49**	—	2.87	2.87
Model	P	7.5*	12.5	—	5.64**	—	11.5	11.5
	Q	10*	25	—	8.25**	—	2.87	2.87
	R	10*	12.5	—	6.89**	—	11.5	11.5
	—	—	13	—	—	—	10.6	10.6

* These figures do not stand for the present time, as there is no natural point in the « steady state model » from which to measure this time, but represent instead the difference — for selected sources — between the time light is emitted and the time it is received, $(T_{obs} - T_{source})/c$.

** These figures represent in the same model, not the radius of curvature — for the space in this model is flat — but the angle effective distance of these selected sources (Eq. 19).

of energy to be expected from the assumed values of H and T_{obs}; and $\rho_H = (3\ c^2/8\ \pi G)\ (H^2/c^2)$. Note that H^{-1} must exceed $3\ T_{obs}/2$ (or $2\ T_{obs}$) for the dust- (or radiation-) filled universe to be closed. The relations between H, T, A and a_o are best seen in figure 1. The last entries in the table do not refer to general relativity but to the so called steady-state model of Bondi, Gold and Hoyle. The letters designate the curves shown in figures 2, 3 and 4. All times are in units of 10^9 years, all distances in units of 10^9 light years, all angles in radians, all densities in units of 10^{-30} gm/cm³.

(2) Tolman's radiation filled universe. The equation of state in this case assumes for the pressure the maximum possible value, one third of the energy density. The metric of course still has the form (1), but now the radius of curvature and the elapsed time are connected by the equation of a circle (fig. 1) :

$$\begin{aligned} T \quad &= a_o (1 - \cos \eta)\ (\doteqdot a_o\, \eta^2/2 \text{ for small } \eta); \\ a(T) &= a_o \sin \eta\ (\eta \text{ from } 0 \text{ to } \pi)\ (\doteqdot a_o\, \eta \text{ for small } \eta). \end{aligned} \qquad (11)$$

The fractional rate of expansion, H, is given by the formula :

$$\begin{aligned} 1/H &= a/(da/dt) = (a_o/c) \sin^2 \eta / \cos \eta \\ &= (T/c) \sin^2 \eta / [\cos \eta\, (1 - \cos \eta)] \geqslant 2\ T/c. \end{aligned} \qquad (12)$$

Equation (12) was used to find from the assumed values of H and T listed in Table 1 first the parameter η and then the maximum and present radii, a_o and a. Let ρ represent the density of energy, transformed into the units of mass density. Then the product ρa^3 is not constant, but varies as $1/a$ in accordance with the principle of adiabatic invariance :

$$\begin{aligned} &\frac{\text{energy per mode}}{\text{proper frequency}} \sim \text{energy X wave length} \\ &\text{or energy X radius of container} \sim \text{constant}. \end{aligned} \qquad (13)$$

The value of the density can be expressed in several alternate forms :

$$\begin{aligned} \rho &= \text{constant}/a^4 = (3\ c^2/8\ \pi\ G)\ (a_o^2/a^4) \\ &= (3c^2/8\ \pi\ G a_o^2)\ (1/\sin^4 \eta) \\ &= (3c^2/8\ \pi\ G)\ (a^{-2} + H^2/c^2) \geqslant (3\ c^2/8\ \pi\ G)\ (H^2/c^2). \end{aligned} \qquad (14)$$

(Compare equation 6). Proceeding to kinematics, we consider as before a source separated from us by the angle χ in the spherical space. When the time parameter of the observer is η_{obs}, the time parameter of the source is calculated as in equation (8), with the result :

$$\eta_{source} = \eta_{obs} - \chi. \tag{15}$$

Formulas (9) and (10) for the red shift and for the fractional volume included within χ are unchanged and with (15) give the indicated curves in figures 2, 3 and 4.

(3) The steady-state model of Bondi, Gold and Hoyle is not compatible with the framework or philosophy of Einstein's general relativity but is included here for comparison. The metric is :

$$dl^2 = -dT^2 + e^{2\,HT/c}\,[dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\varphi^2)]\,. \tag{16}$$

The 3-space at every instant is flat. The density of matter is :

$$\rho = (3\,c^2/8\,\pi\,G)\,(H/c)^2 \tag{17}$$

Relative to an observer at the origin of coordinates, the past coordinate, r, of a source which emits at T_{source} and is detected at T_{obs} is found by putting $dl^2 = 0$ in (16) and integrating :

$$r = (c/H)\,(e^{-HT_{source}/c} - e^{-HT_{obs}/c})\,. \tag{18}$$

The angle effective distance of the source as defined in (7) is :

$$e^{HT_{source}/c}\, r\delta\theta/\delta\theta = (c/H)\left(1 - e^{-H\,(T_{obs}-T_{source})/c}\right). \tag{19}$$

The red shift is determined by the ratio :

$$\frac{\nu_{received}}{\nu_{source}} = \frac{\text{dimensions then}}{\text{dimensions now}} = e^{-H\,(T_{obs}-T_{source})/c} \tag{20}$$

The average number of galaxies per unit proper volume of space is assumed not to change with time. Therefore the number of galaxies out to the angle effective distance (19) is proportional in this model to the volume :

$$\begin{aligned} V &= \int 4\,\pi\,(\text{radius})^2\, d\,(\text{radius}) \\ &= \int_{T_{source}}^{T_{obs}} 4\,\pi\, r^2(T)\, e^{2\,HT/c}\, e^{HT/c}\, dr\,(T). \end{aligned} \tag{21}$$

More convenient for plotting is the dimensionless multiple of V :

$$\begin{aligned} f^* &\equiv H^3V/4\pi c^3 \\ &= \tau - (^3/_2) + 2e^{-\tau} - \tfrac{1}{2}e^{-2\tau} \\ &= (\tau^3/_3) - (\tau^4/_4) + (7\tau^5/_{60}) - \\ \tau &\equiv H\,(T_{obs} - T_{source})/c) \end{aligned} \tag{22}$$

Figures 2, 3 and 4 show how greatly the predictions of the steady-state model differ from those of Einstein's general relativity, and how clearly the effects of the curvature of space should show up on Einstein's theory as the observations are extended to greater red shifts and greater distances.

DENSITY OF MASS AND ENERGY
(L.T.K., Jr.; J.A.W.; R.W.)

In addition to these kinematic predictions, Einstein's theory makes an important prediction about the total density of matter-energy, as indicated by the last column of Table 1. Sandage's value of $H^{-1} = 13 \times 10^9$ years for the present coefficient of expansion leads to a minimum value $\rho_{min} = 10 \times 10^{-30}$ gm/cm^3. If in addition it be assumed that the evolution of the globular clusters has required at least $T_{obs}/c = 6 \times 10^9$ years, then ρ (dust filled universe) $= 46 \times 10^{-30}$ gm/cm^3 and ρ (radiation filled universe) $= 14 \times 10^{-30}$ gm/cm^3. In contrast, Oort's report to this conference estimates from the observable galaxies a density of the order of 0.3×10^{-30} gm/cm^3, with the possibility that this figure might have to be increased by a factor at most 10, probably considerably smaller, to allow for stars and other matter between the galaxies. At the cost of oversimplification, we will write :

$$\rho \text{ agglomerations} = (1 \times 10^{-30}\,\text{gm/cm}^3) \times 3^{\pm 1}. \tag{23}$$

The discrepancy between observation and expectation is disturbing enough to force one to recognize a tacit and very possibly unjustified assumption : that the bulk of the density of matter-energy is agglomerated. This assumption is all the more dangerous because of our ignorance — even our limited imagination — about the process which went on at the early phase when the universe was at maximum density. At that time, for example, typical accelerations produced by gravitational forces were presumably far larger than they are today. Mechanisms for the production of gravitational radiation were more effective. There is no known principle that states that the gravitational radiation given off in the earliest days carries less energy than does the matter which then came into being. No satisfactory method has even been devised to measure the flux of gravitational radiation from outer space either in toto or at any single wave length. Today we know no way to exclude the possibility that the major part of the expected density of 10 to 20 $\times$ 10^{-30} gm/cm^3 is in the form of gravitational radiation. Past history warns of the danger of distrusting the theory of Einstein when it collides with preconceived ideas. He himself tells us of his unhappiness when general relativity predicted that a universe of finite density must have a changing size; of his inventing an artificial new term, with a "cosmological constant", to compensate this "unreasonable" change in size; of the subsequent discovery that the universe really is expanding; and of his conclusion that the cosmological term ought never to have been introduced in the first place; that the implications of the simple straight forward theory ought to have been taken seriously. Now again to take seriously the predictions of general relativity, this time about density, it is not necessary for us to deduce that all the deficit of density is to be accounted for by gravitational radiation, and to resign ourselves to a wait of several generations for a check. Several other contributory sources of density exist which can be investigated sooner!

These additional sources, like gravitational radiation, but in contrast to agglomerations of matter, are all envisaged here as uniformly distributed. It might be supposed that such uniform sources would automatically be included in the mass of a galaxy out to a distance as determined from the velocity v (r) of the rotatory motion at that point :

$$\mathrm{M}\,(r) \sim \mathrm{v}^2 r/\mathrm{G}. \tag{24}$$

or as determined by similar reasoning when the stellar motions also have radial velocity components. However, as noted by H. P. Robertson in a private communication, a test particle moving in a spherical space uniformly filled with matter or energy will be deflected neither to the left nor to the right. The gravitational effects of a uniform density cancel out. Therefore it is only the mass of an agglomeration of density, over and above the uniform background density, that one expects to get from (24). Non-agglomerative contributions to the density must all be determined by their own characteristic and appropriate methods.

Neutrinos, like the only other two known radiations that travel with the speed of light — electromagnetic and gravitational waves — will be expected to be spread uniformly throughout space. The energy output of most stars in the form of neutrinos is of roughly the same order of magnitude as the electromagnetic output. However, less than one percent of the stellar mass is converted into energy in thermonuclear reactions. Therefore the neutrinos from these reactions will make a negligible contribution to the total mass-energy density of the universe. In contrast, gravitational contraction of massive stars appears capable in principle of converting unlimited amounts of mass into neutrinos and antineutrinos. Finally, no limit at all is known to the number of neutrinos which were already present from the earliest days of the universe. So much for theoretical considerations on the neutrino density.

On the experimental side no decisive success has yet been achieved in measuring the flux of neutrinos, such as come from the sun and other fusion reactions. In such reactions in the net one or more surplus bound protons experience the transformation $p + e^- \rightarrow n + \nu$. In contrast, after fission a bound neutron undergoes the change, $n \rightarrow p + e^- + \bar{\nu}$. The antineutrinos produced in this way from a Savannah River fission reactor have been detected and measured by Reines and Cowan and collaborators by a reaction endowed with two identifying time delays :

$$\begin{aligned} &\bar{\nu} \quad + \quad p \quad \rightarrow n + e^+ ; \\ &e^+(t_1) + e^- \quad \rightarrow 2\gamma ; \\ &n \ \ (t_2) + Cd^A \rightarrow Cd^{A+1} + \gamma_{rays}. \end{aligned} \tag{25}$$

With the same reaction Reines and collaborators have determined the counting rate in a deep cave in New Mexico, according to a kind personal communication of August 1957 from Dr. Reines. In this way they set an upper limit to the flux of neutrinos with energy greater than the 1.8 Mev threshold of reaction (25). All of the observed counting rate may of course well be due to cosmic ray effects which penetrate through to the cave and mock up the reaction (25). The detection cross section for a true antineutrino above the threshold is approximately proportional to energy. Therefore the observed counting rate gives an upper limit, not so much on the number flux, as on the flux of energy above 1.8 Mev :

energy flux :

$$(\bar{\nu} > 1.8 \text{ Mev}) \leqslant \text{roughly } 8 \times 10^{11} \text{ Mev/cm}^2 \text{ sec};$$

energy density :

$$(\nu > 1.8 \text{ Mev}) \leqslant \text{roughly } 30 \text{ Mev/cm}^3;$$

equivalent mass density :

$$(\bar{\nu} > 1.8 \text{ Mev}) \leqslant \text{roughly } 5 \times 10^{-26} \text{ gm/cm}^3. \qquad (26)$$

No contribution to these numbers is expected due to the flux of neutrinos from the sun ($\sim$ 1 kW/m^2, or $\sim 10^{12}$ Mev/cm^2 sec) first, because the reaction $\nu + p \longrightarrow n + e^+$ would violate the law of conservation of leptons now presumed to hold; and second, because the great majority of the neutrinos from the sun are calculated to have energies below the 1.8 Mev threshold for (25).

No one has any reason to believe that the antineutrino flux is as high as the limit (26), as stressed particularly by Reines himself. The limit is mainly important because it shows that an improvement in sensitivity by four powers of ten will bring flux determinations to the point where they will mean something for cosmology.

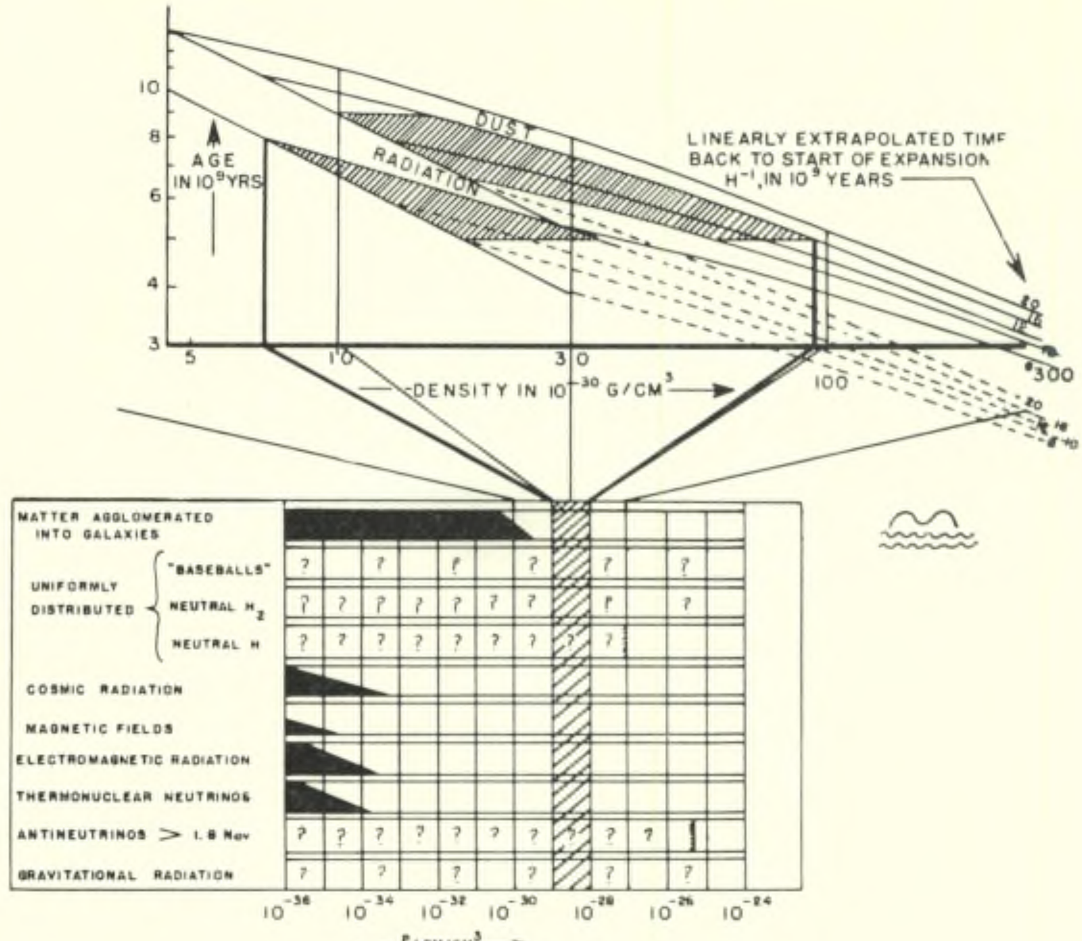
The flux of electromagnetic radiation, mainly in the visible, makes a negligible contribution to the density of matter-energy. Therefore a very rough estimate is justifiable. The solar output of $\sim 3 \times 10^{26}$ watts in a time of $\sim 5 \times 10^9$ years transports away in the form of radiation a fraction of the solar mass of the order of :

$$\frac{(3 \times 10^{33} \text{ ergs/sec})\,(1.6 \times 10^{17} \text{ sec})}{(2 \times 10^{33} \text{ gm})\,(9 \times 10^{20} \text{ ergs/gm})} \sim 3 \times 10^{-4}. \qquad (27)$$

We assume a similar output from the other agglomerated matter in the universe, with a factor of uncertainty of perhaps 10. Looking apart from any electromagnetic radiation present *de novo* in the universe, we arrive at a present average density of electromagnetic mass-energy by multiplying (27) and the ρ agglomeration of (23) :

$$\begin{aligned}\rho_{e.\,m.} &\sim (3 \times 10^{-4} \times 10^{\pm 1}) \times (1 \times 10^{-30} \text{ gm/cm}^3) \times 3^{\pm 1} \\ &\sim (3 \times 10^{-34} \text{ gm/cm}^3) \times 10^{\pm 1} \qquad (28)\end{aligned}$$

The same rough order of magnitude estimate will apply to the output of neutrinos from the same thermonuclear reactions that give the electromagnetic radiation.



AGE IN 10^9 YRS
10
8
6
4
3
5
10
30
100
300
DUST
RADIATION
LINEARLY EXTRAPOLATED TIME BACK TO START OF EXPANSION H^-1, IN 10^9 YEARS
DENSITY IN 10^-30 G/CM^3
MATTER AGGLOMERATED INTO GALAXIES
UNIFORMLY DISTRIBUTED
"BASEBALLS"
NEUTRAL H2
NEUTRAL H
COSMIC RADIATION
MAGNETIC FIELDS
ELECTROMAGNETIC RADIATION
THERMONUCLEAR NEUTRINOS
ANTINEUTRINOS > 1.8 Mev
GRAVITATIONAL RADIATION
10^-36
10^-34
10^-32
10^-30
10^-28
10^-26
10^-24
ρ I GM/CM^3

Fig. 5.

The magnetic field in the galaxy is believed on the basis of polarization effects to have a strength of 10^{-6} gauss or less. It is often assumed that this field is confined to the galaxies. Even if it pervades all space, it can make a contribution to the mass-energy density of at most :

$$\rho_{mag} < (10^{-6}\,\mathrm{gm}^{1/2}\,\mathrm{sec}^{-1}\,\mathrm{cm}^{-1/2})^2/8\,\pi\,c^2 = 4 \times 10^{-35}\,\mathrm{gm/cm}^3 \quad (29)$$

The density of cosmic ray energy near the earth is of the order of 10^{-12} erg/cm^3, or locally :

$$\rho_{cosmic\ ray} \sim 10^{-12}/9 \times 10^{20} \sim 10^{-33}\,\mathrm{gm/cm}^3 \qquad (30)$$

Again confinement to the galaxy is often assumed, but even if the radiation passes relatively freely from one galaxy to another, (30) provides an order of magnitude estimate for an upper limit to the cosmic ray mass-energy density.

About lumps of solid matter from a fraction of a kilogramme to several kilogrammes in mass spread uniformly throughout space we have practically no information, as H. N. Russell has often stressed.

We are indebted to George Field and Martin Schwarzschild for informing us that present evidence (1) places no significant limit on the abundance of neutral H_2 and (2) places no safe limit below 10^{-27} gm/cm^3 on the abundance of neutral H.

Summarizing conceivable sources of density and comparing them with the predictions of Einstein's general relativity, figure 5 shows how far one is today from being able to speak of either discrepancy or check. Probably the further investigation of the individual density components will be among the very important enterprises of the coming decade.

DYNAMICAL INSTABILITIES IN EXPANSION AND RECONTRACTION, (J.B.A.; R.M.; J.A.W.)

The assumption that the curvature of space is uniform and isotropic is of course not a fundamental part of Einstein's theory, but only an idealization which makes it convenient to apply that theory. As an alternative to the assumption of matter spread uniformly as dust, it is simple and natural to consider the idealization of matter collected at a regular lattice of points (fig. 6) whose number may be as low as 5 or as high as 720. The space around each point has approximately the static Schwarzschild character, cut off at a polygonal boundary. These pieces are fitted together to make a closed space after the fashion of the Wigner-Seitz approximation in solid state physics. The condition of smooth matchup of the

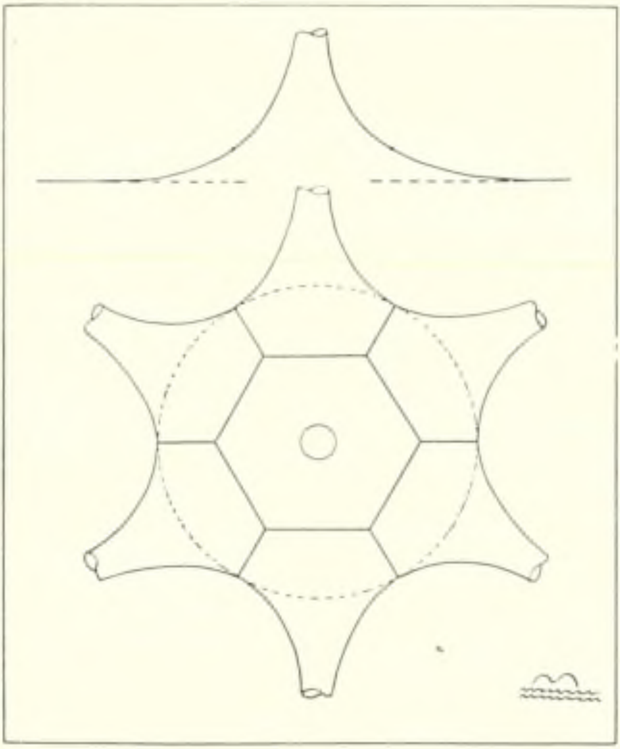


Fig. 6.

space-time metric at the interface between two Schwarzschild zones has the following simple implication : a test particle at the interface moves towards *both* the masses attracting it according to the law of free fall. In other words, the size of each zone has to alter with time. In this way one derives the law of change with time of the effective radius of the universe (3). Compared to a uniform dust-filled universe of the same total mass, the lattice universe expands to approximately the same effective radius before falling together again, and takes approximately the same time for rise and fall. The discrepancy between the two models is the smaller the larger is the number of attracting centers, and is less than 2 percent for 720 masses. Evidently the Friedmann calculation of a (T) (equation 2) makes good sense for a distribution of mass in the universe, whether uniform or not, provided that it has a high degree of symmetry.

What happens when the distribution of mass shows significant departures from symmetry ? Will these asymmetries grow with time or fade away ? This question of the stability of small disturbances in geometrodynamics recalls a similar problem in hydrodynamics.

Consider a practically infinite mass of water at rest, under pressure, but free of the action of gravity. Consider a point in the interior of this mass. At time $t = 0$ let the liquid be driven away symmetrically from this point by a sudden outward impact. The resulting cavity grows in size to a maximum radius and then collapses. The calculated curve of radius as a function of time is very similar to the curves for radius of the universe as a function of time shown in figure 1. We have a kind of mechanical analog for the universe which is easy to observe. The book of Cole (4) presents photographs of an underwater cavity at fractional second intervals. The dynamics of the cavity in this short time interval is very little affected by the pressure gradient of gravity. More important is the effect of residual gas from the explosive. Recompressed towards the end of the phase of contraction, it slows down the inward rush of liquid. The cavity vibrates, not once, but several or many times. The acceleration, directed inward most of the time, reverses direction during the short periods of high contraction. At these times the water behaves like a lake in which g is directed upwards. Departures from ideality that are almost imperceptible multiply up, according to the familar law of Rayleigh-Taylor instability :

$$\frac{d\ (\text{amplitude})}{(\text{amplitude})} = \left[\frac{\text{acceleration}}{\text{wave length}/2\,\pi}\right]^{\frac{1}{2}} dt. \qquad (31)$$

into horns and spikes (figure 7). A look at the bubble at this time suggests, not an ideal mathematical object, but a glove which is turning itself inside out one finger at a time. Does the universe likewise undergo repeated oscillations ? show instabilities ? have short lived localized regions of high curvature and high concentration of matter-energy ? have no single well defined time of maximum contraction ? If so, these past regions of high energy concentration must have important consequences for the subsequent evolution and present structure of the universe.

In a beautiful paper Lifshitz ([5]) analyzed small departures of a dust filled universe away from exact sphericity on the assumption that the universe is always *expanding*. On this basis he concluded " That arbitrary small perturbations of the gravitational field and of the distribution of matter in expanding universe either decrease with time or increase so slowly that they cannot serve as centers of formation of separate nebulae or stars ". However, here we are interested in a different question : the formation of condensations, not in the present phase of expansion of the universe, but *during the stage of contraction towards the end of the previous oscillation.* The equations of Lifshitz serve perfectly well to give a new answer to this new question, at least for the idealized case of a universe that derives all of its mass-energy from dust.

In the notation of Lifshitz $g_{\mu\nu}$ is the unperturbed four dimensional metric of equation (1*b*) (Greek indices 0, 1, 2, 3), γ_{ik} (Latin indices 1, 2, 3) is the 3-space metric :

$$d\chi^2 + \sin^2\chi\,(d\theta^2 + \sin^2\theta\,d\varphi^2), \tag{32}$$

and $g_{\mu\nu} \rightarrow g_{\mu\nu} + h_{\mu\nu}$ describes the perturbation. The unperturbed and perturbed system differ also in density ($\rho \rightarrow \rho + \delta\rho$) and in the four velocity of matter ($u^\sigma \rightarrow u^\sigma + \delta u^\sigma$). The coordinates used to describe the perturbed system are so chosen by Lifshitz as to make :

$$h_{oo} = h_{ok} = 0. \tag{33}$$

Thus the quantity :

$$h = h_\mu^\mu = h_k^k. \tag{34}$$

depends only upon the space-space tensor $h_a{}^b$. The components of this particular tensor — a mixed tensor — are taken by definition to be identical with the corresponding components of the mixed four tensor, $h_\mu{}^\nu$, but any subsequently raising or lowering of indices or covariant differentiation is understood to be performed with respect to the metric γ_{ik} of (32) and its reciprocal γ^{hl}. With this understanding the changes h_j^k in the metric satisfy to the first order of small quantities the linear differential equations $(a \neq b)$:

$$h_{a;k}^{k;b} + h_{k;a}^{b;k} - h_{;a}^{;b} - h_{a;k}^{b;k} + (h_a^b)'' + 2\,(a'/a)\,(h_a^b)' + 2\,h_a^b = 0, \tag{35}$$

$$h'' + 2\,(a'/a)\,h' - h + \tfrac{1}{2}\,(h_{k;j}^{j;k} - h_{;j}^{;j}) = 0, \tag{36}$$

where primes denote differentiation with respect to the time parameter η.

The general perturbation satisfying (35) and (36) is decomposable into tensorial hyperspherical harmonics of various orders, n, on the hypersurface $(\chi,\ \theta,\ \varphi)$ of the unit sphere in 4-dimensional Euclidean space. The harmonics of each order n can be analyzed in and for themselves. Each is expressed as the product of a standard function of the space coordinates (χ, θ, φ), multiplied by an amplitude factor for which equations (35) and (36) give a simple ordinary differential equation. The solutions of these differential equations, written down by Lifshitz only for the case of an open universe, are modified for the problem of interest here by his simple transformation $\eta \rightarrow i\eta$, supplemented by several changes in constants of integration.

The first order perturbations in the metrics of any given harmonic order n, with subsidiary index numbers l and m, fall into four classes :

I, a solution in which $\delta\rho/\rho$ is *even* with respect to inversion of time about the moment, $\eta = \pi$, of maximum expansion :

$$h_a^b = (Q_{;a}^{;b} + \delta_a^b\,Q)\,(1 - \tfrac{1}{2}\,\eta \cot \tfrac{1}{2}\eta)/\sin^2 \tfrac{1}{2}\eta; \tag{37}$$

Q = scalar harmonic with index numbers $n,\ l,\ m$ satisfying :

$$Q_{;a}^{;a} + (n^2 - 1)\,Q = 0; \tag{38}$$

$$\delta\rho/\rho = (1/6)\,[1 + (a'/a)^2]^{-1}\,[h_{b;a}^{a;b} - h_{;a}^{;a} + (2\,a'/a)\,h' - 2\,h] \tag{39}$$

$$= \{-(2/3) + [2 + (\pi - \eta)\cot \tfrac{1}{2}\eta]/\sin^2 \tfrac{1}{2}\eta\}\,Q\,(\chi, \theta, \varphi). \tag{40}$$

II, a solution in which $\delta\rho/\rho$ is *odd* with respect to inversion about $\eta = \pi$:

$$h_a^b = (n^2 - 1)^{-1}\,Q_{;a}^{;b}\,(\tfrac{1}{2} \cot \tfrac{1}{2}\eta + (1/6)\cot^3 \tfrac{1}{2}\eta)\,; \tag{41}$$

$$\begin{aligned}\delta\rho/\rho &= (1/6)\,[1 + (a'/a)^2]^{-1}\,[h_{b;a}^{a;b} - h_{;a}^{;a} + (2\,a'/a)\,h' - 2\,h] \\ &= Q\,(\chi, \theta, \varphi)\cos \tfrac{1}{2}\eta/\sin^3 \tfrac{1}{2}\eta\,;\end{aligned} \tag{42}$$

III, a *rotational* perturbation in which $\delta\rho/\rho$ does not change but the matter is set into a circulatory motion :

$$h_a^b = (\tfrac{1}{2}\cot\tfrac{1}{2}\eta + (1/6)\cot^3\tfrac{1}{2}\eta)\, S_a^b(\chi, \theta, \varphi)\ ; \tag{43}$$

$$\frac{(\text{velocity})^b}{\text{velocity of light}} = S^b(\chi, \theta, \varphi)/24\sin^2\tfrac{1}{2}\eta\ ; \tag{44}$$

$$S_a^b = S_a^{:b} + S^b_{:a}\ , \tag{45}$$

where S^b = harmonic vector field with index numbers n, l, m satisfying :

$$S_{b:g}^{\ \ :g} + (n^2-2)\, S_b = 0;\ S^a_{:a} = 0\ ; \tag{46}$$

and IV, a *gravitational wave* which affects neither the density nor the velocity :

$$h_a^b = G_a^b(\sin\tfrac{1}{2}\eta)^{-1}(d/d\eta)\,[(C_1\sin n\eta + C_2\cos n\eta)/\sin\tfrac{1}{2}\eta] \tag{47}$$

where $G_a{}^b$ = harmonic tensor field with index numbers n, l, m satisfying :

$$\begin{aligned} &G_{a:g}^{b:g} + (n^2-3)\, G_a^b = 0\ ; \\ &G_{ab} = G_{ba}\ ;\ \ G^b_{a:b} = 0\,. \end{aligned} \tag{48}$$

For times separated from moments of collapse by much more than the period of vibration of such a gravitational wave, the disturbances $h_a{}^b$ have amplitudes proportional to $1/\sin^2\frac{1}{2}\eta$; that is, inversely proportional to the radius of the universe, a. The derivatives of the $h_a{}^b$ can be found in order of magnitude by dividing by the reduced wave length, $\lambda/2\pi \sim a/n$, and are proportional to $1/a^2$. Consequently the density of gravitational energy goes as $1/a^4$, just as does the density of electromagnetic energy, and as is to be expected for free radiaton of any type in a spherical universe of radius a.

The time dependence of the first order disturbances in the metric, approximately independent of the order n for gravitational waves in the sense just described, is accurately independent of n for disturbances of types I, II and III. In this regard the stability of the oscillating universe is very different from the stability of the oscillating undersea bubble. The fractional growth of a hydrodynamic disturbances of order n in the time dt, according to (31), is proportional to $n^{\frac{1}{2}}\, dt$. The higher the harmonic, or the shorter the scale of the disturbance, the more rapidly it grows — hence horns and spikes such as typified in figure 7. However, any disturbance in the *metric* of a given type (I, II or III) multiples up with time (fig. 8) without change of form, regardless of its scale and shape, so long as the perturbations in curvature are small compared to the total curvature. It may still give a reasonable first impression of what happens in the stage of high contraction to speak of a glove turning itself inside out one finger at a time. In contrast to the hydrodynamic case where the fingers are sharp and innumerable, here the number and angular extent of fingers bear a reasonably simple relation to small perturbations present in the metric during the era of wide expansion.

Why disturbances of all orders n, of a given type (I, II or III) multiply up at a rate independent of n we have not been able to see in simple intuitive terms except for perturbations of type II. Consider some of the dust, distributed in space with a density $Q(\chi, \theta, \varphi)$, running through its cycle of expansion and contraction at a cotime $\delta T = \delta T_o\, Q(\chi, \theta, \varphi)$ before the rest of a uniform distribution of matter does so. This advance in cotime is connected with an advance in the time parameter η by the equation :

$$\delta T = \tfrac{1}{2} a_o \delta (\eta - \sin \eta) = a_o \sin^2 \tfrac{1}{2}\eta\, \delta\eta. \tag{49}$$

The original density varies with time (equation 6) as the inverse sixth power of $\sin \frac{1}{2}\eta$. Therefore the out-of-step part of the dust contributes a fractional change in density :

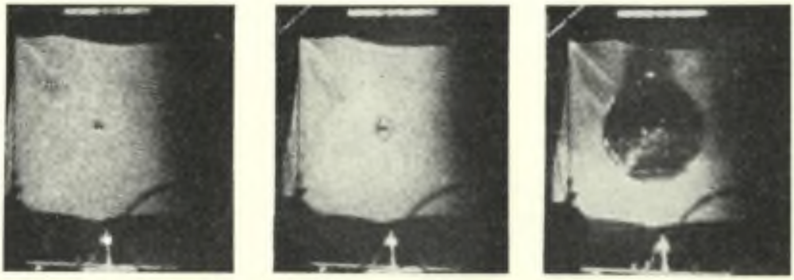
$$\begin{aligned} \delta\rho/\rho &= -6\, \delta(\sin \tfrac{1}{2}\eta)/\sin \tfrac{1}{2}\eta \\ &= -[3\, Q(\chi, \theta, \varphi)\, \delta T_o/a_o] \cos \tfrac{1}{2}\eta/\sin^3 \tfrac{1}{2}\eta. \end{aligned} \tag{50}$$

Except for an obviously free amplitude factor, this result agrees with the general expression (42) for a change in density that is odd with respect to time inversion.

The analogy is very familiar in other connections between the dynamics of a universe of dust and the Newtonian dynamics of a cloud of dust particles endowed with purely radial motions and moving in a preexisting flat space. Provided that the initial outward velocity is low enough so that the cloud will reach a maximum size and fall back together again, the energy equation for the Newtonian problem is close in form to equation (4). Now let some of the particles in the Newtonian analog be given their initial outward velocities a little sooner than standard. Then on recontraction these will implode a little earlier than the others. They will start a new cycle of expansion in some direction while in other directions the contraction is continuing. There will exist no well defined moment for the system as a whole at which contraction ceases and expansion begins.

It seems as impossible in the actual universe as in the Newtonian model to conclude from observations on one phase of expansion that that is the first pulsation that ever took place and the last which will ever occur. The contrary conclusion would seem more natural in the absence of further evidence. In this case it is reasonable to believe (1) that there is no well defined instant at which the present expansion began and (2) that irregularities in the present distribution of matter in the universe ought in part to be correlated with variations in the time at which different parts of the system went through the phase of maximum contraction.

To investigate further the details of implosion would lead into unexplored areas of geometrodynamics. They recall problems in hydrodynamics like cavitation, breaking surf, shock waves and turbulence [6]. However, two points push themselves forward with special insistence. First, the implosion — varied in timing as it may be at different points in space — would seem in any case to



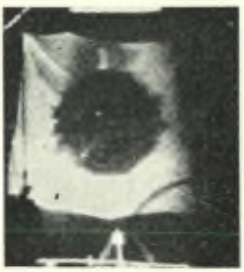
0 1 5.7

Fig. 7.

10.5

15.3

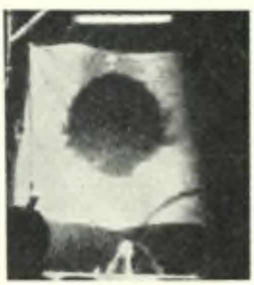
20



24.8

29.6

34.3



39

43.8

48.5

53.2

58

62.2

lead to conditions and densities of matter-energy outside the range of present experience. Second, the gross dynamics of expansion and recontraction will be governed more by the gravitational pull of radiation ($\rho \sim 1/a^4$) than by the attraction of inert matter ($\rho \sim 1/a^3$), certainly during periods of contraction, and perhaps also now and always (fig. 5). On this account Lifshitz's investigation of instabilities needs to be extended from the case of a dust filled universe to the opposite idealized case where pressure and gravitational pull derive entirely from radiation. His paper does consider the case of pure isotropic radiation, describable by a *pressure* ($p = \rho/3$). This idealization is appropriate when the mean free path for the radiation is small compared to the scale of the disturbances in metric and in density under consideration. However, under present conditions the mean free path for all three kinds of radiation — electromagnetic, gravitational and neutrinos — is enormous compared to the distance between galaxies. Therefore, the radiation retains a memory of the place from which it comes which demands a different kind of calculation and a more detailed type of statistical bookkeeping than is provided by the concepts of diffusion and pressure. This more detailed analysis begins with : (1) the world line of one photon as affected by the metric, goes next to (2) the concept of photon density, then to (3) the stress-energy tensor due to this density of photons, and finally in (4) — Einstein's field equations for the metric in terms of the stress-energy tensor — has the means to close the circle of coupled equations. An analysis of this kind has not yet been made; only steps (1) and (2) can be reported here.

The world line of the typical photon can be written in the form :

$$x^\mu = x^\mu(\lambda;\ e, f, g;\ h, m, n) \tag{51}$$

where the parameter λ picks out different points on the world line of the one photon, and the six parameters e, ..., n tell which photon is under consideration. The world line satisfies the familiar equations of a null geodesic :

$$g_{\alpha\beta}\,(dx^\alpha/d\lambda)\,(dx^\beta/d\lambda) = 0. \tag{52}$$

$$(d^2x^\mu/d\lambda^2) + \Gamma^\mu_{\alpha\beta}\,(dx^\alpha/d\lambda)\,(dx^\beta/d\lambda) = 0. \tag{53}$$

The content of these equations can be expressed more conveniently and usefully [2,6] by going to the phase function φ :

$$\varphi = \varphi\,(x^0, x^1, x^2, x^3;\ e, f, g), \tag{54}$$

satisfying the Hamilton-Jacobi equation :

$$g^{\alpha\beta}(\partial\varphi/\partial x^{\alpha})(\partial\varphi/\partial x^{\beta}) = 0. \tag{55}$$

The quantity $k_{\alpha} = \partial\varphi/\partial x^{\alpha}$ can be considered to represent the 4-vector wave number of the photon.

The three *scalar* — and therefore invariant — "conditions of constructive interference" [6] :

$$\partial\varphi/\partial e = h, \quad \partial\varphi/\partial f = m, \quad \partial\varphi/\partial g = n, \tag{56}$$

between waves of slightly different wave numbers are enough to give x^1, x^2, x^3 in terms of x^0 for fixed e, f, g; h, m, n, in agreement with (51), (52) and (53).

The 64 photons designated by the $64 = 2^6$ different choices of parameters :

$$\begin{array}{llllll}
e, & f, & g; & h, & m, & n; \\
e+\delta e, & f, & g; & h, & m, & n; \\
e, & f+\delta f, & g; & h, & m, & n; \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
e+\delta e, & f+\delta f, & g+\delta g; & h+\delta h, & m+\delta m, & n+\delta n,
\end{array} \tag{57}$$

span at any given time x^0 in any arbitrary system of coordinates x^{μ} a six-dimensional region in phase space of volume :

$$\begin{aligned}
dx^1\,dx^2\,dx^3\,dk_1\,dk_2\,dk_3 &= \left[\frac{\partial(x^1, x^2, x^3;\ k_1, k_2, k_3)}{\partial(h, m, n;\ e, f, g)}\right]_{\text{fixed } x^0} de \ldots\ldots dn \\
&= \left[\frac{\dfrac{\partial(x^1, x^2, x^3;\ \partial\varphi/\partial x^1, \partial\varphi/\partial x^2, \partial\varphi/\partial x^3)}{\partial(x^1, x^2, x^3;\ e, f, g)}}{\dfrac{\partial(h, m, n;\ e, f, g)}{\partial(x^1, x^2, x^3;\ e, f, g)}}\right]_{\text{fixed } x^0} de \ldots\ldots dn \\
&= \left[\frac{\begin{array}{ll}
1 \ldots\ldots\ldots\ldots & 0 \ldots\ldots\ldots\ldots \\
0 \ldots\ldots\ldots\ldots & 0 \ldots\ldots\ldots\ldots \\
0 \ldots\ldots\ldots\ldots & 0 \ldots\ldots\ldots\ldots \\
\partial^2\varphi/\partial x^1\partial x^1 \ldots\ldots & \partial^2\varphi/\partial x^1\partial e \ldots\ldots \\
\partial^2\varphi/\partial x^2\partial x^1 \ldots\ldots & \partial^2\varphi/\partial x^2\partial e \ldots\ldots \\
\partial^2\varphi/\partial x^3\partial x^1 \ldots\ldots & \partial^2\varphi/\partial x^3\partial e \ldots\ldots
\end{array}}{\begin{array}{ll}
1 \ldots\ldots\ldots\ldots & 0 \ldots\ldots\ldots\ldots \\
0 \ldots\ldots\ldots\ldots & 0 \ldots\ldots\ldots\ldots \\
0 \ldots\ldots\ldots\ldots & 0 \ldots\ldots\ldots\ldots \\
\partial^2\varphi/\partial e\partial e \ldots\ldots & \partial^2\varphi/\partial e\partial x^1 \ldots\ldots \\
\partial^2\varphi/\partial f\partial e \ldots\ldots & \partial^2\varphi/\partial f\partial x^1 \ldots\ldots \\
\partial^2\varphi/\partial g\partial e \ldots\ldots & \partial^2\varphi/\partial g\partial x^1 \ldots\ldots
\end{array}}\right] de \ldots\ldots dn \\
&= de \ldots\ldots dn.
\end{aligned} \tag{58}$$

This 6-volume evidently does not change with time, a type of result familiar from Liouville's theorem. The number of photons in this 6-volume, divided by its cubical content, defines an invariant dimensionless scalar density (7), N, that remains constant in time, *along the world line of the photons in question* (that is, for fixed $e, f, g; h, m, n$) however much it changes with position in x- and k-space :

$$
\begin{aligned}
& d\,(\text{number of photons}) \\
& = \mathrm{N}\,(e, f, g;\ h, m, n)\, de\, df\, dg\, dh\, dm\, dn \\
& = \mathrm{N}\,(x^0, x^1, x^2, x^3;\ k_1, k_2, k_3)\, dx^1\, dx^2\, dx^3\, dk_1\, dk_2\, dk_3; \qquad (59)
\end{aligned}
$$

$\mathrm{N}\,(e, f, g;\ h, m, n)$ independent of x^0 or λ;
$\mathrm{N}\,(x^\mu;\ k_s)$ dependent on x^0, but invariant at a given point with respect to change of coordinate system. (60)

The response of N *to* perturbations in the metric away from the Tolman value (11), and the influence of N *on* the metric have to be investigated before one can speak in more detail about the instabilities which develop during the phase of contraction of the universe.

MATTER-ENERGY AT HIGH DENSITY; END POINT OF THERMONUCLEAR EVOLUTION (K.H.; M.W.; J.A.W.)

In seeking the consequences of Einstein's theory for the structure and evolution of the universe we have been forced to consider what happens during contraction. Such implosion can be expected to lead not only to the dynamic instabilities just discussed, but also to unprecedentedly high densities of matter and radiation. Such densities pose unsolved problems to general relativity and elementary particle physics. These problems : (1) are significant for the dynamics of the universe; (2) have a bearing on stellar evolution and (3) are also important in their own right. A general survey of the physics of very high densities being beyond our reach today, we shall consider only two highly idealized cases of highly compressed mass-energy — pure radiation and pure-matter — and not try at this premature stage to mix them or fit them into the dynamics of the implosion process.

Radiation begins to behave in unfamiliar ways when its mass-energy density ρ is great enough to curve space significantly or great enough to produce electron pairs. The one effect begins in a region of dimension L when the metric in the region is changed by an amount of the order $\delta g \sim 1$:

$$G \rho L^3/c^2 L \sim 1; \tag{61}$$

the other begins when :

$$\rho \sim m/(\hbar/mc)^3 \sim 10^4 \text{ g/cm}^3; \tag{62}$$

and both effects begin (Table 2) at the same density when L is of the order :

$$\begin{aligned} L &\sim (\hbar/mc)^2/(\hbar G/c^3)^{1/2} \\ &\sim (4 \times 10^{-11} \text{ cm})^2/(1.6 \times 10^{-33} \text{ cm}) \sim 10^{12} \text{ cm}; \end{aligned} \tag{63}$$

a few times the radius of the sun, and when the mass is of the order :

$$M \sim 10^{40} \text{ gm}. \tag{64}$$

TABLE II.

Characterization of clouds of isotopic electromagnetic radiation via Equations (61) and (62).

Mass	Density	Dimensions	Characteristics
10^{39} g = 5×10^5 $M_\odot$ (as example of a mass *less* than 10^{40} g)	0 to 10^4 g/cm^3 10^4 to 10^5 g/cm^3 10^6 g/cm^3	∞ to 5×10^{11} cm 5×10^{11} to 2×10^{11} cm 10^{11} cm	No pairs, $\delta g_{\mu\nu} < 1$ Pairs, $\delta g_{\mu\nu} < 1$ Pairs, $\delta g_{\mu\nu} \sim 1$
10^{41} g = 5×10^7 $M_\odot$ (as example of a mass *greater* than 10^{40} g)	0 to 10^2 g/cm^3 10^2 g/cm^3		No pairs, $\delta g_{\mu\nu} < 1$ No pairs, $\delta g_{\mu\nu} \sim 1$ (geon)

The last entry in Table II recalls the special circumstances where a body of pure radiation is able to hold itself together by its own gravitational attraction into a longlived object, a gravitational-electromagnetic entity or " geon " ([8]).

Such an object acquires additional stability when the radiation gathers together in the form of a circular torus, with half the photons circulating around in each of the two opposed directions, owing to the enhanced gravitational attraction between oppositely directed pencils of radiation. In analysis to date such a geon has always been assumed to be surrounded by a space that is asymptotically flat, or at least not strongly curved. In contrast, a metric that changes strongly in space and time will be more reasonably expected during late stages of contraction of the universe. New idealized situations — such as deformed geons and patterns of electromagnetic and gravitational fields of special symmetry — require investigation before one can easily gain enough insight to see what will happen to pure radiation under asymmetric implosion.

Pure matter and its equilibrium under a gravitational field has been a matter of investigation since early times. Schwarzschild considered a homogeneous incompressible fluid of proper density ρ in static stable equilibrium under the action of its own gravitational field. He chose the system of coordinates in which the metric has the form :

$$dl^2 = -e^{\nu(r)}\,dT^2 + e^{\lambda(r)}\,dr^2 + r^2\,(d\theta^2 + \sin^2\theta\,d\varphi^2). \qquad (65)$$

He found that there is a maximum radius and a maximum mass which cannot be exceeded for any given density. The mass is :

$$M = (4\,\pi\,a^3/3)\,\rho \qquad (66)$$

and the metric factor $e^{\lambda(r)}$ is given by the formula :

$$e^{-\lambda(r)} = \begin{cases} 1 - (8\,\pi G\rho/3\,c^2)\,r^2 = 1 - r^2/R^2 & \text{inside} \\ 1 - a^3/R^2\,r & \text{outside} \\ 1 - a^2/R^2 & \text{at surface.} \end{cases} \qquad (67)$$

Here $2\,\pi\,a$ is the circumference of the sphere and R^2 is an abbreviation for $3\,c^2/8\,\pi G\rho$. The space part of the metric cannot remain regular for all r, according to (67), unless the outer coordinate of the sphere is less than the critical value R :

$$\begin{aligned} a &< R = (3\,c^2/8\,\pi G\rho)^{1/2}, \\ M &< M_{crit} = (4\,\pi/3)\,(3\,c^2/8\,\pi G\rho)^{3/2} = (4\,\pi/3)\,R^3. \end{aligned} \qquad (68)$$

It is often said that this result shows that one cannot bring together in a limited region of space an unlimited amount of mass.

It is necessary to distinguish between amount of mass and amount of matter, however; or between the gravitational attraction of the sphere on a remote test particle, and the number of nucleons in the sphere. Equation (66) for the mass looks obvious but actually has a subtle interpretation : The volume of the sphere is not $(4\pi a^3/3)$ but :

$$
\begin{aligned}
V &= \int (r\,d\theta)(r \sin\theta\, d\varphi)(e^{\frac{1}{2}\lambda}\, dr) \\
&= 4\pi \int_o^a (1 - r^2/R^2)^{-\frac{1}{2}} r^2\, dr \\
&= 2\pi R^3 [\arcsin (a/R) - (a/R)(1 - a^2/R^2)^{\frac{1}{2}}] \\
&= (4\pi a^3/3)[1 + (3/10)(a^2/R^2) + (9/56)(a^4/R^4) + \ldots]\,. \qquad (69)
\end{aligned}
$$

The proper mass of the sphere — the sum of the masses of the individual pieces before assembly under gravitational attraction — is :

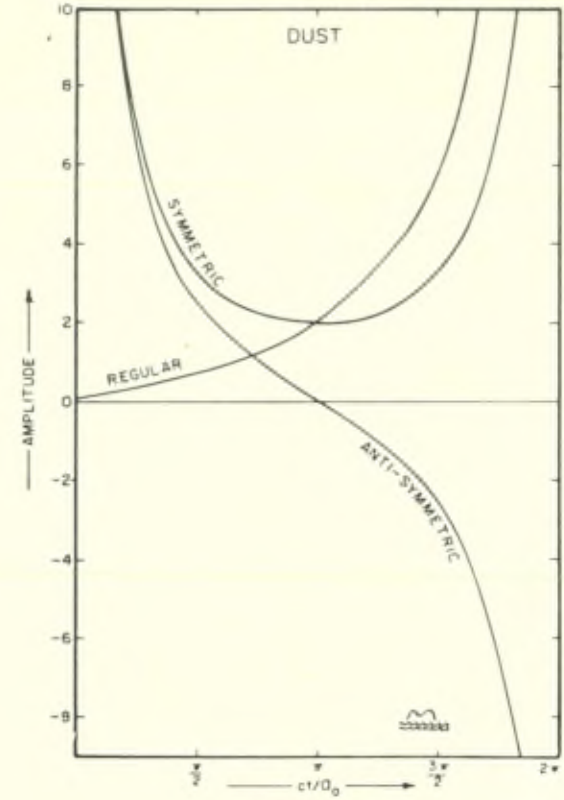
$$M_{original} = V\rho = M\{1 + (3/5)(GM/c^2a)[1 + (15/14)(GM/c^2a) + \ldots]\} \qquad (70)$$

Of course the observable mass — the mass after assembly — is less by an amount which represents in mass units the gravitational potential energy of the system :

$$
\begin{aligned}
\delta M_{grav} = M_{orig} - M &= (3/5)(GM^2/c^2a)[1 + (15/14)(GM/c^2a) + \ldots] \\
&= (3/5)(GM_{orig}{}^2/c^2a)[1 - (9/70)(GM_{orig}/c^2a) + \ldots]. \qquad (71)
\end{aligned}
$$

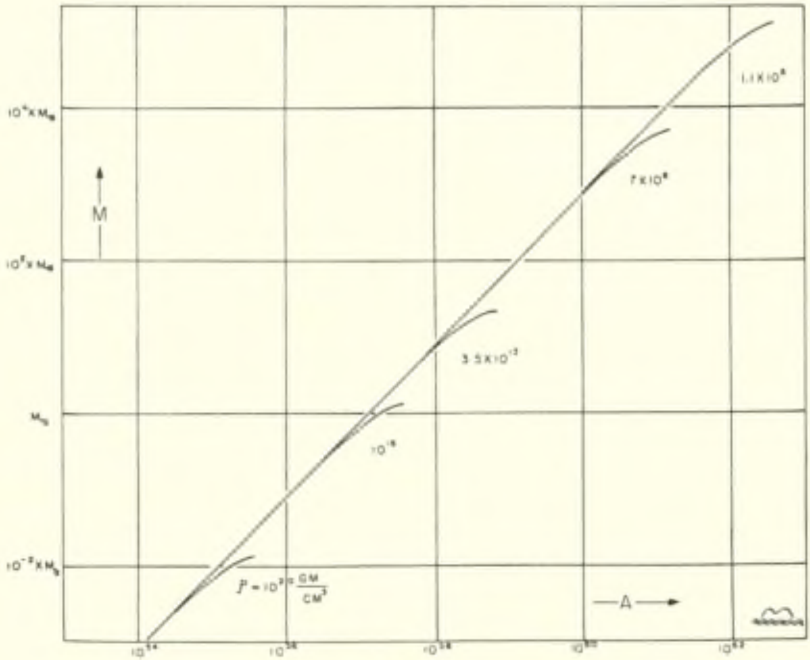
Here the first term, after multiplication with $-c^2$, agrees with the familiar Newtonian expression for the gravitational potential energy of a uniform sphere.

Starting with a fluid of a given density, and a mass a little less than the critical mass, what will one observe as he lets fall on the system drop after additional drop of the same fluid, removing at each stage the heat energy derived from the gravitational work done? For each gram added, a fraction will go towards increasing the mass of the sphere and a fraction will be radiated away as heat of combination. The radiated fraction will approach one as the mass of the sphere approaches the critical mass, and very little will go into increasing the mass of the system. It is therefore tempting to guess (1) that the amount of proper mass $M_{original}$ or the number of nucleons,



DUST
SYMMETRIC
REGULAR
ANTI-SYMMETRIC
AMPLITUDE
10
8
6
4
2
0
-2
-4
-6
-8
$\frac{\pi}{2}$
π
$\frac{3\pi}{2}$
2π
ct/a_0

Fig. 8.



M
A

Fig. 9.

can be increased beyond all limit, and (2) that the net mass of the system, M, approaches the critical mass only asymptotically :

$$M \longrightarrow M_{crit} \quad \text{as} \quad M_{original} \longrightarrow \infty \qquad (72)$$

This outcome would make it unnecessary ever to worry about what will happen when one adds the final critical drop of fluid. The number of nucleons could be infinite and still leave the total mass finite. But *this guess is wrong.*

Figure 9 shows net mass as a function of original mass, or number of nucleons, for selected but fixed values of the total density. The curves plotted there on the basis of equations (66) and (69) show not only (1) that the final net mass is limited to a value less than M_{crit} but also (2) *the allowable original mass before assembly*, M_{orig}, *is also limited* :

$$M_{orig} < M_{crit} \left\{(3/2)\,[\arcsin(a/R) - (a/R)(1 - a^2/R^2)^{\frac{1}{2}}]\right\}_{a = R}$$
$$= (3\,\pi/4)\,M_{crit} = 2.3562\,M_{crit} \qquad (72)$$

The increment in final net mass per unit increase in the mass assembled has the value :

$$dM/dM_{orig} = [1 - a^2/R^2]^{\frac{1}{2}} = [1 - (M/M_{crit})^{2/3}]^{\frac{1}{2}}. \qquad (73)$$

This quantity, the fraction of the added mass not radiated away, approaches zero at the upper limit as expected. But even *this limit cannot be attained.*

If the limit $M = M_{crit}$ could be attained, one could operate in principle an idealized machine to manufacture matter at no cost in energy as follows : A gram of matter added to the sphere will give off its entire mass as radiation. A gram of antimatter added to a similar but remote critical sphere of antimatter will give off its entire mass as radiation. The radiant energy from both sources can be caught and used to create in the empty space far from the spheres another gram of matter, and another of antimatter, to be dropped on the two spheres to keep the process going.

The reason why this idealized process of creating matter costs no energy is easy to see. The normal extremal energy for a positive (or negative) energy particle is $+\,mc^2$ (or $-\,mc^2$). However, in a static gravitational field these limits are changed to :

$$E_{\pm} = \pm\, mc^2 (-g_{00})^{\frac{1}{2}} = \pm\, mc^2 e^{\frac{1}{2}\nu}. \qquad (74)$$

or outside the sphere :

$$E_{\pm} = \pm mc^2 [1 - a^3/R^2 r]^{\frac{1}{2}}. \tag{75}$$

and at its surface :

$$\begin{aligned} E_{\pm} &= \pm mc^2 [1 - a^2/R^2]^{\frac{1}{2}} \\ &= \pm mc^2 [1 - (M/M_{crit})^{2/3}]^{\frac{1}{2}}. \end{aligned} \tag{76}$$

Evidently it would cost no energy to create a pair at the surface of the sphere if its mass had the critical value M_{crit}. The positive and negative energy seas would have no more separation than they do for neutrinos — zero separation.

This critical condition of zero separation of the two seas of electrons is already attained, at the *center* of the sphere, however, when the mass M is 16 percent short of M_{crit}. A value of M greater than 0.8382 M_{crit} leads to a singular and physically unacceptable behavior of the metric, in the sense that $(-g_{oo})^{\frac{1}{2}} = e^{\frac{1}{2}\nu}$ becomes negative for some value of r. The maximum sphere is that for which $e^{\frac{1}{2}\nu}$ goes to zero just at the origin. In comparison to the Newtonian values :

$$e^{\frac{1}{2}\nu} \doteq \begin{cases} 1 - (3\,a^2/4\,R^2) + (r^2/4\,R^2) \text{ for } r < a \\ 1 - (a^3/2\,R^2 r) \text{ for } r \geqslant a, \end{cases} \tag{77}$$

the accurate values found by solving Einstein's field equations are :

$$\frac{E_+ - E_-}{2\,mc^2} = e^{\frac{1}{2}\nu} = \begin{cases} (3/2)(1 - a^2/R^2)^{\frac{1}{2}} - (1/2)(1 - r^2/R^2)^{\frac{1}{2}} \text{ for } r < a; \\ (1 - a^3/R^2 r)^{\frac{1}{2}} \text{ for } r \geqslant a. \end{cases} \tag{78}$$

This expression for the separation of positive and negative energy states just goes to zero at the origin when $a^2 = a^2_{max} = (8/9)R^2$. The *maximum sphere* in this sense has these properties (figure 10) :

$$\begin{aligned} a_{max} &= (8/9)^{\frac{1}{2}} R = 0.9427\, R; \\ M_{max} &= (8/9)^{3/2} M_{crit} = 0.8382\, M_{crit}\,; \\ M_{max\ orig} &= (3/2)\,[(\pi/2) - \arcsin(1/3) - (8^{\frac{1}{2}}/9)]\, M_{crit} \\ &= 1.3750\, M_{crit}\,; \\ dM/dM_{orig} &= 1/3\,; \\ e^{\frac{1}{2}\nu} &= \begin{cases} \frac{1}{2}[1 - (1 - r^2/R^2)^{\frac{1}{2}}] & \text{for } r < a \\ (1/4)(r^2/R^2) + \ldots\ldots & \text{for } r < a \\ (1/3) & \text{for } r = a \\ [1 - 8^{3/2} R/9^{3/2} r]^{\frac{1}{2}} & \text{for } r \geqslant a. \end{cases} \end{aligned} \tag{79}$$

The general expression for the pressure in the interior of the Schwarzschild sphere, in dimensionless units,

$$(8\pi GR^2/c^4)\,p = 3\,p/\rho c^2 = \frac{3(1-r^2/R^2)^{\frac{1}{2}} - 3(1-a^2/R^2)^{\frac{1}{2}}}{3(1-a^2/R^2)^{\frac{1}{2}} - (1-r^2/R^2)^{\frac{1}{2}}}, \quad (80)$$

reduces for the maximum sphere to the form :

$$3\,p/\rho c^2 = \frac{3(1-r^2/R^2)^{\frac{1}{2}} - 1}{1-(1-r^2/R^2)^{\frac{1}{2}}}$$
$$\doteq 4\,R^2/r^2 \text{ for } r \ll R, \quad (81)$$

with an unacceptable singularity at the origin. If, however, the radius is less than the critical radius by a non-zero fraction ε, no matter how small, then the pressure is everywhere perfectly finite :

$$a = a_{max}(1-\varepsilon)\,;$$
$$3\,p/\rho c^2 = 4/[(4\,\varepsilon^{\frac{1}{2}})^2 + (r/R)^2]\,(\text{for } r < a)\,. \quad (82)$$

Therefore configurations of all masses and radii up to limits just short of M_{max} and a_{max} are perfectly acceptable from a physical point of view.

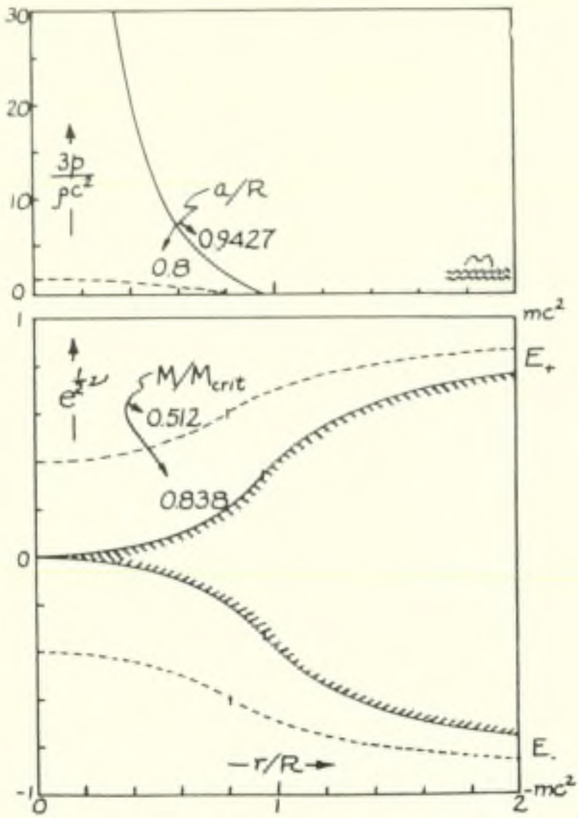
The Schwarzschild idealization of a uniform incompressible fluid provides a most helpful starting point for discussing the problem of gravitational binding. However, it is plainly incompatible even with special relativity to assume a truly incompressible fluid, for such a medium would propagate signals with a velocity greater than the speed of light ([9]). Compressibility must therefore be assumed. But as soon as the fluid begins to compress, the density goes up and the allowable critical mass goes down (figure 9). Therefore the situation is critically dependent upon the equation of state of the fluid.

Landau ([10]) has considered in general terms the case where the pressure that sustains a cold star derives from an ideal Fermi gas :

(1) pressure due either to electrons or to nucleons;

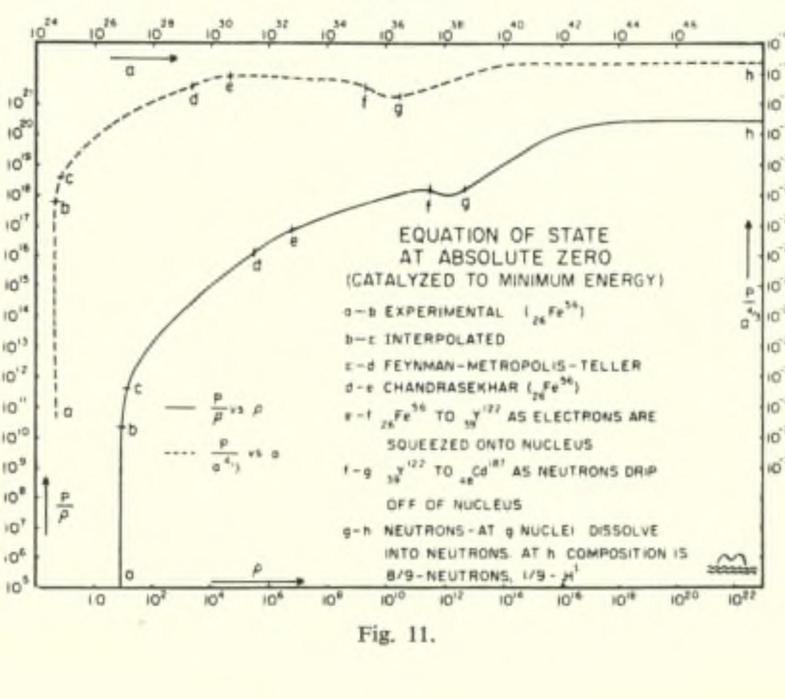
(2) ideal Fermi particles; no correction for interactions between particles;

(3) limiting case where the A particles in the sphere of radius a are squeezed so tightly that the linear dimensions of the region of confinement of one particle, $\sim a/A^{1/3}$, are smaller than the Compton wave length of the particle; or the energy of the particle is relati-



a/R
0.9427
0.8
M/Mcrit
0.512
0.838
mc²
E+
E-
-mc²
r/R

Fig. 10.



EQUATION OF STATE
AT ABSOLUTE ZERO
(CATALYZED TO MINIMUM ENERGY)
a–b EXPERIMENTAL (26Fe56)
b–c INTERPOLATED
c–d FEYNMAN-METROPOLIS-TELLER
d–e CHANDRASEKHAR (26Fe56)
e–f 26Fe56 TO 39Y122 AS ELECTRONS ARE
SQUEEZED ONTO NUCLEUS
f–g 39Y122 TO 48Cd187 AS NEUTRONS DRIP
OFF OF NUCLEUS
g–h NEUTRONS – AT g NUCLEI DISSOLVE
INTO NEUTRONS. AT h COMPOSITION IS
8/9 – NEUTRONS, 1/9 – H1

Fig. 11.

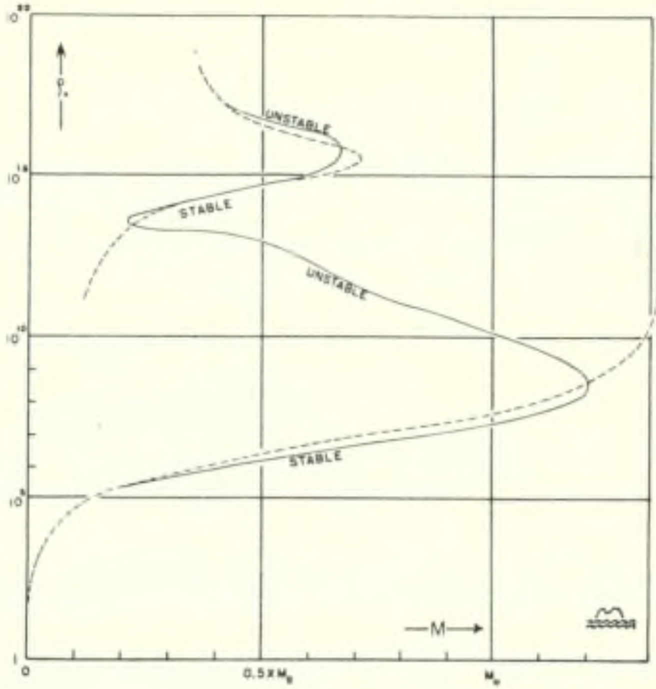


Fig. 12.

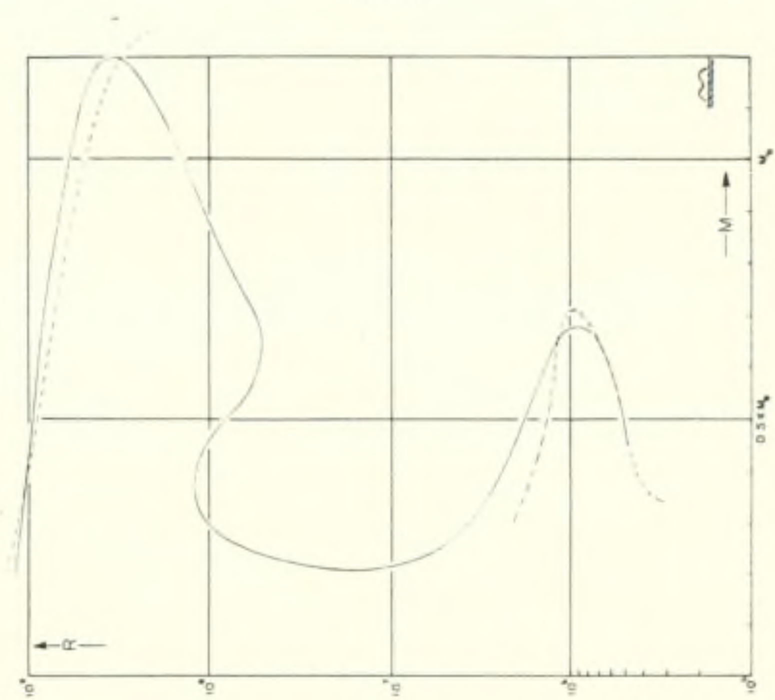


Fig. 13.

vistic, and its rest plus kinetic energy is approximately proportional to its momentum,

$$E \doteq cp \sim \hbar c A^{1/3}/a \tag{83}$$

and independent of its mass.

(4) Newtonian treatment of gravitation; no general relativity corrections; gravitational energy per particle of order :

$$E_{grav} \sim - G(AM_p)\, M_p/a \tag{84}$$

Landau notes that the Fermi energy of the particles dominates over the gravitational energy when the total number of particles is small; then an equilibrium is reached at a non-relativistic energy. This energy increases with number of particles and pressure. For high total mass the energy per particle is pushed into the relativistic limit. Then (83) applies. The total energy per particle is roughly represented by an expression of the form :

$$[(\hbar c/G) - A^{2/3}\, M_p^2] \quad (A^{1/3}\, G/a) \tag{85}$$

The cold degenerate star becomes unstable when the number of particles reaches a limit of the order A_{crit}, where :

$$\begin{aligned} A_{crit}^{1/3} &= (\hbar c/G)^{1/2}/M_p = 2.2 \times 10^{-5}\, g/1.6 \times 10^{-24}\, g \\ &= 1.3 \times 10^{19} \\ A_{crit} &= 2.2 \times 10^{57} \\ M_{crit} &= A_{crit}\, M_p = 3.7 \times 10^{33}\, g \\ &= 1.85 \times (1.99 \times 10^{33}\, g) = 1.85\, M_{\odot} \end{aligned} \tag{86}$$

The limiting mass in this approximation is independent of the type of particles responsible for the pressure; it is the same when it is the outer electrons that are being crushed, and when the nuclei themselves are in contact and the nucleons are being crushed together. Of course the densities are of totally different orders of magnitude for the two critical configurations, each at the very verge of instability. In a more detailed analysis the critical masses at the two crushing points turn out to differ by a factor about 2.

Chandrasekhar made a detailed calculation ([11]) of the equilibrium under gravitational forces assuming :

(1) Densities sufficiently low (10 to 10^8 g/cm^3) that electron pressure is decisive and the nucleons contribute only inert mass, expressed in an « electron effective molecular weight » or « molecular weight per electron ».

$$\mu = \Sigma\, n_i\, A_i / \Sigma\, n_i\, Z_i. \tag{87}$$

(2) Accurate treatment of electron pressure spanning the entire range from non-relativistic to relativistic energies.

(3) Newtonian theory of gravitation; no general relativity corrections.

On these simplifying assumptions he found that the critical mass for crushing of the electrons is :

$$\begin{aligned} M_{chandra} &= (2.015/2)\,(3\pi)^{\frac{1}{2}}(\hbar c/G)^{3/2}/\mu^2\, M^2{}_p \\ &= 1.14 \times 10^{34} g/\mu^2 = 5.73\, M_{\odot}/\mu^2 \end{aligned} \tag{88}$$

His results at lesser masses give a connection between central density and total mass that is displayed by the lower dashed curve in figure 12. The curve is drawn for the case $\mu = 56/27 = 2.07$. The case of more interest, ${}_{26}Fe^{56}$, corresponds to $\mu = 56/26 = 2.15$. It can be obtained from the dashed curve in the diagram by a small change in the M scale (proportional to μ^2) and in the ρ scale (proportional to μ). The Chandrasekhar curve clearly does not display the second crushing point.

The second crushing point has in turn been treated by Oppenheimer, Serber and Volkoff ([12]) (upper dashed curve in fig. 12) in a way which neither brings into evidence the first crushing point nor was intended to do so. They assumed :

(1) All electrons squeezed onto protons to make neutrons; all nuclear structure dissolved away; nuclear binding negligible in comparison with Fermi kinetic energy of compressed neutron gas.

(2) Accurate treatment of pressure of ideal gas of free neutrons spanning entire range from non-relativistic to relativistic energies (identical with Chandrasekhar's equation of state for electrons apart from a scale change) : $p = p(\rho)$ or $\rho = \rho\,(p)$.

(3) Accurate general relativity formulation of equations of hydrostatic equilibrium,

$$dM(r)/dr = 4\pi\rho(p)\,r^2, \tag{89}$$

$$dp(r)/dr = -\frac{[\rho(p) + c^{-2}p]\,G\,[M(r) + 4\pi c^{-2}p(r)\,r^3]}{r\,[r - 2GM(r)/c^2]} \tag{90}$$

Starting with one or another assumed density $\rho = \rho_0$ at the center of the system, Oppenheimer and Volkoff integrated these equations and found no solution for masses greater than :

$$M_{\text{Opp Vol}} = 0.7\ M_{\odot}, \tag{91}$$

as indicated by the upper dashed curve in figure 12.

To try to connect up the two crushing points with each other and resurvey the whole situation, we examined the consequences of an equation of state that attempts to span the whole region from normal densities to supranuclear densities (fig. 11). We assume :

(1) Cold matter ideally catalyzed to the end point of thermonuclear evolution. This end point at ordinary pressures is Fe^{56}. At higher pressures the nuclear composition of minimum energy is altered. This region was treated by making use of the usual semi-empirical formula for nuclear binding energies. At very high densities nuclear forces were overlooked and all particles were treated as belonging to ideal Fermi gases. (K. H.).

(2) General relativity equations of hydrostatic equilibrium (89) and (90). These equations were integrated numerically on the MANIAC computer of the Institute for Advanced Study with the kind collaboration of Mrs. Barbara Weymann and the helpful advice and support of Dr. Hans J. Maehly. (M. W.). The preliminary and tentative account of the results given here in figures 12 and 13 is to be followed later — it is intended — by a more detailed account elsewhere (K. H., M. W., J. A. W.).

The numerical integrations show for the first time both crushing points on a single curve for M as a function of ρ_0. This curve agrees as closely as can be expected in the appropriate ranges of density with the results of Chandrasekhar and Oppenheimer and Volkoff. The agreement emphazized more strongly than ever the conclusion that followed already from the work of Oppenheimer and Volkoff : *assembly of an amount of mass that exceeds in order of magnitude the* *mass of the sun and catalysis of this matter to the endpoint of energy evolution results in a condition which lies at the untamed frontier between elementary particle physics and general relativity.*

Of all the implications of general relativity for the structure and evolution of the universe, this question of the fate of great masses of matter is one of the most challenging. Moreover, the issue cannot be escaped by appealing to stellar explosion or rotational disruption, for the issue as it presents itself today is one of principle, not one of observational astrophysics.

Won't the star explode? Let it! Let the cycle of thermonuclear reactions proceed in whatever way later more detailed analysis shall reveal. Let one or another reaction suddenly come into prominence and possibly lead to instability. Let it cause one or more explosions. Let an explosion drive off any fraction of the mass of the system. Simply catch the ejected matter and extract its kinetic energy. Let it fall back on the star. Then the original number of nucleons, A, is restored; but the mass-energy of the system drops. Ultimately the star gets tired. It can't eject matter. It can't radiate photons. It can't emit neutrinos. It comes into the absolutely lowest state possible for an A-nucleon system under the dual action of nuclear and gravitational forces. However impossibly long the time may be to get to this final state, this is the state in which we are interested as a matter of principle.

If the star originally possesses angular momentum, then even more energy can be extracted from it in its passage to the final state.

It may not be necessary to found this discussion entirely on idealized experimentation and issues solely of principle. It is conceivable that astrophysical evolution may occasionally or even often lead to contracting stars substantially more massive than the sun.

M. Schwarzschild has recently analyzed theoretically the behavior of stars with initial masses of 30, 60, 120 and 220 $M_{\odot}$. He finds pulsational instability at $\approx 70\ M_{\odot}$ and above. It is not clear whether the pulsations will be large. One does not see any large pulsating stars. It may therefore be that such a star sheds an outer shell of unconverted matter. If a star survives pulsational instability at the start of H burning, Schwarzschild finds that it remains stable for the next 2 or 3×10^6 years. A star of mass 50 $M_{\odot}$ can burn in this time 50 percent of its hydrogen quietly and without instability. Problems about the working of the convective core come in at an earlier date for smaller stars. It would be interesting to follow through the later evolutionary stages of a star of mass 50 $M_{\odot}$ and find how much of its mass remains at the start of gravitational contraction.

If a star is rotating and its core starts to undergo gravitational contraction to superdensities, it is quite conceivable that the core may never face the issue of rotational disruption. Magnetic fields will couple it to the outer shell. This coupling will work to ensure a common angular velocity for core and shell. The moment of inertia will come predominantly from the shell. Therefore it may not undergo great changes. Consequently it is quite possible that the rotation of the core may not speed up significantly as the core contracts.

It is clear that a full treatment of stellar evolution will take many years, with or without allowance for rotation. Physical chemistry has taught one to look at the final equilibrium state of a reaction before one asks the much more complicated questions of mechanism and rate of approach to that equilibrium. The equilibrium in the reaction

$$2H_2 + O_2 \leftrightarrows 2H_2O$$

is well known; but the mechanism and rate are fantastically involved and only partly understood after a half a century. Therefore one has a powerful motivation to ask once more our question, *What is the final equilibrium state of an A-nucleon system under gravitational forces when A is large?*

Perhaps there is no equilibrium state : this is the proposal of Oppenheimer and Snyder ([13]). They consider a collection of particles separated from their common center by distances of the order of the solar radius. They note that the fall towards a K. Schwarzschild singularity will take only a few hours as measured by an observer on one of the particles, but will take forever as measured by a remote stationary observer. They suggest the same indefinitely prolonged fate for a star whose mass exceeds the critical mass.

A new look at this proposal today suggests that it does not give an acceptable answer to the fate of a system of A-nucleons under gravitational forces :

(1) No mechanism of release of the gravitational energy into the surroundings is taken into account. The particles are considered to

convert gravitational potential energy into kinetic energy, but not into heat and radiation. Therefore this approach excludes from the start any decrease in the massenergy of the system and rules out *by definition* any approach to an equilibrium, if there is one. The mass of the system as viewed by a distant observer remains forever the same. In actuality the particles will collide, give off heat, lose speed and slow down their contraction.

(2) The particles are envisaged as falling into a K. Schwarzschild singularity. However a Schwarzschild singularity does not give an adequate representation of the forces sustained by a particle at high compression. The forces between nucleons come in in a most vital way. Of course it is not clear to what consequences these forces lead. The present review recalls that hard core forces at the one extreme of an incompressible liquid are as incapable of sustaining the system as are the pressures of a perfect Fermi gas at the opposite extreme. But it does appear that *any answer is incomplete that does not deal with the ultimate consitution of a nucleon.*

(3) The particles are envisaged as « cutting themselves off from the rest of the universe. » (13). This expression would seem to suggest that the particles lose their effect on the rest of the universe. However, the picture under discussion demands at the same time that they maintain an unchanged gravitational pull on a distant test mass — the direct opposite of losing their effect.

Considering today the final state of a system of many nucleons, we must rule out an equilibrium mass equal to the mass of the unassembled particles. The final mass must be very substantially smaller than the original mass; it must also be finite, and limited to a fixed upper bound, no matter how many nucleons are introduced into the system.

If we are to reject as physically unreasonable the concept of an indefinitely large number of nucleons in equilibrium in a finite volume of space, it seems necessary to conclude that the nucleons above a certain critical number convert themselves to a form of energy that can escape from the system : radiation. If the energy were to escape in the form of particles, we could in principle extract the energy from the emerging particles and then let them fall back on the system. The build up of these particles on the system would then ultimately lead back to the paradoxical situation from which an escape is sought.

Radiation presents no such difficulty. However low its energy, it can always escape from the system by travelling radially outwards. No escape is apparent except to assume that the nucleons at the center of a highly compressed mass must necessarily dissolve away into radiation — electromagnetic, gravitational, or neutrinos, or some combination of the three — at such a rate or in such numbers as to keep the total number of nucleons from exceeding a certain critical number.

In view of the absence of any acceptable alternative equilibrium, it appears desirable to take seriously this possibility of nucleonic disruption and explore its consequences. Dissolution of nucleons into neutrinos at very high pressures would be a process fully compatible with the laws of conservation of momentum and energy. It would violate the law of conservation of nucleon number, but leave unaffected most other conservation laws. The present lower limit to the life of the nucleon against spontaneous breakup, $T_{\frac{1}{2}} > 4 \times 10^{23}$ years, [14] refers to pressures of a so much lower order of magnitude that no inconsistency arises. A motion picture of a large mass of nucleons dissolving away under high pressure into free neutrinos presents a fantastic scene when run backwards. Sufficiently many neutrinos of the right helicity coming together from all directions into one region of space over a short time interval form themselves into nuclear matter.

To put matters into perspective, it is necessary again to say that questions of nucleon formation and nucleon destruction lie outside the bounds of present theory. However, it is possible to recall at the same time the often emphasized point that the number of nucleons in the universe is a dynamic variable, the value of which may not have had the same magnitude in the successive oscillations of the universe. On the contrary, conditions of superdensity would seem to be particularly favorable for altering the number of nucleons in the universe.

CONCLUSION (J.A.W.)

In conclusion, general relativity makes important predictions and poses important problems about the structure and evolution of the universe. Among the predictions one of the most striking — and probably before long one of the easiest to test — is the existence of an upper limit to the angle effective distance of any galaxy (figs. 2 and 3) and the presence of fainter and brighter galaxies at any lesser angle effective distance.

Among the problems some of the most challenging are : (1) to observational science, the approximate magnitude of the so-far unexplored sources of mass-energy (fig. 5); (2) to geometrodynamical analysis, what goes on in an unsymmetric implosion and reexpansion of the universe (a redo of fig. 7!) and (3) to elementary particle physics, the fate of a great number of nucleons under the action of their mutual gravitational attraction.

ACKNOWLEDGMENT

We are indebted to many colleagues for illuminating discussions, among them H. Bondi, Jan Oort, Robert Oppenheimer, Martin Schwarzschild, Lyman Spitzer and George Volkoff; and to Hans Maehly and his collaborators at the MANIAC for their help with the calculations reported in figures 12 and 13.

Fig. 1. — Radius a of ideal spherical universe as a function of cotime $T = ct$ since start of expansion. The upper curve applies to the case of a universe filled with inert matter exerting no pressure and the lower curve applies to the opposite limiting case of a radiation filled universe where the pressure has the maximum possible value. The upper curve is a cycloid generated by rolling a circle around an angle of 2π. The lower curve is a half circle where the angular parameter goes from 0 to π. The different ratio between cotime elapsed and maximum radius in the two cases can be stated in this form : an observer near the end of the period of recontraction will in the upper case receive light which has made the entire circuit of the universe starting from a point near him while in the lower case he will be receiving light which has come only from the antipodal point of the universe. The time, H^{-1}, associated with the Hubble constant, is found in either case by extrapolating back the tangent to the curve $a(T)$, as indicated in the diagram.

Fig. 2. — Relation between red shift and angle effective distance (in 10^9 light years) for the spherical universe of Einstein's general relativity filled with dust (smooth curves) and filled with radiation (dashed curves). The dot-dash curves give the predictions of Bondi-Gold-Hoyle model. The existence of a maximum angle effective distance is a decisive consequence of Einstein's general relativity. By way of illustration consider an observer at the North Pole of the earth who launches rockets to two towers located at the same latitude but separated by a kilometer. If they are 10 km from the North Pole the launching directions must differ by 1/10 radian. Precisely the same angular separation of launching directions will also serve, however, if the two towers are 10 miles from the South Pole. In both cases the angle effective distance, defined as the size of the object divided by the apparent angular diameter, has the value 1 km/0.1 radian = 10 km. The two towers will obviously have the maximum angle effective distance from the North Pole when they lie at the equator, 10,000 km travel distance to the south. The angle effective distance is $6360 \sin 90° = 6\,360$ km. In the case of the universe, light coming from two sources of the same size and intensity and at the same angle effective distance, one close, the other far away, will obviously have started at very different times and have experienced very different red shifts. For this reason the most remote one will naturally appear much fainter even though it has the same apparent size as the nearer one. Observations are still insufficient to show whether such an effect occurs as expected. — From the analogy between the curved surface of the earth and the curved space of general relativity, where the two towers stand for the two extremities of one galaxy, it is not right to conclude that a galaxy far enough away to have zero angle effective distance is antipodal to our own. The curved surface of the earth was envisaged as static while the curved universe is expanding at a rate significant in comparison with the speed of light. The most ancient signal that reaches the observer was emitted practically at the start of the expansion of the universe, at a moment when the spherical space had a very small radius. The time that has elapsed between then and now has not been sufficient for the signal to come from the antipodal point. It comes from one or another of the points which are separated from our own location by an angle on a sphere which is equal to the present value of the angular parameter which measures the present time: the angle through which the cycloid-generating circle has turned in the upper example in figure 1 (max η between 0 and 2π). In the lower example in figure 1 the angle η that governs the time from the start of expansion to the present instant is measured on the half circle itself and has a value somewhere between 0 and π. More generally $\eta = \int_{o}^{now} dT/a(T)$. Objects at the maximum angle η from us on the sphere gave out light — which we today receive from them — very near the start of the expansion of the universe. They should appear enormously magnified in size but very faint and very reddened over the entire celestial sphere according to Einstein's general relativity.

Fig. 3. — Fraction f of the galaxies comprised within any specified angle effective distance (in 10^9 light years), according to Einstein's general relativity for a spherical universe filled with dust (smooth curves) or filled with radiation (dashed curves). The curves are double valued. The galaxies that lie at a given angle effective distance from the observer are divided into two classes, one much farther away than the other. The so-called steady state model of Bondi-Gold and Hoyle is unbounded so that the fraction f cannot be defined. Quantity f^* (dot-dash curves) which is plotted in this case nevertheless gives a dimensionless measure of the number of galaxies up to the given angle effective distance.

Fig. 4. — Relation between red shift of galaxies at a given distance and the fraction f of all galaxies which are included within that distance, according to Einstein's general relativity. In contrast the dot-dash curve plots for the Bondi-Gold-Hoyle steady state universe a dimensionless measures, f^*, of the number of galaxies out to a given distance.

Fig. 5. — Density of matter-energy required to hold the universe together compared to known and yet-to-be-studied sources of density. The curves in the upper diagram tell how much present density is required for any specified value of the present Hubble constant ($H^{-1} = 8, 10, 12, 16, 20 \times 10^9$ years) and any present value of the age of the universe. The calculated density differs appreciably according as the source of the density is primarily inert matter (dust; smooth curves) or radiation (dashed curves). The cross-hatched regions assume 5 and 9×10^9 years as lower and upper limits on the age and 10 and 16×10^9 years as lower and upper limits on the inverse Hubble constant. The density deduced from these limits lies between 7 and 100×10^{-30} g/cm^3 as indicated by the cross-hatched region in the lower diagram. There the black bars indicate known sources of mass-energy and approximate uncertainty limits on present values. Among the potential sources of mass-energy for which no present estimates are available are two (neutral H and anti-neutrinos) where approximate upper limits can be given. In view of present day ignorance about several potential sources of mass-energy it is impossible to say that there is any contradiction between the facts and the predictions of Einstein's general relativity. According to a late communication by George Field (October 1958) the limit on hydrogen atoms is reduced to

$$[\sim 2 \times 10^{-29} \text{ g/cm}^3]$$

if the ionization does not exceed a critical limit of about 20 percent.

Fig. 6. — Schematic representation of relation between uniform dust model and model where mass is concentrated in a regular 3-dimensional lattice of points. For ease of visualization the resulting variably curved space is here conceived as imbedded in a flat Euclidean 4-space. Of course the fourth dimension has nothing to do with time! Some of the higher dimensionality has been thrown away to permit representation in the plane of the paper. Above : Schwarzschild metric due to a single mass. The dots mark the intended place of cut. Below : polygonal pieces of many such Schwarzschild metrics joined — with slight readjustment around the boundary of each lattice zone — to make a closed space. An effective radius of curvature of this closed space is defined by the dashed line. This radius is shown in reference 3 to have very nearly the same value and very nearly the same rate of change with time as the radius of a uniform dust-filled space of the same total mass.

Fig. 7. — Photographs of an underwater explosion bubble against a white background at 0, 1, 5.7, ... milliseconds after detonation, showing the dynamical instability and the formation of prongs and spikes during the contraction phase. The 0.55 pound Tetryl charge was exploded 300 feet below the surface. From Cole, *Underwater Explosions*, courtesy of the author.

Fig. 8. — Amplitude of small disturbances in density as a function of the time parameter ct/a_0 given by equation 42 for a disturbance odd (anti-symmetric) with respect to inversion of time about the moment ($\eta = \pi$) of maximum expansion, by equation 43 for a disturbance even (symmetric) with respect to time reversal and by such a linear combination of the two that the result is regular at the start of expansion. The amplitudes are given in arbitrary units.

Fig. 9. — Mass of an incompressible fluid sphere as a function of the mass before assembly. The curves differ from each other according as one or another value is assumed for the density of the incompressible Schwarzschild fluid. The density of nuclear matter, a little over 10^{14} g/cm^3, corresponds to a curve a little lower than the middle curve. The independent variable actually plotted is not the mass before assembly but a quantity proportional to it, the number of nucleons. According to the reasoning in the text it is not legitimate to follow one of these idealized curves to its upper limit where the tangent runs horizontally and where an added nucleon radiates away all its mass. Instead it is necessary to stop at a number of nucleons 16 percent short of the upper limit indicated in this diagram to keep the pressure finite at the center of the fluid mass and to keep the seas of positive and negative energy states from merging with each other.

Fig. 10. — State of an idealized incompressible fluid held together by its own gravitational attraction. Above, pressure in dimensionless units as a function of distance from the center, also in dimensionless units; below, the metric factor $e^{\frac{1}{2}\nu}$ that measures the separation of positive and negative energy states. The heavy curve applies to the case where the sphere has the maximum physically acceptable mass : at the center the pressure goes to infinity and the separation of positive and negative energy seas goes to 0. The dashed curves apply to a smaller mass.

Fig. 11. — Equation of state of cold matter catalyzed to the end point of thermonuclear evolution. The smooth curve gives the ratio P/ρ in cm^2/sec^2 as a function of the density ρ in g/cm^3. The dashed curve gives the ratio $P/a^{4/3}$ in g cm^3/sec^2 as a function of the number of nucleons/cm^3, a. Until the density arrives at a function of the order of 10^6 g/cm^3 the pressure is insufficient to squeeze electrons onto the nucleus or to alter the position of the minimum on the packing fraction curve and the composition is therefore 100 percent Fe^{56}. Higher pressures alter the nuclear composition of minimum energy in the direction of heavier nuclei with a higher ratio of neutrons to protons. Eventually (f) the point is reached where the nucleus can hold no more neutrons and the equilibrium composition is a mixture of free neutrons, degenerate electrons with a high Fermi energy, and the nuclear species appropriate to the given pressure. The equilibrium was calculated by use of the familiar semi-empirical mass formula. At density of the order of 10^{14} g/cm^3 the medium becomes as compact as nuclear matter and the further course of the pressure density relation becomes more uncertain. If a hard core repulsion begins to dominate then it appears reasonable to consider as idealized representation of the equation of state an incompressible fluid (Figs. 9 and 10 for the Schwarzschild fluid). In the present curves the opposite idealized limiting behavior was assumed : a perfect Fermi gas at very high pressures. The equilibrium goes asymptotically to a mixture of neutrons, protons and electrons in the ratio 8:1:1.

Fig. 12. — Mass of a star made out of cold catalyzed matter in units of the mass of the sun, $M_{\odot}$, 1.987×10^{33} g, as a function of the central density, ρ_0, in g/cm^3 as deduced from 45 integrations performed on the MANIAC computer of the Institute for Advanced Study with the kind collaboration of Dr. H. J. Maehly and Mrs. Barbara Weymann. To start an integration, a value was assumed for the central density and the general relativity equations of hydrostatic equilibrium were then integrated step by step out to the point where the pressure went

to zero. In the integration the numerical equation of state of figure 11 was employed. The lower dashed curve shows the connection between mass and central density which was deduced by Chandrasekhar without the use of general relativity, where the pressure derives entirely from an ideal electron gas. (See text for value inserted into his formula for effective mass per electron). The upper dashed curve represents the results of Oppenheimer and Volkoff (*Phys. Rev.*, **55**, 378, (1939) derived by integrating the general relativity equation of hydrostatic equilibrium for an idealized pure Fermi neutron gas. So far as is known a curve has never before been available to connect the region where the pressure is mainly of atomic origin with the region where the pressure is mainly of nuclear origin. The two transitions on this curve from stable equilibrium to unstable equilibrium mark what might be called two « crushing points ». The lower curve marked « stable » is of course, where it runs under the upper curve marked « stable », actually metastable with respect to contraction to a high density; however, a potential hill has to be surmounted to pass from the less compact condition of stability to the more compact one. From that more compact configuration it appears conceivable by surmounting another potential hill to pass to a configuration of still lower energy; but to say much about this condition of lower energy is beyond the present power of general relativity and elementary particle physics.

Fig. 13. — Connection between outer radius, R, in cm, and the mass of the 45 configurations of cold catalyzed matter which are summarized in figure 12.

REFERENCES

(1) A. Einstein : « The Meaning of Relativity », *Princeton Univ. Press*, Princeton, New Jersey, 5th Edition (1955), p. 107 and p. 127.

(2) For a summary see for example L. Landau and E. Lifshitz : « The Classical Theory of Fields » (translated by M. Hammermesh), *Addison-Wesley Press, Cambridge and Reading* (1951); also the forthcoming article of O. Heckmann in the new edition of Springer's « Handbuch der Physik ».

(3) R. Lindquist and J. A. Wheeler : *Rev. Mod. Phys.* **29**, 432 (1957).

(4) R. H. Cole : « Underwater Explosions », *Princeton University Press*, Princeton, New Jersey (1948).

(5) E. Lifshitz : *Journal of Physics of the U.S.S.R.*, **10**, 116 (1946).

(6) A table of analogies between geometrodynamics and hydrodynamics appears in E. Power and J. A. Wheeler : *Rev. Mod. Phys.*, **29**, 480 (1957).

(7) We are indebted to G. E. Tauber and J. W. Weinberg for the opportunity to study their unpublished article : « The Internal State of a Gravitating Gas ». It defines a density in phase space considering the totalized amount of world line arc length contained in the typical 8-volume $dx^0dx^1dx^2dx^3dp_0dp_1dp_2dp_3$.

This definition appears unusable in the present connection (1) because p_0 for photons is uniquely determinded by $p_1p_2p_3$, and (2) because the world line of a photon always has zero arc length. This question was investigated by J. B. Adams, A. B. Senior thesis (unpublished) : « Investigations for a Covariant Transport Equation for Non-Interacting Particles », *Princeton University* (May 1958).

(8) J. A. Wheeler : *Phys. Rev.*, **97**, 511 (1955).

(9) J. Weinberg : *Phys. Rev.*, **75**, 338 (A) (1949).

(10) L. Landau : *Phys. Zeits. Sowjetunion*, **1**, 285 (1932).

(11) S. Chandrasekhar : *Monthly Not. Roy. Astr. Soc.*, **91**, 456 (1931); **95**, 207 1935); **95**, 226 (1935). See also M. Rudkjobing : *Kgl. Danske Vid. Sels. Math. Fys. Medd.*, **27**, No. 5 (1952).

(12) Order of magnitude analysis : J. R. Oppenheimer and R. Serber, *Phys. Rev.*, **54**, 540 (1938); detailed general relativistic treatment, also numerical integration : J. R. Oppenheimer and G. M. Volkoff, *Phys. Rev.*, **55**, 374 (1939).

(13) J. R. Oppenheimer and H. Snyder : *Phys. Rev.*, **56**, 455 (1939).

(14) F. Reines, C. L. Cowan, Jr. and H. W. Krusse : *Phys. Rev.*, **109**, 609 (L) (1958).

Discussion of Wheeler's Report

Dr. Oppenheimer. — I do not know whether non rotating masses much heavier than the sun really occur in the course of stellar evolution, but if they do, I believe their collapse can be described within the framework of general relativity; the case of cold free fall can be solved.

Dr. Hoyle. — Is it not likely that this very large neutrino flux that you mention is due to neutrino emission from the sun, rather than from a general cosmical background?

Dr. Wheeler. — It is necessary first of all to stress that the high flux mentioned by Reines is not an observed flux; it is only an upper limit to the flux of neutrinos which would be consistent with the observed underground counting rate. Some of that counting rate is certainly due to events in which penetrating cosmic rays manage to mock up the pattern of time-delayed pulses characteristic of neutrinos. In any case whatever neutrinos are present can hardly come from the sun. First, the equipment of Reines and collaborators is sensitive to the *anti*-neutrinos which come from decaying fission products and from any processes in which neutrons change to protons. In contrast the thermonuclear reactions in the sun change protons to neutrons and give off *neutrinos*, to which the Reines equipment is believed in principle not to be sensitive. Second, the upper limit of the energy of the neutrinos from the thermonuclear reactions in the sun is too low to surmount the energy threshold associated with the detecting process of Reines. Therefore any effect that is real would seem by necessity to be due to antineutrinos that do *not* come from the sun.

Dr. Oppenheimer. — Would not the simplest assumption about the fate of a star of more than the critical mass be this, that it undergoes continued gravitational contraction and ultimately cuts itself off more and more from the rest of the universe?

Dr. Wheeler. — It is very difficult to believe « gravitational cut-off » is a satisfactory answer to the problem for the following reasons :

(1) The only past analysis of this question considered as initial state a neutron gas that was started at rest and allowed to fall freely. During that process gravitational potential energy was converted into kinetic energy of the system. Therefore, this treatment gave possibility for the total mass of the system ever to decrease. The gravitational attraction of the matter, as seen by a distant observer, never changed.

(2) The assumed rapid collapse would in actuality lead to shock waves and heating. This thermal energy can be dissipated as radiation with a resultant decrease in the mass of the system.

(3) One soon comes to densities at which it is simply impossible to make any theoretical analysis of the final state because of our ignorance of the properties of nuclear matter at high density. (The question of « gravitational cut-off » is discussed a little more in the version of the report than it was in the original oral presentation.)

World Models

by **E. SCHÜCKING** and **O. HECKMANN**

1. THE FUNDAMENTAL PROBLEM OF COSMOLOGY

1. On methodology.

Since the days of Sir Isaac Newton scientists have been guided by a principle which might be formulated as follows : A theory constructed on a sound foundation of empirical data ought not to be discarded unless there new facts turn up that cannot be fitted into the framework of this theory.

This very precept was adhered to by the scientists who at the turn of the last century revolutionized physics. Planck introduced the quantum of action because the observed energy distribution of black body radiation was not accounted for by theory. Einstein formulated special relativity because the Michelson-Morley experiment failed to reveal an absolute motion of the earth. The steady progress of science, in our opinion, was possible only because scientists thought it not permissible to put forward new theories unless new data forced them to abandon the older concepts.

Without this principle physics would become a hopeless chase, for there are innumerable ways of modifying known laws of nature by playfully introducing into their mathematical description terms as yet not detectable by observation. Progress would be paralysed if experimental scientists were held back by the labor of verifying the reality of such correction terms introduced at will.

It would seem that some cosmologists have abandoned the aforementioned principle of methodology. Ten years ago Bondi, Gold, and Hoyle launched their steady-state theory. They denied the validity of the laws of local conservation of energy and momentum because models of a homogenous universe with isotropic expansion and vanishing Λ-term in the field equations of Einstein's theory of gravitation would have implied an age of the universe difficult to reconcile with other numerical data of astronomy and nuclear physics then current. After recalibration of the distance scale of extragalactic objects it became evident that this apparent discrepancy had resulted from erroneous astronomical data. At any rate, it could have been removed by retaining the Λ-term in the field equations. There is no reason to doubt the validity of the conservation laws of energy and momentum. We believe, therefore, it is sound policy to refrain from theorizing along the lines of Bondi, Gold, and Hoyle until there is strong empirical evidence for continuous creation of energy and momentum.

In order to understand the large-scale structure of the world, we prefer to retain the experimentally and observationally well established laws of physics. Moreover, we assume tentatively that these laws, though established here and now in the domain of the solar system, are valid everywhere and for all times. In particular we assume that all so-called fundamental constants of nature are invariable. Accordingly we exclude explicitly Jordan's theories of gravitation and Milne's cosmology. The assumptions of the invariability of the constants of nature over large regions of space-time is supported by the measurements of line-shifts in the spectrum of the radio source Cygnus A, as pointed out by Minkowski. The distance of this object is about 700 million light years, if we accept the Hubble constant $H = 75$ km/sec Mpc (Sandage). Comparing the lineshifts in the visual region (Minkowski) (1) with the shift of the 21 cm absorption line of hydrogen (Lilley) (2) one is forced to conclude that seven hundred million years ago in the Cygnus A region the fine-structure constant had the same value as *hic et nunc*. For the ratio of optical and radio frequencies is essentially determined by the square of the fine structure constant.

2. **The characteristical initial value problem.**

All observations of the universe refer to events lying on the light-cone which reaches into the past. These data furnish certain initial conditions for the hyperbolic field equations of Einstein's theory of gravitation. But so far the mathematical problems underlying this observational approach to cosmology have not been solved. We simply do not know which initial data on a segment of the light-cone are required for the determination of the field in a certain region of space-time. At any rate, it is clear from the general nature of the field equations that all inferences from observation refer exclusively to the past, more precisely to the shaded region of figure 1. The volume of this region depends on the penetrating power of the telescopes. Only after the characteristical initial value problem has been solved will it be possible to examine the distribution of matter, some 10^9 years ago, over a finite region of space and to test its homogeneity.

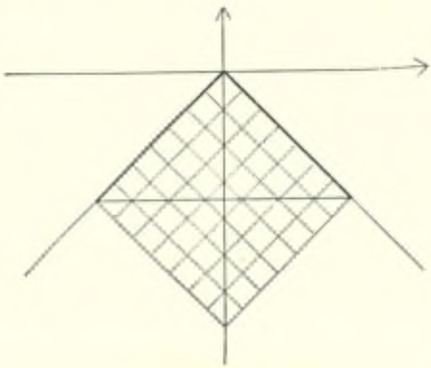


Fig. 1

In this two-dimensional " world-map " space-like directions have been drawn horizontally. The vertical axis is the world-line of an observer. His observations refer to events on the light cone. The shaded area is the space-time volume that can be investigated by solution of the characteristical initial value problem of the field equations.

3. The world-postulate.

Several decades ago cosmologists proposed a far reaching hypothesis about world-structure in the large, the so-called « world-postulate ». In the frame work of Einstein's theory of gravitation this postulate states that not only the continuum of space-time but also the momentum-energy-vector are invariant with respect to a six-parameter continuous group. This postulate asserts the permanent homogeneity of certain three-dimensional subspaces and affords isotropy of expansion for each observer. In a certain sense the world-postulate demands too much because it is formulated without regard to the field equations. It is sufficient, in fact, to postulate merely the spatial homogeneity and isotropy of the metric of a three-dimensional subspace and its first derivatives with respect to time at a certain instant of time $t = t_0$. This postulate is a special initial condition for the field equations which govern the temporal development of the model. The unrestricted form of the world-postulate introduces redundant conditions, it might even lead to contradictions with the field equations. Hoyle as well as Jordan were obliged to modify Einstein's field equations because their assumptions about the temporal evolution of world models contradicted these equations.

It is not necessary to review here those world-models which in the frame work of Einstein's theory obey the six parameter world-postulate. They are well known. However, because the world-postulate is satisfied by nature only approximately, if at all, more general world-models command a considerable interest provided they obey the world-postulate only approximately. Moreover, only the study of these more general models can reveal how sensitive the homogeneous and isotropic models are to altering the underlying assumptions. In particular they will greatly modify our conception of the early phases of the universe.

2. CLASSIFICATION OF WORLD-MODELS

1. Separation of the four vector of material flow.

A clear-cut classification of several world-models without any postulate of homogeneity is possible by investigating the four-vector of flow of incoherent matter. Evidently the tensor of the first derivative of the flow-vector admits of being separated as follows :

$$
\begin{aligned}
u_{\mu\|\nu} &= \left(g_{\mu\nu} - u_{\mu} u_{\nu}\right) \frac{\mathrm{DR}}{\mathrm{Ds}} + w_{\mu\nu} + q_{\mu\nu} \quad , \\
\frac{\mathrm{DR}}{\mathrm{Ds}} &= \frac{\mathrm{R}_{|\lambda}}{\mathrm{R}}\, u^{\lambda} = \frac{1}{3}\, u^{\lambda}_{\|\lambda} \quad , \\
w_{\mu\nu} &= \frac{1}{2}\left(u_{\mu\|\nu} - u_{\nu\|\mu}\right) = \frac{1}{2}\left(u_{\mu|\nu} - u_{\nu|\mu}\right) = - w_{\nu\mu} \quad , \\
q_{\mu\nu} &= \frac{1}{2}\left(u_{\mu\|\nu} + u_{\nu\|\mu}\right) - \frac{1}{3}\left(g_{\mu\nu} - u_{\mu} u_{\nu}\right) u^{\lambda}_{\|\lambda} \quad . \\
q^{\mu}_{\cdot\mu} &= 0 \, .
\end{aligned}
\tag{1}
$$

Here greek indices run from 0 to 3, $g_{\mu\nu}$ is the metrical fundamental tensor and the covariant derivative is indicated by two vertical bars — the common partial derivative by one bar. The symbol $\frac{\text{« D »}}{\text{« Ds »}}$ designates the invariant derivative along the world line with the tangential vector u^λ. The conservation laws and Einstein's « 00 »-equation then give :

$$\frac{4\pi}{3}\rho R^3 = \mathfrak{M} = \text{const. along each world line} \quad ,$$

$$\frac{D}{Ds}\left(R^2 w_\mu\right) = \left(R^2 w_\mu\right)_{\|\lambda} u^\lambda = q_{\mu\nu} R^2 w^\nu \quad ,$$

$$\frac{D^2R}{Ds^2} - \frac{R}{3}\left(\Lambda - q_{\mu\nu}q^{\mu\nu}\right) - \frac{2}{3} R w_\mu w^\mu + \frac{G\mathfrak{M}}{R^2} = 0, \quad (2)$$

$$w^\mu = \frac{1}{\sqrt{-g}}\,\varepsilon^{\mu\nu\lambda\sigma}\, w_{\nu\lambda} u_\sigma, \qquad w^\mu u_\mu = 0 .$$

w_μ is the vector of angular velocity conjugated to the tensor $w_{\mu\nu}$, ρ the density of matter. $\varepsilon^{\mu\nu\lambda\sigma}$ a totally skew-symmetric tensor with $\varepsilon^{0123} = 1$.

These four-dimensional tensors are reduced to three-dimensional ones in a co-moving local inertial system. These tensors are the well known terms of expansion, rotation, and shear in a general velocity field. The spatial components u_i of u_μ form the three-dimensional velocity vector. By definition we have :

$$u_{i|k} = \frac{R_{10}}{R}\,\delta_{ik} + w_{ik} + q_{ik} \quad ,$$

$$\frac{R_{10}}{R} = \frac{1}{3}\,u_{i|i} \;, \quad \omega_{ik} = \frac{1}{2}\left(u_{i|k} - u_{k|i}\right), \quad (1a)$$

$$q_{ik} = \frac{1}{2}\left(u_{i|k} + u_{k|i}\right) - \frac{R_{10}}{R}\,\delta_{ik} \quad ,$$

$$\omega_i = \varepsilon_{ijk}\, w_{jk} .$$

Locally the conservation laws and the « 00 »-equation for incoherent matter are transformed exactly into the corresponding equations of Newtonian theory in a comoving local inertial system. These Newtonian equations have the following form ([3]) :

$$\frac{4\pi}{3}\rho R^{3} = \mathfrak{M} = \text{const.} > 0$$

$$\left(\dot{R}^{2} w_{i}\right)^{\cdot} = q_{ij}\left(R^{2} w_{j}\right) \qquad , \quad (2a)$$

$$\ddot{R} - \frac{R}{3}(\Lambda - q_{ij}q_{ij}) - \frac{2}{3}R w_{i} w_{i} + \frac{G\mathfrak{M}}{R} = 0 .$$

G is the constant of gravitation and a dot indicates the derivative with respect to the time t. The equations (2a) are locally valid independently of any postulate of homogeneity. They have the following meaning. The first is the equation of continuity for the conservation of matter. The second furnishes the vortex theorems of Helmholtz. They state that the vortex lines are frozen into the matter and that the angular velocity of a comoving surface element perpendicular to the vector of angular velocity is inversely proportional to the area of this element. In general therefore cosmic matter will show a precession when compared with a comoving inertial system. Only if this vector of angular velocity points in the direction of one of the principal axes of the tensor q_{ij} does this precession disappear.

The third equation shows how the temporal expansion is influenced by shear and rotation. This equation was already discussed by A. Raychaudhuri ([4]) in a special system of coordinates in Einstein's gravitational theory. It shows that the rotation of the world matter counteracts the gravitational forces and may possibly remove the singularity of the well-known world models. The shear however works in the opposite sense.

2. Types of world-models.

If shear and rotation of incoherent matter are to vanish everywhere, there will be only one type of solutions of the field equation, namely the well known homogeneous and isotropic world models. Our knowledge of homogeneous, but unisotropic models is still quite incomplete. Those homogeneous models in which the angular velocity is not zero do not fulfill Mach's principle. The universal angular velocity is revealed in these cases by the rotation of a linear oscillator relative to a coordinate system with axes orientated in the direction of some galaxies. The only model of this kind, but without shear, is that of Gödel ([5]). The classification of homogeneous world models dates from the work of Bianchi ([6]) at the end of the last century. The solution of the field equations by Bianchi's models ist equivalent to the solution of certain systems of ordinary differential equations. But there arise considerable complications in the course of the cumbersome computations. Nevertheless several solutions are available.

3. SPECIAL SOLUTIONS

1. **Inhomogeneous solutions.**

There exist very many investigations of the spherically symmetrical solutions which need not be considered here. We should mention only one very simple inhomogeneous model, which might be used in discussing the phenomenon of clusters of nebulae (fig. 2).

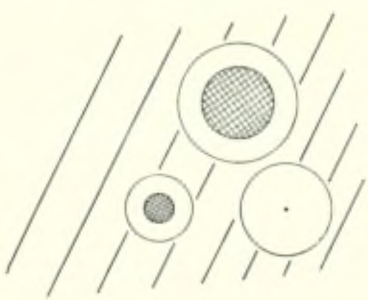


Fig. 2

Spherical condensations of matter are situated in spherical holes which are expanding at the same rate as the general field between the holes,

Consider in a homogeneous universe with isotropic expansion an infinite number of nonoverlapping spherical domains the material content of which is symmetrically condensed into smaller spheres or even mass-points. The matter outside the boundaries of these domains is then expanding at the same rate as before. But the dynamics of a particle moving between the boundary of a domain and its spherical nucleus is unaffected by the general expansion. The « swiss cheese model » satisfies the field equations as was shown by A. Einstein and E. Strauss ([7]). The model can be treated easily in its Newtonian version ([8]).

2. **Homogeneous models.**

K. Gödel discovered in 1949 ([5]) a static homogeneous world model the matter of which is in « absolute » rigid rotation. This model is of the greatest interest for the interpretation of Einstein's theory of gravitation. Because of its static character we shall exclude it here. Instead we consider three anisotropic models. The first and the second one show shear but not rotation. The third one shows shear and rotation.

The *first* model is the generalized model of an expanding space with zero corvature. The line element is :

$$\begin{aligned} ds^2 &= dt^2 - [R_1(t)dx^1]^2 - [R_2(t)dx^2]^2 - [R_3(t)dx^3]^2 \ , \\ u^\mu &= \delta^\mu_0 \end{aligned} \tag{3}$$

For $\Lambda = 0$, for instance, one gets the following solutions of Einstein's field equations :

$$\begin{aligned} R_1 &= t^{\frac{1+2\sin\alpha}{3}} \left(\frac{qG\mathfrak{M}}{2}\, t - a\right)^{\frac{1+2\sin\alpha}{3}} , \\ R_2 &= t^{\frac{1+2\sin\left(\alpha+\frac{2\pi}{3}\right)}{3}} \left(\frac{qG\mathfrak{M}}{2}\, t - a\right)^{\frac{1-2\sin\left(\alpha+\frac{2\pi}{3}\right)}{3}} , \\ R_3 &= t^{\frac{1+2\sin\left(\alpha+\frac{4\pi}{3}\right)}{3}} \left(\frac{qG\mathfrak{M}}{2}\, t - a\right)^{\frac{1-2\sin\left(\alpha+\frac{4\pi}{3}\right)}{3}} , \\ \frac{4\pi}{3}\, \rho\, R_1R_2R_3 &= \mathfrak{M} = \text{const.},\ a = \text{const.},\ \alpha = \text{const.} \end{aligned} \tag{4}$$

This solution can easily be interpreted in a comoving local inertial system in Newtonian theory. If x_i are the Cartesian coordinates of a galaxy with fixed (Lagrangian) coordinates ξ_i we may easily derive from (4), in the case $\alpha = 0$:

$$\begin{aligned} X &= t^{\frac{1}{3}} \left(\frac{qG\mathfrak{M}}{2}\, t - a\right)^{\frac{1}{3}} \xi_1 , \\ X_2 &= t^{\frac{1+\sqrt{3}}{3}} \left(\frac{qG\mathfrak{M}}{2}\, t - a\right)^{\frac{1-\sqrt{3}}{3}} \xi_2 , \\ X_3 &= t^{\frac{1-\sqrt{3}}{3}} \left(\frac{qG\mathfrak{M}}{2}\, t - a\right)^{\frac{1+\sqrt{3}}{3}} \xi_3 . \end{aligned} \tag{4a}$$

The expansion is asymptotically isotropic for $t \to \infty$. But for $t \to 0$ every finite distribution of matter degenerates into an infinite line if $a < 0$.

We can draw similar conclusions for the *second* model which shows an expanding space with negative curvature. The line element can be written :

$$ds^2 = dt^2 - R^2(t)[(dx^1)^2 + S^2(t)e^{2x^1}(dx^2)^2 + S^{-2}(t)e^{2x^1}(dx^3)^2],$$
$$u^\mu = \delta^\mu_0 \tag{5}$$

We get a solution if R obeys the generalized « Friedmann-equation » :

$$\dot{R}^2 = 1 + \frac{\Lambda}{3} R^2 + \frac{2\mathfrak{M}G}{R} + \frac{a^2}{3R^4}$$

with : (6)

$$\chi_\rho = \frac{6\mathfrak{M}G}{R^3} \quad , \quad S(t) = \exp a \int \frac{dt}{R^3} \, , \quad a = \text{const.}$$

For $a = 0$, this is the well known model with isotropic expansion. For $a \neq 0$, we get new models with expansion and shear which show for $\Lambda = 0$ the same instability of the solutions for $t \to 0$ as in the case of zero curvature.

Accordingly a slight anisotropy in the expansion causes a considerable change in a « big bang » theory. For $\Lambda = 0$ we obtain in both models solutions of the field equations for an empty space-time with a Riemannian tensor different from zero. These solutions have singularities which possibly could represent matter. Their physical interpretation has not yet been cleared up [9].

The *third* model is given by the following « ansatz » which contains Gödel's model as a special case. This ansatz gives solutions which depend on several parameters :

$$ds^2 = dt^2 + 2e^{x^1}dtdx^2 - c_{11}(t)(dx^1)^2 - 2c_{12}(t)\, e^{x^1}dx^1dx^2$$
$$- [1 - \alpha c_{11}(t)]e^{2x^1}(dx^2)^2 - S^2(t)(dx^3)^2 \, , \tag{7}$$

$$u^\mu = \delta^\mu_0 \, , \quad \alpha \neq 0 \, , \quad \alpha = \text{const.}$$

The three unknown functions C_{11}, C_{12}, and S are to be solutions of the following differential equations :

$$\frac{1}{4R^2}[(\dot{c}_{12}+1)^2+\alpha(\dot{c}_{11})^2-4\alpha c_{11}] + \frac{\dot{S}}{S}\left(\frac{2c_{12}+c_{11}}{2R^2}-\frac{\dot{R}}{R}\right) = -\Lambda - \frac{\alpha^2}{RS} ,$$

$$\dot{c}_{12}\, c_{11} - \dot{c}_{11}\, c_{12} - c_{11} = -\alpha \frac{R}{S} , \tag{8}$$

$$\frac{\ddot{R}}{R}+\frac{\ddot{S}}{S} - \frac{1-(\dot{c}_{12})^2-\alpha(\dot{c}_{11})^2}{2R^2} = \Lambda - \frac{\alpha^2}{2RS} .$$

It is :

$$R^2 = c_{11} - \alpha\left(c_{11}\right)^2 - \left(c_{12}\right)^2, \; \rho = \frac{\alpha^2}{RS}, \; \omega = \frac{1}{\sqrt{2}\,R} . \tag{9}$$

Here ω is the angular velocity of matter around the x^3 axis.

We have not yet succeeded in solving the system (8) of three ordinary differential equations for the unknown functions c_{11}, c_{12}, and S. But we might mention that general rotation of matter would have the following consequences for a world-model :

1) The singulartity for $t = 0$ might disappear.
2) Without giving up homogeneity we could explain the alleged phenomenon that we observe more clusters of galaxies in one half of the sky than in the other (10).
3) A small difference in the constant of precession, depending on whether it is derived from observations of the distant galaxies or the planetary system could be understood.

REFERENCES

(1) *Astrophysical Journal*, **123**, 373 (1956).
(2) *Astrophysical Journal*, **123**, 172 (1956).
(3) *Zeitschrift für Astrophysik*, **38**, 95 (1955); **40**, 81 (1956).
(4) *Physical Review*, **98**, 1123 (1955).
(5) *Reviews of Modern Physics*, **21**, 447 (1949).
(6) *Mem. Soc. Ital. Sci.*, Ser. IIIa, **11** (1897).
(7) *Reviews of Modern Physics*, **17**, 120 (1945).
(8) *Handbuch der Physik* LIII, Artikel Newtonsche und Einsteinsche Kosmologie.
(9) Proceedings of the Cambridge Philosophical Society, Vol. **53**, Part 4, 836 (1957).
(10) Proc. Internat. Congr. Math. (Cambridge, Mass.), 30 Aug. bis 5. Sept. 1950, 1 p. 175 (1952).

Discussion of Heckmann's Report

Dr Pauli. — I would like to ask which solutions remain, if one assumes the cosmological constants to be zero. There exists a brief summary of Gödel on generalisations of his solution (at a Canadian Mathematical Congress) which is difficult to understand because of its shortness.

Dr Heckmann. — Gödel has not given any explicit solution beyond the static one, which demands $\Lambda \neq 0$. As I said, our investigations are still very incomplete; but I have no doubt that there exist many nonstatic and nonstationary solutions with rotation, shear and expansion in which Λ may be zero.

Dr Wheeler. — Does your analysis allow for the circumstance that ω_i and g_{ij} cannot be both specified independently and arbitrarily, because both must be derivable from the same velocity field u_i ?

Dr Heckmann. — The splitting of the vector field u_i was used only to visualize the classification of world models. In the solutions reported, however, the coordinate system was always chosen so that $u^{\mu} = \delta^{\mu}_{a}$. Your problem therefore never arose in our work.

Dr Oppenheimer. — If we accept the views of Hoyle and his collaborators on element formation, then there was a time some 10^{10} years ago when the physical state of the universe — its nuclear equilibrium — were radically different from today. For instance, ω would have to be very small.

II. SURVEY OF EXPERIMENTAL DATA ON THE UNIVERSE

Distribution of Galaxies and the Density in the Universe

by J. H. OORT

1. Possible state of the universe at the epoch of formation of the galaxies

One of the most striking aspects of the universe is its inhomogeneity. We see in the first place that the matter is concentrated in big clumps, the galaxies, varying in mass (*) from about 5×10^{11} for a few exceptionally massive ones to perhaps 10^7 for the smallest that have been observed. The lower limit is probably determined by observational limitations, in reality there may be large numbers of still less massive systems. We do not know how far the present concentration into galaxies reflects the unevenness in the distribution of matter in the universe at the epoch when the galaxies became separate units. For most galaxies must have contracted considerably since their segregation from the rest of the universe. Yet they can furnish important information on conditions at the time of their formation. In spherical galaxies the degree of concentration to the centre can teach us something about the amount of small-scale turbulent motion in the universe. Most galaxies deviate greatly from the spherical shape and have a considerable angular momentum. The total angular momentum must have been present in the primeval clump of material from which the galaxy has contracted. The angular momentum together with the strength of the concentration of mass towards the centre of a galaxy contain information on the proportion between *regular* large-scale rotation and the *irregular* currents in the part of the universe which contracted into that galaxy. So

* In this report all masses will be expressed in that of the sun as a unit.

far as I know, little or nothing has yet been done to outline the state of motion in the early universe from such data. At present all we can do is to state that most galaxies are endowed with considerable angular momentum, but that in general there is also a pronounced concentration of density toward the centre. The character of a galaxy would appear to be determined by the measure of orderliness of the large-scale currents in the clump of the universe from which it contracted. If this clump rotated approximately like a solid body, but of very irregular shape, the subsequent evolution of the galaxy would presumably be slow : it would contract little and remain irregular during a long time; it might form a system like the Magellanic Clouds. If, on the other hand, there was little ordered rotation to start with, the primeval gas would contract very much, resulting in a system of high central density. The orbits of the stars formed in it would have relatively short periods of revolution. We would thus obtain a condensed system of great regularity. This is what is shown by the elliptical galaxies. The spiral galaxies would be intermediate between the two extremes described.

On the basis of the above reasoning the masses and outer radii of irregular systems of the Magellanic-cloud type, or of the very late spirals, might give us an indication of the density in the universe at the time the galaxies were formed. Assuming that spheres with radii corresponding to the outer radii of these galaxies were in contact with each other, we would arrive at a density of the order of 10^{-25} g/cm^3; the uncertainty is of the order of a factor of ten. In considering this guess it should be remembered that it concerns those parts of the universe in which galaxies were formed. In view of the enormous unevenness in the distribution of galaxies the overall average density in the universe may have been rather lower. Evidently, the information we can obtain in this way is of limited significance.

The galaxies occur very often in groups, containing from two to several thousand members. In the case of a double or triple nebula, where the components often touch each other, the axes of rotation of the components do not usually show much relation. This would seem to indicate that the dimensions of the « rotational elements », or the diameters of the « currents » in the primeval universe were of the order of size of galaxies.

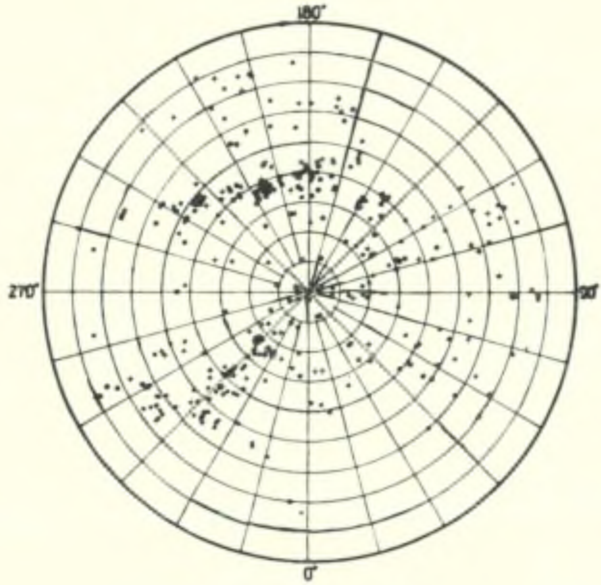


Fig. 1

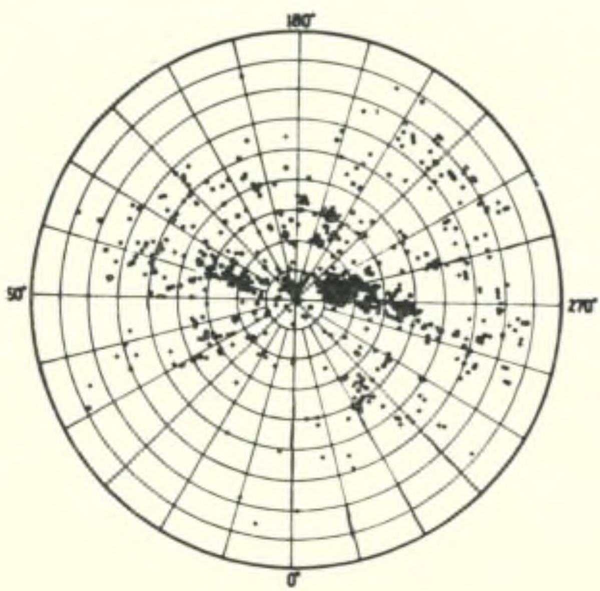


Fig. 2

2. Distribtion of galaxies within 40 million parsecs. The Virgo cluster.

In order to obtain insight into the character of the unevenness of the universe we consider first the distribution of galaxies brighter than about the 13th photographic magnitude. The distribution over the sky is shown in figures 1 and 2 taken from an investigation by Harlow Shapley and Adelaide Ames ([1]). The galactic poles are at the centres, the circles are at intervals of 10° galactic latitude; the numbers on the peripheries indicate galactic longitudes. If we provisionally adopt a value of 75 km/sec per 10^6 pc for Hubble's constant H (*) the radius of the volume surveyed in these pictures is about forty million parsecs, or hundred thirty million light years. *Some* of the striking unevenness in distribution is due to the effect of light-scattering by obscuring clouds in the Galactic System. These clouds are mainly confined to a thin layer around the galactic plane. They explain the practical absence of galaxies below 10° latitude and their scarcity between 10° and 20°. In the regions above 30° latitude the influence of irregular obscuration is probably small, and practically all the irregularities observed must be attributed to real unevenness in the distribution of galaxies in space.

The most outstanding feature is a dense concentration between about 72° and 80° northern galactic latitude and between 240° and 270° longitude, the so-called Virgo cluster of galaxies. The galaxies in this cluster are fairly uniformly distributed over a patch of about 7° diameter, but many galaxies outside this patch are evidently related to it. In particular it has a tail towards lower latitudes, which can be seen rather clearly down to $b = + 40°$, and may well extend still farther. The longish stretch of galaxies extending over the galactic pole to about 30° latitude on the other side must also be considered as being in some measure connected with the Virgo cluster, as it has much the same velocity and distance. The diagrams show several other large features, probably without direct relation to the Virgo cloud and its annexes, the most striking ones being two long stretches of galaxies in the southern galactic hemisphere. Within these generally denser regions, as well as elsewhere, we find subgroups of various sizes and populations. We also see large

* This same value will be used throughout the present report for obtaining distances.

spaces which are almost devoid of galaxies, the largest one extending over nearly one fourth of the northern galactic hemisphere.

Even a superficial inspection of figures 1 and 2 shows that the large majority of the galaxies in this part of the universe are linked together in larger or smaller « structures » and that there is no such thing as a regular « field » on which the structures we see are superimposed. This impression is confirmed when the third dimension is added, using radial velocities as distance indicators. In surveys of other parts of the universe much of the clearness of the Shapley-Ames picture is lost, because many structures are superimposed in surveys to larger distances. But all available evidence points to the conclusion that unevenness of the same type exists throughout the observable universe, and that the near-by region which we have been considering is in no way exceptional.

Some astronomers believe that the major part of the features which I have described forms some sort of superstructure, which they have called the Supergalaxy. Whether or not this is a useful concept I do not know. As far as total population is concerned, the space covered by Shapley and Miss Ames' survey appears to be fairly normal. There is also no clear deviation between the average expansional velocity of this region and that of the universe at large. I do not, therefore, see much reason to suppose that there exists a deeper relation between the features in this region than that which we have described and which is depicted in the plots published by Shapley and Ames.

What have the features considered above been in the past?

The central Virgo cluster presents an interesting problem. The concentration is so strong that it is plausible to assume that it has been a separate concentration at least since the epoch at which its galaxies were formed. It is difficult to imagine how it could have been brought together at a later date (*). Radial velocities are known for 35 galaxies in the central region of 7^0 diameter; they show an average residual velocity of 550 km/sec. Rough allowance for possible superimposed non-members would

* An alternative theory of clusters of galaxies has, however been proposed by Lemaître. It is described in his communication on p. 163

make the true internal velocity 520 km/sec; the distribution is practically Gaussian. If the velocity distribution is isotropic (*), the corresponding average internal space motion would be 1,040 km/sec, or 1.1×10^6 pc per 10^9 years. The distance of the cluster is 15×10^6 pc, and its radius 1.0×10^6 pc. If the gravitational attraction of the cluster on its membres were negligible it could not, therefore, have existed more than 900 million years. On various grounds it seems almost impossible to admit that it could be so young. To all appearances the galaxies in the cluster seem to be about as old as the expanding universe. We are thus led to conclude that it is held together by gravitation.

If we accept the theory that the universe has expanded from a considerably smaller radius, it is very unlikely that the Virgo cluster as such could have been formed with approximately its present size. If, for instance, one would contract the configuration shown in Shapley and Miss Ames' catalogue to the time that the radius of the universe was one fourth of its present value, leaving the central part of the Virgo cluster unaltered, most of the large-scale features we have noted would have been denser than this central part of the Virgo cluster. It would then be hard to understand why these other features would *not* have kept together by gravitation like the Virgo cluster. We must conclude that at the epoch of its formation the Virgo cluster must have been much smaller and that at that epoch it was endowed with similar expansional motion as the rest of the universe. Only, the local density excess must have been so large that the sum of the potential energy of the cluster when considered as a separate unit and the kinetic energy of the expansion plus the random motions was negative, so that the cluster could not expand to infinity.

The Virgo cluster furnishes important data bearing on the mass density in the universe. Judging by the measured internal motions the time of revolution at the outer boundary is 5×10^9 years. If the age is of the order of 9×10^9 years the cluster would have had time to order itself to some extent up to about this radius. This is in rough agreement with what we observe : the large deviations

* The fact that there is no strong central cencentration within the central Virgo cloud seems to indicate that there is no great preference for radial motions in this cloud.

from circular symmetry start outside this radius. As a working hypothesis we assume that in the more regular central part of the cluster there is an approach to dynamical equilibrium, so that the virial theorem may be applied to estimate the mass density. We find that the total mass within a radius of 1.0×10^6 pc ($3\frac{1}{2}^\circ$) is 5×10^{14} solar masses, while the mean density within this same sphere is 8×10^{-27} g/cm^3. If the mass is estimated from the observed galaxies and extrapolated to the fainter ones we obtain a total of only 0.18×10^{14}. The mass estimates of the galaxies are quite uncertain, and so is the extrapolation to the faint galaxies. Yet, the estimate indicates that, if the virial theorem holds, the mass of « intergalactic » stars or other matter would considerably exceed that of the galaxies. We shall see in section 4 that other, denser, clusters do *not* show this same discrepancy.

For the attempt to estimate the total mass density in the universe which will be made in the last section of this report, the partition between galaxies and intergalactic stars is irrelevant; all we need to know is the mass density corresponding to a given density of intrinsically bright galaxies. The Virgo cluster data may serve as a partial basis for estimating this ratio.

The Virgo cluster is certainly not the only structure among the bright galaxies that has maintained itself against the general expansion. The same must hold for most of the double and the multiple galaxies in this catalogue, as well as for several larger groups having space densities of the same order as the Virgo cluster.

3. *The large clusters of galaxies and their distribution. Large-scale homogeneity of the universe.*

The largest-scale unevenness of the distribution in the catalogue of Shapley and Miss Ames measures some twenty million pc. It is of interest to inquire whether there exists unevenness on a still larger scale. To investigate this we may consider the large survey which is being made at the Lick Observatory ([2]), ([3]).

This survey, which is now practically finished for the whole sky observable from Mt Hamilton, gives a wealth of information on the structure of the universe as well as on the absorption of light within our Galactic System. It extends to about $18^m.3$, i.e. roughly to a ten times greater distance than the Shapley-Ames catalogue, or to four hundred million parsecs.

Like the bright survey just discussed the Lick survey is strongly affected by galactic absorption. Though this is greatly reduced at latitudes above 40^0, some of the irregularity observed in the galactic polar caps may still be due to absorption within the Galactic System. However, there can be no doubt that above this latitude the more striking features are generally due to real variations in the space density of galaxies.

The picture shown by the survey is very complicated, as was to be expected when irregularities similar to those found in the Shapley-Ames catalogue are superimposed upon each other. The most striking features are large clusters of galaxies. Many of them are much richer than the Virgo cluster. These rich clusters differ from the Virgo cluster also in that they are more compact and generally show a strong central concentration. Often they are round and show great regularity in their central parts. An example of such a compact and rich cluster will be considered in some detail in section 4.

It was shown by Neyman, Scott and Shane ([4]) that the clustering tendency in this survey extends over distances of about 4^0. For larger distances there is in general no positive correlation between the counts. The catalogue, therefore, gives no evidence of unevenness on a very much larger scale than the irregularities observed in the Shapley-Ames catalogue. At the distance to which the survey extends 4^0 corresponds with about 30 million parsecs — a little less than the radius of the volume surveyed in figures 1 and 2.

The clusters in the Lick survey are often associated in small groups containing from two to four clusters, but the dimensions of these aggregates do not generally exceed the 4^0 indicated above.

The rich clusters themselves have been the object of two large surveys made on plates of the National Geographic Palomar Sky Survey, one by Zwicky and collaborators, the other by Abell. These extend to distances of about twice that reached in the Lick survey and lend themselves particularly well for investigating whether there is any very-large-scale structure in the universe.

At distances where effects of curvature of the universe may become sensible individual galaxies become so small that even with the largest instruments and under the best conditions the larger ones are only barely distinguishable from stars. Near this limit counts of galaxies must therefore become seriously incomplete. There appears to be no way in which the measure of incompleteness can be determined with sufficient accuracy.

The big clusters would seem to hold out a better hope for being discoverable with the same completeness at moderate and at very large distances. Their diameters, as found from the brighter members, are about 6 million pc. At a distance corresponding to a velocity of recession of c/2 the angular diameter would still be as much as 10′. A very great advantage of the clusters beside the fact that the completeness limits of surveys of clusters can probably be better defined than those for surveys of single galaxies, is that their diameters furnish a second measurable parameter in addition to the brightness of their members. For individual galaxies diameters are unmeasurable at distances relevant for the large-scale structure of the universe.

The above points have been stressed in particular by Zwicky ([5]), ([6]). With his collaborators he has searched for clusters of galaxies on large parts of the National Geographic Society-Palomar Observatory Sky Survey. He asserts that the category of what he calls « rich compact clusters » can be surveyed with the same completeness at moderate and at large distances, and that it forms a rather homogeneous group with respect to linear dimension. Many of the compact clusters are spherically symmetrical. Zwicky believes that these clusters are fairly evenly distributed over the observable universe. The apparent unevenness in their distribution over the sky, even in latitudes where galactic absorption effects must be fairly small, is attributed to effects of absorption in intergalactic space. This absorption, according to Zwicky, would be particularly strong within the large clusters themselves. The data published so far in support of these opinions are still rather meagre.

The average number of clusters per plate as recorded in this survey is roughly 35. As a Sky Survey plate covers about 40 square degrees, the total number for the whole sky above 30° galactic latitude would be of the order of 20,000. The mean distance of the most distant 25% of these clusters may be about 19^9 pc or 3×10^9 light-years (recessional velocity = c/4). The smallest clusters have diameters of about 3 minutes of arc.

An independent search for rich clusters has been made by Abell (7). He restricted his investigation to a smaller number of clusters (about 1/10th of that which will be included in Zwicky's survey) which enabled him to complete his search over the whole part of the Palomar Sky Survey that is not seriously affected by galactic obscuration. He took pains to specify as accurately as possible the criteria to be fulfilled for a cluster to be included in his catalogue, so as to make it reasonably homogeneous. The number of clusters selected to constitute a representative sample was 1,682. The limiting distance of his catalogue corresponds to a velocity of about 60,000 km/sec. Up to this distance the number of clusters increases with the 3rd power of the distance (as inferred from the photo-red magnitude of the tenth brightest member of each cluster). Abell comments on the extremely rapid way in which the number of clusters increases as we consider objects that are less and less rich. While there are 1,682 clusters in his list containing 50 or more galaxies with magnitudes in the interval between m_3 and $m_3 + 2$ (m_3 denoting the magnitude of the third brightest member), there are only 75 that have 130 or more members in this interval, and only 1 with 300 or more.

Abell made an extensive study of the unevenness of the cluster distribution. He found that the maximum deviations from a random (Poisson) distribution occur for fields with diameters of about 5°, or 60 million pc, in his more distant groups. This is of the same general order as what was found in the Lick survey from individual galaxies. It should be pointed out that the actual density variations observed in these big cells is not very large, and is still further reduced when we consider the largest volumes for which a significant comparison can be made from the available material. Table 1 gives the numbers of clusters in Abell's most distant group (average distance 700 million pc) in the regions above 60° galactic latitude.

For these regions the average number of clusters is 54, and the corrected average deviation from the mean is 16%. The average deviation for a random distribution would be 11%. Part of the deviation is undoubtedly due to galactic absorption; for instance

TABLE 1

Galactic longitude	Galactic latitude	
	$+60^\circ$ to $+90^\circ$	-60° to -90°
0° to 90°	60	67
90° to 180°	60	54
180° to 270°	46	
270° to 360°	37	

this is probably responsible for the low value in the quadrant 270^0 to 360^0. It appears, then, that on this scale (diameters of volumes about 350 million pc) there remains little, if any, real variation in mean density. As this volume is still small compared to the dimensions of the universe, we may conclude that all available evidence supports the concept of a universe which is homogeneous on a large scale.

It should be mentioned that a similar conclusion had already been obtained by Hubble 25 years ago from counts of individual nebulae down to about the same distance ([8]). The measurements of the velocities of recession of distant nebulae point to the same conclusion.

Like Shane and Wirtanen, Abell finds a number of cases where the clusters appear to form groups. The most striking agglomeration is one centred at 8^0 longitude and -66^0 latitude, which contains at least 10 clusters all at approximately the same distance in an area of $2^0 \times 3^0$, or in a space of 20 million pc diameter. There are several other less concentrated groupings which would seem to be real. It should be pointed out that Zwicky states that there are no real groupings among what he calls compact rich clusters. However, for larger cells the fluctuations in Abell's material are relatively slight, and the controversy is not too serious. It does not affect the above conclusions concerning the homogeneity of the universe.

Turning once more to the cluster phenomenon we may ask whether the majority of galaxies are members of clusters. Zwicky believes that most galaxies occur in large clusters, which he describes as « space fillers ». The best information may again be obtained from the sample of space covered by the Shapley-Ames catalogue. Most of the galaxies in this space may indeed be said to « belong » to a few large structures. But these structures are far from regular. The galaxies belonging to more or less regular and possibly stable groups, like the Virgo cluster proper and some smaller compact groups, form only a small fraction of the total. It might be expected that conditions in other parts of the universe would be similar, although some volumes of space contain clusters that are richer than the Virgo cluster.

4. The coma cluster

The Coma cluster may be considered as a typical specimen of the rich compact clusters of galaxies. Counts to various magnitudes have been made by several investigators, a.o. by Zwicky (5), by Shane and Wirtanen (2), and by Omer. Within about 1^{0} radius the cluster is fairly regular and strongly condensed towards the centre. Though there is some controversy about the outer parts, the various investigators agree that it extends to at least $1^{0}.7$. Zwicky has adduced evidence that the galaxies between $16^{m}.5$ and 19^{m} show a much wider distribution than the galaxies brighter than $16^{m}.5$. He indicates that there exists an excess density of faint galaxies around the Coma cluster up to distances of about 6^{0}. He has also shown that the density distribution corresponds with that of the equilibrium configuration given by an Emden sphere (9).

Radial velocities are available for 23 galaxies in the condensed central part. They show a Gaussian distribution with a dispersion of 1,000 km/sec.

Taking again $H = 75$ km/sec. per 10^{6} pc, the distance of the cluster is 89×10^{6} pc.

I have investigated anew whether the density distribution in the inner part (within $r = 1^{0}$, or 1.6 million pc) corresponds to what one would expect for a cluster in dynamical equilibrium under its own gravitation. The regularity and the strong central concentration of the cluster gives ground for the expectation that this inner part has indeed approached a steady state. The second and third columns of table 2 show the numbers of galaxies per square degree as used in these computations (the suffix *br* refers to galaxies brighter than $16^m.5$, while *f* refers to galaxies between $16^m.5$ and $19^m.0$). The data have been mainly taken from Zwicky; the central density is a compromise between that of Zwicky and that of Shane and Wirtanen. Beyond $r = 15'$ the surface density of the bright galaxies appears to vary as $r^{-1 \cdot 6}$, while that of the faint ones varies as $r^{-0 \cdot 9}$. Up to $r = 1^\circ$ there is no serious difference between the distributions from these two sources. The space densities can be easily computed from the surface densities. In order to obtain mass densities we need to know the average mass of a galaxy. I suppose that the mass of a galaxy is proportional to its luminosity and that, if mass and light are expressed in the mass and the luminosity of the sun as unit, the ratio of mass to light is 50 for elliptical and S*o* galaxies, 20 for S*a* and S*b*, and 7 for S*c* galaxies. These numbers are estimated from provisional rotation measurements of a few bright galaxies, and are subject to very great uncertainty. In the Coma cluster most of the galaxies are of types E or S*o*; I have provisionally adopted an average mass-to-light ratio of 50 for the galaxies in this cluster. The true ratio may possibly differ from this value by as

TABLE 2

Densities and gravitational potential in the Coma cluster.

r (′)	n_{br} (per sq. degree)	n_f	Δ_{br} (10^{-28} g/cm³)	Δ_f	$\Phi(r) - \Phi(0)$ (cm/sec)²	$\log \frac{\Delta_{br}(0)}{\Delta_{br}(r)}$
0	1,580	1,170	427	13	0	0
10	530	890	201	10	34×10^{14}	.328
15	350	690	78	7	61×10^{14}	.740
20	260	550	36.6	4.4	83×10^{14}	1.068
30	145	350	13.1	2.0	116×10^{14}	1.514
40	94	250	6.1	1.2	139×10^{14}	1.847
60	46	153	2.2	0.5	170×10^{14}	2.284

much as a factor 2. Using the luminosity distribution as derived by Holmberg for the galaxies in the Virgo cluster, I find that the average luminosity of a galaxy in the bright group is 1.2×10^{10}, in the faint group 0.09×10^{10}. The mean masses become 6×10^{11} and 0.5×10^{11} solar masses, respectively.

With these numbers we obtain mass densities as shown in columns 4 and 5; they are expressed in units of 10^{-28} g/cm^3. The contribution by the fainter group is small. It seems probable that the still fainter, unobserved galaxies will likewise contribute little. The values of the potentials $\Phi(r)$ minus $\Phi(0)$ computed from these densities are shown in the 6th column.

If we suppose the velocity distribution to follow Maxwell's law, with a mean square space velocity σ^2, the density in a steady state is given by :

$$\frac{\Delta(r)}{\Delta(0)} = e^{-3\{\Phi(r) - \Phi(0)\}/\sigma^2}. \tag{1}$$

This is the same formula as for an isothermal gas sphere. Comparing the values of $-\log_{10} \dfrac{\Delta(r)}{\Delta(0)}$ for the bright group, as shown in the last column of the table, with $\Phi(r) - \Phi(0)$ we see that the condition of the linear relation required by formula (1) is fulfilled very nicely, considering the uncertainties of the observations. This gives support to the hypothesis that the inner part of the cluster is approximately in dynamical equilibrium, thus confirming what Zwicky had found before. From this comparison we find $\sigma = 1{,}010$ km/sec; the velocity dispersion in one co-ordinate would then be 583 km/sec. This is remarkably close to the observed dispersion of 1,000 km/sec. In reality the agreement is probably still better, because it is likely that the velocity distribution in a cluster of galaxies deviates from a Maxwellian distribution in that the motions are probably preponderantly radial. It can be shown that, if these radial motions have a Gaussian distribution (and observations show this to be the case), the equation (1) will still apply to a good approximation. As the observed radial velocities are all of galaxies very near the centre of the cluster, the observed dispersion must then be practically that of the space motions in the cluster, and therefore be equal to σ.

For a galaxy whose maximum distance from the cluster centre is 1^0 and which moves radially, the time between two successive passages through this maximum distance is 4.6×10^9 years. If the age of the cluster is 9×10^9 years, there will have been a fair amount of mixing up to this distance. However, at distances of 5^0, up to which Zwicky has observed signs of an excess of faint galaxies, re-arrangement of galaxies since the birth of the cluster must have been practically negligible. All we can expect at these large distances is a highly irregular and expanding cloud of galaxies : a large unevenness in the universe, of which the Coma cluster forms a kind of core. The relation might be somewhat analogous to that between the Virgo cluster and its long extensions over the whole of the northern galactic hemisphere.

From the data given above the time of relaxation for the bright galaxies in the part within 30′ from the centre of the Coma cluster is found to be 56×10^9 years if the velocity distribution is isotropic. In case the motions are predominantly radial the relaxation time would be somewhat shorter, but in any case the exchange of energy between individual galaxies must have been relatively slight. It certainly cannot be responsible for the difference in distribution of bright and faint galaxies indicated by Zwicky's counts. Such differences probably reflect the initial conditions at the time the galaxies were born.

In the Coma cluster there seems, judging from the above analysis, to be little mass in addition to that contained in its bright galaxies. How do conditions in other compact clusters compare with this? According to the virial theorem the mass of a cluster is given by

$$\mathfrak{M} = 2\overline{v^2}\ R, \tag{2}$$

where $\overline{v^2}$ is the mean of the squares of the space velocities, and R is a sort of mean radius defined by :

$$R^{-1} = \overline{r_{ij}^{-1}}\ , \tag{3}$$

r_{ij} being the distance between two galaxies in the cluster ; the average to be taken over all pairs.

On the hypothesis that various compact clusters have the same type of density distribution as the Coma cluster the above formula will allow us to make rough comparisons of the mass-luminosity ratios. The data which are actually available are very scanty and have not been sufficiently discussed. As far as I have been able to make out from a superficial inspection, the ratios in three other clusters appear to be much the same as that in Coma. The only known outstanding difference is presented by the relatively loose and little concentrated Virgo cluster discussed in section 2.

5. The average density in the universe

Let $\Phi(M)$ represent the number of galaxies per unit of absolute magnitude per pc^3 averaged over so large a volume that the unevenness of the distribution has dropped out. If $A(m)$ is the number of galaxies per magnitude per square degree, again averaged over a sufficiently large volume, we have, remembering that

$$M = m + 5 - 5 \log r :$$

$$A(m) = 0.14 \times 10^{0.6m} \int_{-\infty}^{+\infty} 10^{-0.6M} \Phi(M) dM . \qquad (4)$$

$A(m)$ can be found from counts of galaxies. If, in addition, we know the *shape* of the luminosity function $\Phi(M)$, we can use (4) to determine the total number of galaxies of each absolute magnitude per cubic parsec.

Hubble ([10]) has derived the following expression for the average number of galaxies brighter than m :

$$\log N(m) = 0.6 (m - \Delta m) - 9.052. \qquad (5)$$

In this formula Δm is a correction depending on the red-shift. Neglecting this correction for the present rough estimate for $m = 18$, and neglecting similarly the fact that at each limiting magnitude a certain proportion of the galaxies must have remained undiscovered (these effects may to some extent be counterbalanced by a systematic error in the magnitude scale used by Hubble), we find for $m = 18$:

$$A(18) = 84. \qquad (6)$$

Holmberg ([11]) has derived the luminosity function in the Virgo cluster. Applying a correction of -0.9 to his absolute magnitudes in order to conform to the distance scale used in the present report, we find the results indicated in table 3 under $\log \Phi_H + k$; k is the

TABLE 3

Distribution of absolute magnitudes of galaxies.

M_{pg}	$\log \Phi_H + k$	$\log \Phi_{corr} + k'$	$10^{-0.6M} \Phi(M)$	Total lum.
— 22.0	0.00	—	—	—
— 21.0	0.70	— 0.07	$3.4 \times 10^{12-k'}$	2.0×10^{-11}
— 20.0	1.18	0.64	$4.4 \times 10^{12-k'}$	4.0×10^{-11}
— 19.0	1.47	1.14	$3.5 \times 10^{12-k'}$	5.1×10^{-11}
— 18.0	1.63	1.45	$1.8 \times 10^{12-k'}$	4.2×10^{-11}
— 17.0		1.62	$0.7 \times 10^{12-k'}$	2.4×10^{-11}
extrap.			$0.5 \times 10^{12-k'}$	4.2×10^{-11}

unknown constant to be added to reduce to an average cubic parsec in the universe.

We must first verify whether the $\Phi(M)$ found for the Virgo cluster satisfies the condition that it must yield the correct average distance for galaxies of a given apparent magnitude. This average distance is given by

$$\overline{r(m)} = 10^{0.2\,m+1} \frac{\int_{-22}^{+\infty} 10^{-0.8M}\ \Phi(M) dM}{\int_{-22}^{+\infty} 10^{-0.6M}\ \Phi(M) dM} \tag{7}$$

For $m = 18$ we obtain :

$$\overline{r(18)} = 5.9 \times 10^8 \text{ pc.} \tag{8}$$

Some extrapolation of $\Phi(M)$ to fainter absolute magnitudes had to be used. It is unlikely that this part of the luminosity distribution can contribute in an important manner to the integrals in (7). Even if we use the ever-increasing form for $\Phi(M)$ proposed by Zwicky (viz. $\Phi(M) = c \times 10^{0.2M}$) the contribution by the magnitudes fainter than those shown in table 3 is practically negligible.

The investigation by Humason, Mayall and Sandage ([12]) gives for « field » nebulae of the 18th magnitudə $v = 26{,}000$ km/sec, corresponding with $\bar{r} = 3.5 \times 10^8$ pc. Comparison with (8) shows that the distance computed by means of Holmberg's luminosity function is 1.67 times too large. It is not altogether surprising to find that Holmberg's Φ(M) for the central part of the Virgo cluster yields too large a proportion of intrinsically bright objects, for there appears to be a tendency for the brightest galaxies to be concentrated to the centre of this and other clusters. In order to bring Φ(M) into accordance with the field nebulae I have shifted it to fainter absolute magnitudes by an amount of $5 \log 1.67 = 1^m.1$. Although the form chosen for this correction is arbitrary, this may not have too much influence on the results for the average density. The values corrected in this manner are shown in table 3 under $\log \Phi_{corr} + k'$. The fourth column shows the values of the integrand in (4). Integrating (4) from -21 to $+\infty$ (using again Zwicky's expression to extrapolate to fainter M; cf. the line « extrap. » in table 3) we obtain :

$$A(m) = 2.00 \times 10^{12-k'+0.6m}. \tag{9}$$

Equating this to (6), we find $k' = 21.18$. The corresponding values of the total luminosity per cubic parsec in the universe, expressed in the luminosity of the sun as a unit, are in the last column of table 3. Extrapolating in the same way as above we find the average total luminosity per pc^3 in the universe to be 2.2×10^{-10} times that of the sun.

In order to compute the mass density we assume again that the mass of a galaxy is proportional to its light, and that the massto-light ratios are the same as those adopted in section 4. According to estimates by Hubble the relative proportions of elliptical galaxies, S*a* + S*b*, and S*c* + Irr are about 0.2, 0.4 and 0.4. These numbers refer to galaxies of a given apparent brightness and are probably approximately valid for the present computation, where the intrinsically bright galaxies are the determining factor. With this mixture the average value of the ratio of mass to light becomes 21, and the total mass density in the universe, in so far as it is contained in galaxies, is 4.6×10^{-9} solar masses per pc^3, or 3.1×10^{-31} g/cm^3. The real mean density may be higher, because of the existence of stars or matter distributed in between the galaxies. So far, the only way in which we can obtain information on this is by the dynamics of clusters of galaxies. The data at our disposal indicate that in some rich clusters the amount of intergalactic matter may be negligible. In the Virgo cluster, on the other hand, we found that a factor of approximately 25 had to be applied to the total mass of the galaxies in order to give the mass required by the virial theorem. Although it is true that the Virgo cluster is the least regular of the clusters considered, and doubt may be entertained as to the applicability of the virial theorem in this case, it seems nevertheless difficult to escape the conclusion that there must be very considerable additional mass. However, in view of the large discrepancy between the observed mass and that inferred from the motions, we shall have to reconsider whether the central Virgo cluster *is* really so concentrated a group as it *appears* to be. For the present it remains an open question whether or not a factor of the order of ten or more must still be applied to the density estimated above.

The principal other uncertainty affecting the determination of the density is in the assumed value of H. If the true value of H is $75/\varkappa$ all distances will be increased by a factor $\varkappa$. The density in the universe will then be multipled by a factor $\varkappa^{-2}$.

REFERENCES

(1) Harlow Shapley and Adelaide Ames, « A Survey of the External Galaxies Brighter than the Thirteenth Magnitude », *Harvard Obs. Annals*, **88**, No.2 (1932).

(2) C.D. Shane and C.A. Wirtanen, « The Distribution of Extragalactic Nebulae », *A. J.*, **59**, 285 (1954).

(3) C.D. Shane, « The Distribution of Extragalactic Nebulae II », *A. J.*, **61**, 292 (1956).

(4) J. Neyman, E.L. Scott and C.D. Shane, « On the Spatial Distribution of Galaxies. A Specific Model », *Ap. J.*, **117**, 92 (1953).

(5) F. Zwicky, « Morphological Astronomy », Berlin-Göttingen-Heidelberg (1957).

(6) F. Zwicky, « Statistics of Clusters of Galaxies », *Proc. Third Berkeley Symp. on Math. Stat. and Prob.*, III, 113 (1956).

(7) George O. Abell, « The Distribution of Rich Clusters of Galaxies », *Astrophysical Journal Supplements*, **3**, 211 (1958), No 31.

(8) E.P. Hubble, « The Distribution of Extragalactic Nebulae », *Ap. J.*, **79**, 8 (1934).

(9) F. Zwicky, « On the Clustering of Nebulae. I », *Ap. J.*, **95**, 555 (1942).

(10) E.P. Hubble, « Effects of Red Shifts on the Distribution of Nebulae », *Ap. J.*, **84**, 517 (1936), and « The Observational Approach to Cosmology », Oxford (1937).

(11) Erik Holmberg, « A Photographic Photometry of Extragalactic Nebulae », *Meddel. Lund Observ. Ser.* II, No. 136 (1958).

(12) M.L. Humason, N.U.Mayall and A.R. Sandage, « Redshifts and Magnitudes of Extragalactic Nebulae », *A. J.*, **61**, 97 (1956).

Discussion of Oort's Report

Dr Gold. — The upper limit of density of matter should also be mentioned. A uniform background would not show in any of the dynamical effects in clusters — those measure only the excess mass in the cluster over the background density.

The movement of galaxies, through the intergalactic gas, gives a limit, through the dynamical pressure that would be exerted onto the central sheet of galactic gas. This would suffer greater displacements from the gravitationally defined central plane than are in fact observed if densities of more than a few times 10^{-27} were the rule.

No spectroscopic observations would be more sensitive in the case of hydrogen, with only little impurities.

It seems to me that we are entitled to make our guess about the mean density of matter anywhere in the permitted range : between about 10^{-31} and 5×10^{-27}, and I cannot see that one end or the other of this range is more probable on observational grounds.

Dr Oort. — It is evident that a uniform background would not show up in the dynamical effects discussed. But I find it somewhat difficult to believe that the violent irregularities in the distribution of galaxies would not also be reflected in that of the intergalactic medium.

Dr Shapley. — I can supplement in some small points the presentation by Dr Oort.

1° It may be significant that the more distant clusters of galaxies have had 10^9 years less « time to organize themselves » than the nearby Virgo cluster, yet they are more compact and symmetrical.

2° The virial theorem cannot rigorously apply to the Virgo cluster because of its strong asymmetry = the rich « Centaurus tail » signathe ırregularities.

3° All types of galaxies are found in the Virgo cluster; it is otherwise in the Coma cluster, and possibly in other compact systems.

4° The Harvard results on the distribution of 100.000 galaxies give, for the space density : $m_t = 15,\ 16 \pm 0{,}02$, leading to a density of 50 galaxies per square degree rather than the 84 deduced by Oort.

Dr Oort. — I agree that the virial theorem does not strictly apply to such irregular systems. But, unless you want to assume that the Virgo galaxies are ten times younger than the universe, it has to apply *approximately*.

The Coma cluster is so nearby that its age is practically equal to that of the Virgo cloud.

Dr Schatzman. — Starting from the Neymann, Scott and Shane statistical analysis of the clustering of galaxies, I got, with the same light to mass ratio as yours, a space density 6 times greater as yours. The difference is significative.

Dr Oort. — The difference may come from the use of a different value of Hubble constant.

Dr Schatzman. — This does not appear possible, in as much as the star counts by Shane were calibrated by comparison with Hubble's counts and my density was based on the latter.

Including the change in Hubble constant, I still find a discrepancy by a factor 3.

Radio-Astronomical observations which may give information on the structure of the universe

by A. C. B. LOVELL

1. INTRODUCTION

In 1947 Bolton and Stanley discovered that the radio emission from the direction of Cygnus was emitted by a localised source. Subsequently Ryle and Smith found a more intense source in Cassiopeia. In the ensuing ten years the number of known radio sources has increased very rapidly and the existence of several thousand is now established. Opinion as to the nature of these radio sources, which are superimposed on a much more general distribution from the Galaxy, has fluctuated markedly during this time. At the beginning of this period it was believed that the phenomena were confined to the Galaxy but this view hat to be abandoned after the identification of M31 as a radio source by Hanbury Brown and Hazard in 1950, and the appreciation of the true nature of the Cygnus source as a remote extragalactic collision.

At present the following objects have been identified as radio sources in the Galaxy :

(a) The supernovae of 1054, 1572 and 1604;

(*b*) The peculiar gaseous nebulosities in Cassiopeia, Puppis, Gemini, Auriga and the Cygnus Loop ;

(*c*) About 15 emission nebulae of ionized hydrogen surrounding hot stars which are detectable as radio sources in the centimetre waveband.

Apart from these few cases the attempt to associate the bulk of the radio sources with Galactic objects, either on an individual or statistical basis, has failed. In this report we shall be concerned first of all with the known extragalactic radio objects and then with the problem of the unidentified radio sources in so far as they may give information about the structure of the universe.

2. THE KNOWN EXTRAGALACTIC RADIO SOURCES

(a) *Normal galaxies.* — About 16 normal galaxies have so far been identified as radio sources in the surveys made by Hanbury Brown and Hazard (1), and by Cambridge (2) in the northern hemisphere, and by Mills (3) in the southern hemisphere. The list includes 6 S*b*, 6 S*c*, 1 S*bc* and 3 of the Magellanic type. No normal elliptical galaxies have yet been identified as radio sources. The relation between the radio and photographic magnitudes of these nebulae is shown in figure 1. A constant ratio of radio to optical emission would give a line of slope unity. The line of slope unity which gives the best fit corresponds to $(m_R - m_p) = + 1.4$.

(b) *Abnormal extragalactic objects.* — In 1951 Baade and Minkowski (4) (5) identified the radio source in Cygnus with the collision of two spiral galaxies. Compared with the radio emission from

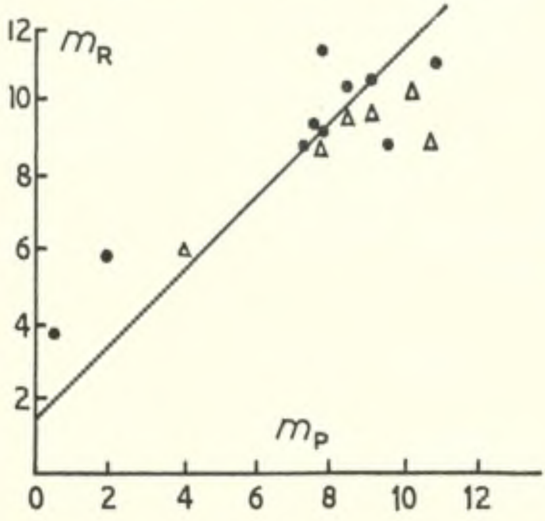


Fig. 1. — The relation between the radio and optical emission of all galaxies what have been identified as radio emitters.
● = Southern galaxy. — Δ = Northern galaxy.

normal galaxies the ratio of the radio to light emission is increased by a factor of 10^6. Thus, whereas the apparent photographic magnitude of the galaxies is $+18$ the radio source is the second strongest in the sky on metre wavelengths. The radio source in Perseus has also been identified with the object NGC 1275 considered by Minkowski to be a case of two galaxies in collision. Baade and Minkowski have also concluded that the radio source in Centaurus is associated with the object NGC 5128 which they believe to be another case of a galactic collision, one of which is a spiral seen edge-on. There are three other probable cases of radio sources associated with interacting galaxies. These are NGC 1316 associated with the radio source Fornax A (03S3A), and the radio sources Hercules A (16NOA) and Hydra A (09S1A).

One other abnormal object has been identified as a radio source. This is the Virgo galaxy M 87 (NGC 4486) associated with the radio source Virgo A (12N1A). This object is discussed in Section 8.

(c) *Clusters.* — It might be expected that clusters of galaxies would appear as identifiable radio sources at distances far exceeding those cut with their individual galaxies could be detected. The present position in this respect is as follows. In 1952 Hanbury Brown and Hazard detected the Perseus and Ursa Major II clusters as radio sources, but the radio emission is substantially more than might be expected from the aggregate of the emission from the individual galaxies. In the case of the Perseus cluster later measurements of the angular diameter of the radio source led to the conclusion that the majority of the radiation is associated with the particular object NGC 1275 ([7]) as mentioned in (*b*) above. In the southern hemisphere Mills surveyed the most probable clusters for radio emission but concluded that in only two cases was it reasonable to assume that a radio source was related to the cluster. These were Clusters I and II in Reticulum, both showing an abnormally high radio emission over that to be expected from the aggregate of the individual galaxies.

More recently the integrated emission from the Coma cluster has been studied ([8]) ([9]) and it seems likely that our knowledge of the radio emission from clusters will extend rapidly in the near future.

3. THE PROBLEM OF THE UNIDENTIFIED RADIO SOURCES

Apart from the relatively few radio sources identified either as galactic or extragalactic objects the basic observational data consist mainly of the number-intensity counts of the sources in the Cambridge (2) (10) and Sydney surveys (11). A certain amount of limited information is also available about the angular diameter of the sources. In one respect only is there any measure of agreement between these two major surveys, namely that the bulk of the unidentified sources is extragalactic. The evidence in favour of this view may be summarised as follows.

The surveys show that the radio sources with an appreciable angular extent are concentrated within a few degrees of the galactic equator. For example in the 1936 sources of the Cambridge survey 30 of large angular diameter (20′ to 120′) were within 5 degrees of the equator. The belief that these extended sources are galactic objects is strengthened by the identification of some of them with the diffuse galactic nebulosities Cassiopeia, Puppis, Gemini and Auriga.

On the other hand the remaining sources (1906 in the Cambridge survey) are of small angular diameter, of the order of one minute of arc or less, and show an isotropic distribution. Further, the few sources which have been identified at high galactic latitudes are associated with extragalactic objects. The possibility that the unidentified sources are galactic objects with an isotropic distribution must, of course, be examined for any particular survey. The only published case where the numbers are sufficient to receive a statistical treatment is the Cambridge survey and, as will be shown in Section 4, the authors have discussed this question. In view of the criticisms which have been made of this survey it is perhaps too early to state categorically that the unidentified sources of small angular diameter are extragalactic. The possibility that the Galaxy may contain large numbers of radio sources of star-like dimensions, as distinct from the already identified extended nebulosities, cannot be entirely dismissed until a satisfactory explanation has been found for the general background radio emission of the Galaxy.

4. THE ANALYSIS OF THE CAMBRIDGE SURVEY

Ryle and Scheuer ([10]) investigated the spatial distribution of the sources by two methods :

(*a*) The first method was originally used by Mills ([12]) and by Bolton, Stanley and Slee ([13]). This consists of plotting, for each of a number of areas of sky, log N against log I, where N is the number of sources per unit solid angle with an intensity greater than I. The slope of this curve can then be expected to give information about the variation of spatial density of the sources with distance. For example, if the objects responsible for the radio emission have a density of n_0 per unit volume in a universe which is static and Euclidean then :

$$N = \frac{4}{3}\pi n_0 \left(\frac{W_0}{I}\right)^{3/2}$$

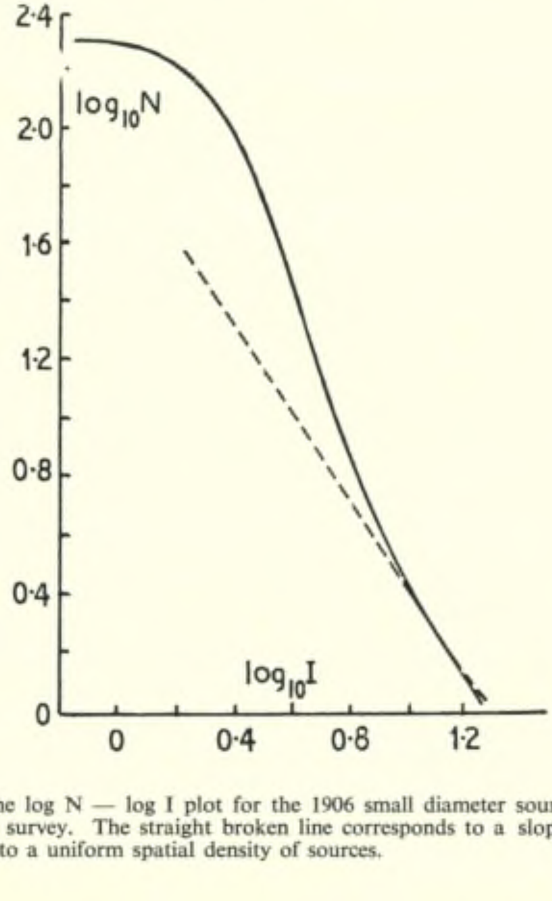


Fig. 2. — The log N — log I plot for the 1906 small diameter sources in the Cambridge survey. The straight broken line corresponds to a slope of —1.5 equivalent to a uniform spatial density of sources.

where W_0 is the power per unit bandwidth per steradian radiated by each source. In this case the slope of the log N — log I curve would be —1.5. The actual log N — log I curves for the 1906 sources of the Cambridge survey are shown in figure 2.

(*b*) Because of the limitations of the resolving power of the Cambridge interferometer, the confusion between adjacent sources may influence seriously the counts of the weaker sources. Scheuer ([14]) has shown that under these circumstances it is possible to obtain information about the distribution by measuring the magnitude D of the envelope of the interference pattern produced by the interferometer. In the absence of intense confusing sources the probability distribution P(D) of the envelope of a large number of weak sources would be a Rayleigh distribution. In the case of an actual record with intense sources deductions can be made about the radial distribution of the sources at great distances from the P(D) of the composite record. The curves of the probability distribution of the recorded traces are shown in figure 3, compared with the theoretical curves for a uniform spatial distribution of sources.

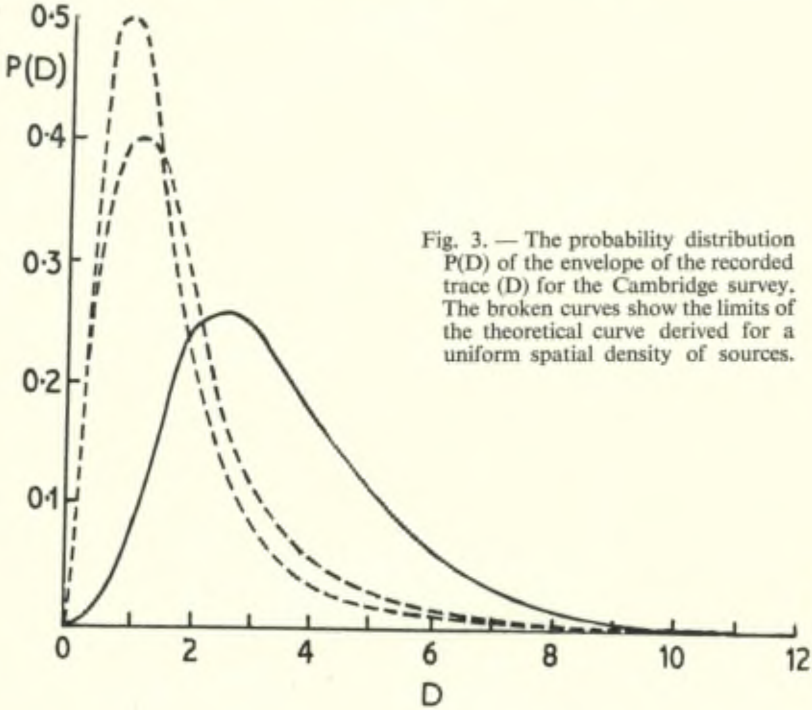


Fig. 3. — The probability distribution P(D) of the envelope of the recorded trace (D) for the Cambridge survey. The broken curves show the limits of the theoretical curve derived for a uniform spatial density of sources.

The conclusion to be drawn from both (*a*) and (*b*) is that the sources do not have an uniform spatial density but that the distribution is such that the spatial density, or the luminosity, increases progressively with distance. Ryle and Scheuer first consider the case where the average luminosity is independent of distance. They show that the observed probability curve P(D) can be used to derive the distance r_2 at which the spatial density ceases to increase and must remain constant or decrease. For this value they derive $r_2 = 1.3\, r_0$ where r_0 is the distance at which all sources can be resolved. r_0 was taken as the distance at which N = 125 and the flux = 2×10^{-25} watts $m^{-2}(c/s)^{-1}$.

On the assumption that the sources are responsible for the general galactic radiation it is shown that this value of $1.3\, r_0$ is 2,000 parsecs. Hence if the sources are in the Galaxy the solar system must be situated at the centre of a spherical region in which the density of the sources increases continuously outwards to radial distances of at least 2,000 p.s. On the other hand if the spatial density is constant and if it is assumed that the luminosity increases progressively with distance, then the minimum radius of the spherical region would be 600 p.s. Because these distances are small compared with the dimensions of the Galaxy the suggestion that the unidentified sources have a Galactic origin is discarded.

In terms of the identified extragalactic radio sources a model based on the emission from normal extragalactic nebulae is clearly out of the question in view of their low intensity, and an explanation must be sought in collisions or other peculiar objects in which the radio luminosity may be orders of magnitude more than that of the local galaxy. Other work suggests that the intensity of the integrated extragalactic radio emission corresponds to a brightness temperature of about 400°K (at $\lambda = 3.7$ m), and in terms of the emission from normal nebulae and collisions it is suggested that the value of r_2 must be at least 10^8 p.s.

On this basis Ryle and Scheuer consider that the implications of the Cambridge survey must be discussed on a cosmical scale. The steady state theories are immediately rejected because they imply that the average absolute luminosity and spatial density of radio sources is everywhere equal and unchanging with time and the slope of the log N — log I curve must always be less than 1.5.

An explanation is therefore sought in terms of evolutionary theories on which the finite size of the region of increasing apparent density would have a natural interpretation as corresponding to the earliest stages of galactic evolution, when the spatial densities of the nebulae were high, and hence collisions frequent. Thus the total number of colliding galaxies, and hence of detectable radio sources, must have been much greater in the past then at the present time. A slope in the log N — log I curve of greater than 1.5 would be a natural consequence, the precise value depending on the evolutionary model considered. It is easy to show that a collision of the Cygnus type gives an absolute radio luminosity of the order of magnitude required for the interpretation of the Cambridge log N — log I curve on this basis.

The local spatial density of these sources at the present time would be about 2×10^{-26} p.s.$^{-3}$, and the theory offers an immediate explanation for the difficulty of identifying radio sources since only a few would be within reach of the 200 inch telescope.

5. THE SYDNEY SURVEY

The survey of radio sources in the southern hemisphere has been made by the « Mills Cross » aerial ([15]), which produces a pencil beam of 50 minutes of arc. Compared with this the effective collecting area of the Cambridge interferometer is about 2 deg. $\times$ 16 deg. For comparison with the Cambridge survey Mills and Slee ([11]) have chosen an area bounded by declinations $+10^{o}$ and -20^{o} and by right ascensions 00^{h} and 08^{h} as common to the two surveys. In this common area the Cambridge list includes 227 sources and the Sydney list 383.

These catalogues are almost completely discordant. If the catalogues were completely uncorrelated with a random distribution of sources it can be shown that 42 chance coincidences would be expected. The actual number is 62 and hence a certain number represent genuine observations of real sources. By considering only sources of high flux density of approximately the same value in both catalogues Mills and Slee list only 12 examples which they consider to be an observation of the same physical object in both catalogues. It is therefore clear that instrumental effects must play a decisive effect in determining the positions and intensities of the sources in at least one of these surveys. Mills and Slee produce arguments to show that the errors are in the Cambridge catalogue because of the comparatively poor resolution and consequent confusion in the interferometer patterns. They consider that the level of reliability in the Sydney catalogue is at 8×10^{-26} w.m^{-2}(c/s)$^{-1}$, ten times lower than in the Cambridge catalogue. It is well known that at present the argument is unresolved, and it seems probable that other approaches to the problem will be needed before the differences can be settled. However, the Sydney group consider that their sources represent real concentrations of radio emission, the majority being physically discrete and have made a statistical analysis on this basis.

The comparison of this analysis with the Cambridge plots is given in figure 4. In favour of their own results Mills and Slee

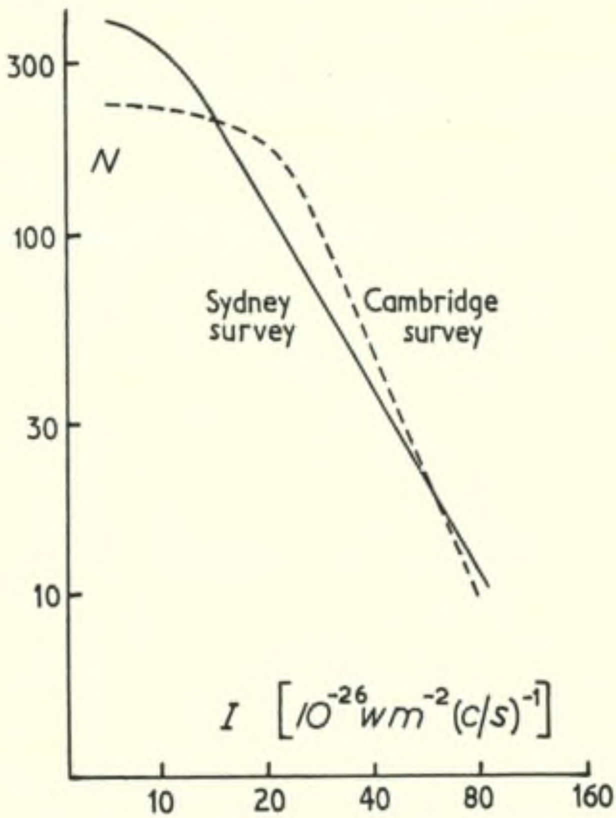


Fig. 4. — A comparison of the log N — log I plots for the Cambridge and Sydney surveys.

emphasise that the greater slope of the Cambridge plot is the result of an excess of sources with flux densities between 2×10^{-25} and 8×10^{-25} w.m^{-2}(c/s)$^{-1}$ where the Sydney catalogue is claimed to be reliable. Apart from the differences between the slopes of the two curves, the Sydney plot has a slope of —1.7 which seems to be significantly different from the slope of —1.5 which would apply if the sources were uniformly distributed in a static Euclidean universe.

The authors investigated the possible causes of this divergence. They first exclude the regions within 12½ deg. of the galactic plane which contains the concentration of strong, extended sources with a probable galactic origin. The log N — log I curve for high galactic latitudes is show in figure 5. This includes a point for the strong

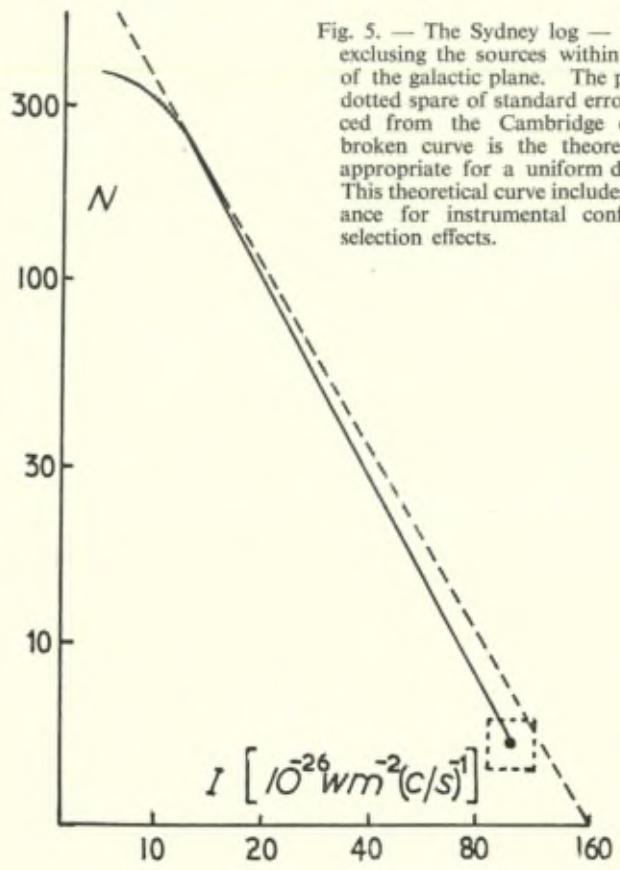


Fig. 5. — The Sydney log — log I plot exclusing the sources within 12½ deg of the galactic plane. The point in the dotted spare of standard errors is deduced from the Cambridge data. The broken curve is the theoretical curve appropriate for a uniform distribution. This theoretical curve includes the allowance for instrumental confusion and selection effects.

sources derived from the Cambridge data which the authors consider to be reliable. The slope is still three probable errors greater than —1.5, but an allowance for the instrumental confusion and selection effects gives the broken line of figure 5 as the source count expected for a uniform distribution of sources. It is considered that the small remaining differences in slope probably arise because the local space density of sources is below average, and that no special cosmological principles need to be invoked to explain the results of the survey. The authors also suggest that this remaining difference could arise from a significant clustering of the sources.

6. THE JODRELL BANK COMBINED PENCIL BEAM AND INTERFEROMETER SURVEY

The work described in sections 4 and 5 above reveals a disturbing state of affairs in which two carefully executed series of measurements give results which are quite discordant for the same region of sky. The conclusion that either one or both of the surveys bears little relation to the observable universe seems inescapable. The major differences between the surveys which might give rise to the discordant results are :

(i) Differences in fundamental technique, i.e., interferometer versus pencil beam;
(ii) Difference in overall beam widths of the radio telescopes;
(iii) Different zenith angle of the comparison area of the sky;
(iv) Different observers.

The crucial question is, of course, whether if (ii) (iii) and (iv) are eliminated the interferometer and pencil beam surveys are consistent. This question has been investigated at Jodrell Bank ([16] by using the 218 ft transit telescope both singly as a pencil beam instrument and in combination as an interferometer on a frequency of 92 Mc/s. The solid angle of the pencil beam was 7.5 square degrees and of the interferometer 12.5 square degrees. Thus (ii) is approximately satisfied compared with the differences between Sydney and Cambridge where the solid angles were 0.55 square degrees and 11.8 square degrees respectively (this allows for the fact that each half of the Cambridge interferometer consisted of two elements). The field of view was restricted by the limited beam shifts available on the transit telescope and covered a range of declinations from 26°N to 80°N. A solid angle of 4,250 square degrees was common to both surveys, and only sources in this region were considered in the comparison, which was made by the same observer. Thus points (iii) and (iv) are satisfied.

The analysis gave 83 sources in the pencil beam records and 104 sources in the interferometer records. There were in addition 9 extended sources in the pencil beam records which are not relevant to the present comparison. A direct comparison of the two lists gives 39 positional agreements within the limits of experimental error. 11 coincidences would be expected if the lists were completely random. 8 sources in both lists had flux densities greater than 100×10^{-26}w.m^{-2}(c/s)$^{-1}$ and these showed complete agreement; further comparison of flux densities on the two lists indicated that the 39 positional coincidences were probably genuine. Hence there were 41 sources on the pencil beam list and 61 on the interferometer list without mutual coincidences. The authors' conclusions with regard to these were as follows :

a) **The 41 sources in the pencil beam list without coincidence in the interferometer list.**

(i) A further inspection revealed that 5 sources were associated with interferometer traces originally rejected as doubtful. These were weak sources with flux densities less than 34×10^{-26}w.m^{-2}(c/s)$^{-1}$ These sources should therefore probably appear in the list of genuine coincidences.

(ii) 25 were in positions where the interferometer records were so confused that it was not possible to establish the identity of a given source.

(iii) 10 sources failed to appear on the interferometer records although above the confusion limit and well established on the pencil beam records. The rejection of these sources by the interferometer is probably a function of the diameter of the sources.

b) **The 61 sources in the interferometer list without coincidence in the pencil beam list.**

(i) 2 weak sources could be identified on the pencil beam records and should probably appear in the list of genuine coincidences.

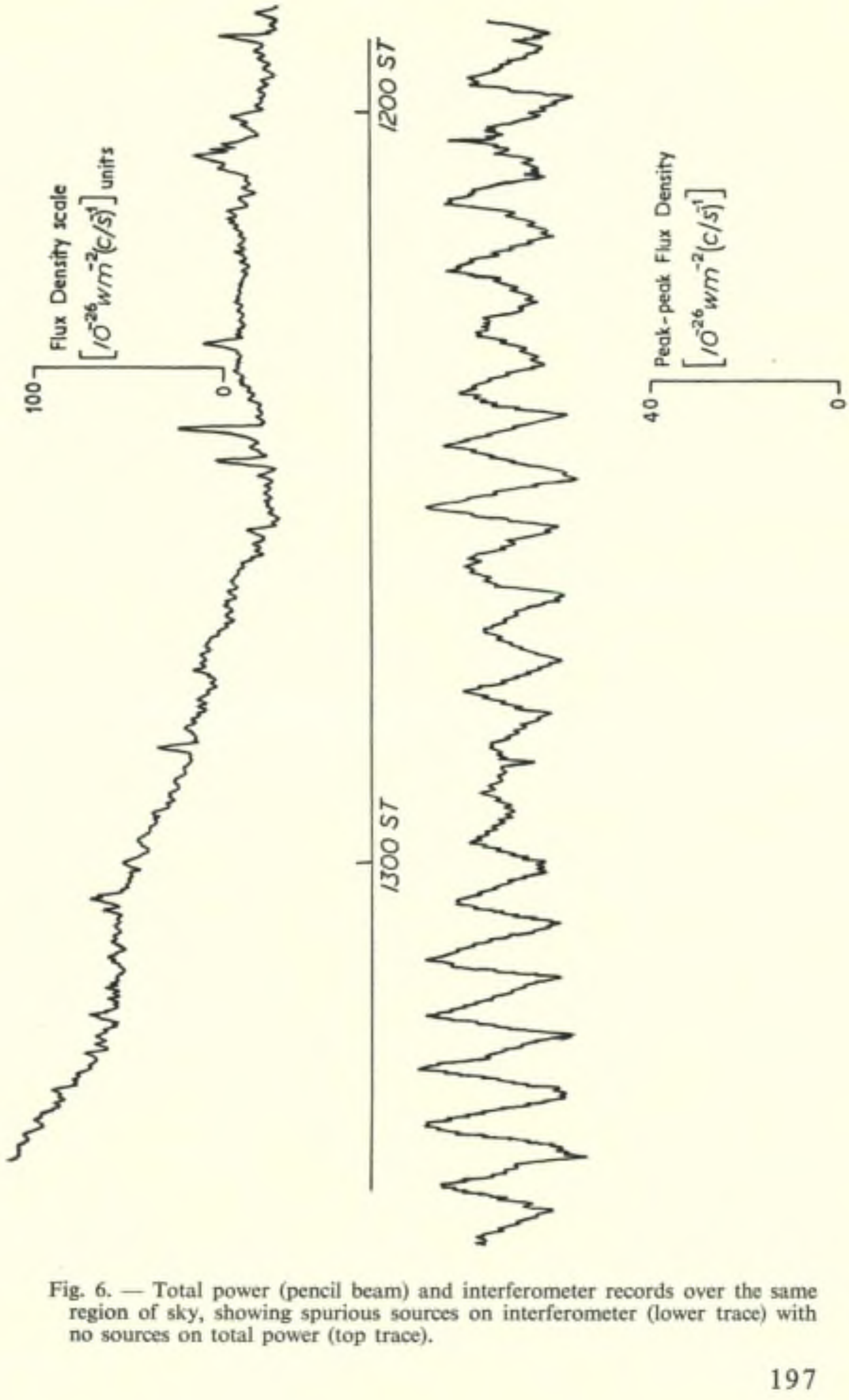


Fig. 6. — Total power (pencil beam) and interferometer records over the same region of sky, showing spurious sources on interferometer (lower trace) with no sources on total power (top trace).

A further 7 sources lie in confused regions and their coincidence cannot be excluded.

(ii) 14 sources lie in regions of the 9 extended sources and may have been confused on the pencil beam records.

(iii) 38 sources could not be detected on the pencil beam records, although the indicated flux densities were such that their appearance on the pencil beam records would be expected.

In respect of (*a*) and (*b*) the authors conclude that actual sources can occur on the pencil beam records and not on the interferometer either because of confusion effects or because they have large diameters. On the other hand the appearance of sources of appropriate flux density on the interferometer records and not on the pencil beam cannot be explained unless they are spurious, and, in fact, the authors conclude that at least the 38 sources in (*b*) (iii) are spurious.

A careful inspection of the records in the neighbourhood of the suspected spurious sources revealed that the more intense of these were positioned between two adjacent pencil beam sources where the latter record showed a minimum. The beating pattern between these sources had been interpreted as intense sources on the interferometer records. An illustration of this effect is shown in figure 6. A further investigation of the discrepancies showed that they occurred below 60×10^{-26}w.m^{-2}(c/s)$^{-1}$, and that the most important factor in determining confusion effects at a given flux density is the average number of sources per beam width at that level. In this case the interferometer with a beam of 12.5 square degrees became unreliable at a source density of 11 per 3,500 square degrees or one source per 25 beam widths.

This is an important conclusion because the figure is lower by 5 or 10 times than has been assumed in the past. When applied to the Cambridge and Sydney surveys the conclusion is :

(i) The source density of 1 per 25 beam widths occurs at a flux density of 56×10^{-26}w.m^{-2}(c/s)$^{-1}$ in the Cambridge survey. (This is, in fact, nearly the same as the figure of 50×10^{-26}w.m^{-2}(c/s)$^{-1}$, estimated by Mills and Slee as the reliability level of this survey.) The slope of the log N — log I curve below this level must therefore be considered to be very doubtful.

(ii) The same source density on the Sydney survey occurs at a level of 14×10^{-26}w.m^{-2}(c/s)$^{-1}$ and the slope of the log N — log I curve above this level can be considered to be genuine.

7. THE CAMBRIDGE SURVEY ON 159 Mc/s

Although the work described in Section 6 was too limited to give new information about the slope of the log N — log I curve it indicates rather definitely that interferometer surveys may give results which are considerably less reliable than pencil beam surveys where weak sources are involved. The dangers attending the free interpretation of interferometer records have been emphasised by the discrepancies between the original Cambridge survey and a more recent one ([17]) carried out by modifying the interferometer to work on 159 Mc/s instead of 81.5 Mc/s. In its new form the overall reception pattern covered 8 deg. $\times$ 1 deg. with interference fringes in two planes at right angles as before.

Of 143 sources in the original 81.5 Mc/s catalogue between declinations 40°N and 50°N there were only 35 coincidences in the new survey, indicating that the original level of reliability of flux densities for identification was set too low and should be increased to 40×10^{-26}w.m^{-2}(c/s)$^{-1}$ at 81.5 Mc/s. (This may be compared with the estimate in Section 6 that the Cambridge catalogue was unreliable below 56×10^{-26}w.m^{-2}(c/s)$^{-1}$).

Data have been given for two regions :

$$[+37^{\circ} < \text{Dec} < +52^{\circ},\ 00^{h} < \text{R.A.} < 24^{h}\ ;$$

and : $$-10^{\circ} < \text{Dec} < +10^{\circ},\ 00^{h} < \text{R.A.} < 08^{h}]\ ,$$

the second being common to the original Cambridge survey and the Sydney survey. The authors state that the agreement between their two surveys is no better than that between the original and the Sydney survey; but that better agreement is found between this new survey and the Sydney one. The measure of agreement as a function of flux density is indicated in the following table.

Flux density 10^{-26}w.m^{-2}(c/s)$^{-1}$	No. of sources in Sydney catalogue	No. of coincidences with new Cambridge survey	Expected number of random coincidences
I > 30	28	16	1.4
30 > I > 20	49	13	3.6
20 > I > 15	59	4	4.3
15 > I	71	6	5.4
Totals. . .	207	39	14.7

The log N — log I plots for this common region are shown in figure 7, which also gives the plot for the original 81.5 Mc/s survey

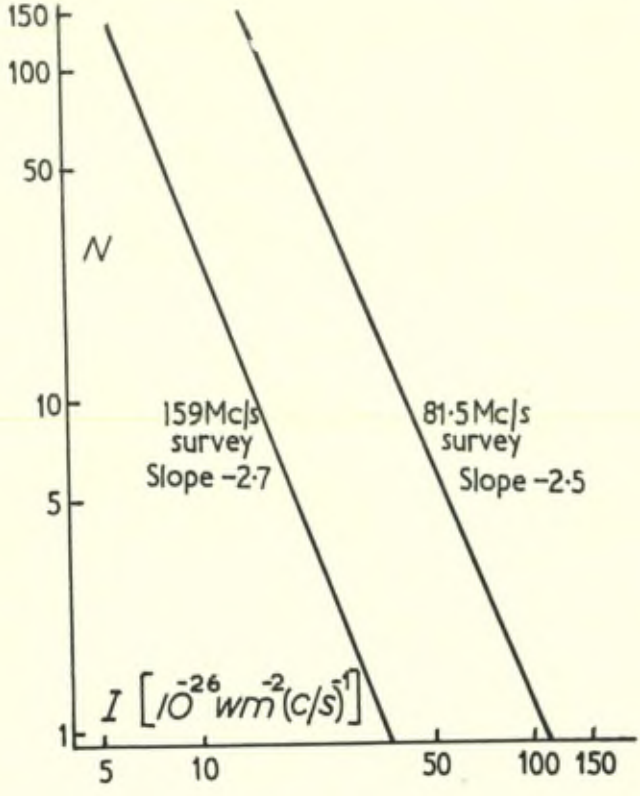


Fig. 7. — The log N — log I plots for the Cambridge surveys on 81.5 Mc/s and 159 Mc/s in a region common to the Sydney survey.

for this particular region. Whereas the original overall slope of the first survey was —3 the slope for this particular region was —2.5, compared with —2.7 in the new survey. The slope of the new survey in the other area analysed was —2.2. These results do not therefore alter the conclusion of the original survey, although the authors recognise that the statistical significance of the new data is not very good.

8. GENERAL DISCUSSION

It seems that the only safe conclusion to draw from the work described above is that, as yet, there are no radio astronomical observations which can influence significantly the existing views on the large scale structure of the universe. The Sydney survey which is open to least criticism yields results which can be explained without invoking any special cosmology. The more extensive survey in Sydney has not yet been published, but it is understood that it does not alter significantly the conclusions in Section 5.

At this stage it seems important to investigate in more detail the question as to whether the observations do, in fact, relate to processes on a cosmical scale. This question revolves around the nature of the unidentified radio sources and it may be useful to recapitulate the essential evidence.

(*a*) The sources have an overall isotropic distribution and are of small diameter (less than a few minutes of arc). The few established Galactic identifications are of large diameter and are associated with nebulosities concentrated in the plane. The circumstantial evidence is therefore in favour of an extragalactic origin for the unidentified sources. Unfortunately the nature of the background emission from the Galaxy is unresolved, and the possibility that this type of source may play some part in the more generally distributed components cannot be altogether excluded.

(*b*) The sources which have been identified at high galactic latitudes are associated with extragalactic objects. About 16 of these are normal galaxies, but from the radio-luminosity relationship it is clear that this type of object cannot figure prominently at the existing sensitivity limits of the radio surveys. If the radio sources are the result of processes at great distances, then special objects must be invoked with a very high radio-luminosity ratio.

(*c*) The possibility of a cosmical explanation for the radio surveys is therefore based on a limited knowledge of these few special objects which have been identified as radio sources, and a certain amount of evidence which indicates that some of the unidentified sources may be similar.

The further discussion of the question raised in (*a*) does not fall within the scope of this report. In respect of the positive evidence under (*c*) the identifications which have been summarised in Section 2 will now be discussed in more detail.

(i) **Colliding or interacting galaxies.**

(a) *The Cygnus collision.* — The identification of the second strongest radio source in the sky (19N4A — Cygnus A) with the colliding galaxies can be regarded as well established. An accurate position for the radio source was first determined by F.G. Smith [18] and measurements by Mills show agreement to within 1 second of R.A. and 1 minute of arc in declination. The photographs of this region by Baade and Minkowski [5] showed the unusual nebula with a double nucleus in this position. The spectrum was compatible with that of interstellar gas in a highly excited state, and the view of Baade and Minkowski that the object represents a face-on collision of two spiral galaxies is generally accepted. On the distance scale then extant, Baade and Minkowski estimated the distance to be 3×10^7 p.s. The association of the radio source with this object received further confirmation when Lilley and McClain [19] succeeded in measuring the hydrogen line red shift in 1955. The direct measurement of the frequency shift of the 21 cm line in emission is not possible with contemporary techniques; on the other hand, if the peripheral hydrogen gas is associated with the collision, then the continuous spectrum of the radio source should be absorbed in this gas at the displaced red shift frequency. Lilley and McClain observed this absorption at a frequency of about 1,340 Mc/secs, corresponding to a hydrogen line shift of 80 Mc/secs. This is compatible with the recession of 16,800 km/sec derived from the optical data.

The most detailed measurements of the structure of the radio source have been made by Jennison and Das Gupta [20], and by Jennison and Latham [21]. According to their measurements the source consists of two radiating centres each of size 45 sec $\times$ 35 sec, the spacing between the centres being 1 min 25 sec. The photographic image of the nebulae is within an elliptical area about 18 $\times$ 30 sec of arc, and hence the size of the radio source is considerably greater than might be expected. This result is not yet explained. Presumably the radio emission arises from the interaction of the gas in the two nebulae and it is not unreasonable that this should extend over a region which is large compared with the visible nucleus. The lack of emission from the central region where the collision would be expected to be most active remains a somewhat puzzling feature.

Estimates of the power output from the source are that optically (absolute magnitude —17.5) the total emission is 5.4×10^{35} watts or somewhat less than that of the Galaxy. The total radio emission can be estimated on the basis of an assumed spectrum as 4×10^{35} watts, which is surprisingly high in comparison with normal nebulae. On the basis of a relative velocity of 500 km/sec Baade and Minlowski have given the available total energy as 10^{52} joules, which for a full face-on collision might become available at the rate of 10^{37} or 10^{38} watts.

(b) *The Perseus collision NGC 1275.* — The identification of the radio source 03N4A was originally made with the Perseus cluster. Subsequently it was shown ([22]) that 75 % of the radiation from the source came from an area of 1 minute of arc and the remainder from an area of about 2 degrees. Minkowski ([23]) considers that the object responsible for the concentrated emission is NGC 1275 in the cluster which he identifies as a clear case of two galaxies in collision. The large diameter component is believed to be the aggregate radiation from the cluster. The identification is particularly important because the collision is the only known example suitable for a detailed optical investigation. The system consists of a tightly wound spiral of early-type with a velocity of + 5,200 km/sec and a strongly distorted late-type spiral with a velocity of + 8,000 km/sec. The spectral data give interesting details about the nature of the collision which is believed to have occurred at an angle of 15 degrees to 20 degrees between the galactic planes of the two systems. Minkowski estimates the duration of the collision as 10^6 years with a total energy emission of 10^{47} ergs, which is still a small fraction of the total kinetic energy available of 10^{49} to 10^{50} ergs where the relative velocities of collision are 3,000 km/sec.

(c) *The Centaurus collision NGC 5128.* — The identification of the radio source 13S4A with the object NGC 5128 appears to be satisfactory. Baade and Minkowski suggest that the system represents an interaction between an SO galaxy which is primarily responsible for the optical emission, and a late-type spiral seen edge-on which is responsible for the dark obscuring band. Emission lines compatible with excitation of interstellar gas lend some support to this view. The radio emission consists of a central localised source of diameter 3 min × 6½ min, surrounded by a large diameter source of size 5 deg × 3 deg. The distribution of the radio emission has not yet been satisfactorily associated with the optical features of the nebula ([24]).

(d) *The Hercules collision.* — The radio source 16NOA — Hercules A — is regarded by Minkowski ([23]) as a clear case of association with strongly interacting galaxies. The 200 inch photographs show a double galaxy with high excitation forbidden line emission. The red shift gives a distance of 1.4×10^8 p.s., which is 1.5 times the distance of the Cygnus collision. The existing measurements of the angular diameter of the radio source (2.5 min) show the typical excess over the size of the optical image (0.5 min).

(e) *The possible collision in Hydra.* — The radio source 09S1A — Hydra A, shows very good positional agreement with a faint double galaxy. On the other hand there are no indications from the spectra that a collision may be in progress. The angular sizes of the radio source and the galaxy are about 1½ minutes and ½ minute respectively. There appear to be certain radio abnormalities associated with this source ([25]) and the association cannot be considered as well established.

(ii) Unusual galaxies of unspecified type.

In addition to the 5 cases of collisions listed in (i) there are two cases of galaxies with abnormally high radio emission which are not believed to be cases of collision.

(a) *The Fornax galaxy NGC 1316.* — The discussion of this object and of its association with the radio source 03S3A (Fornax A) has been very confused. NGC 1316 was at one time believed to be similar to NGC 5128 but this suggestion has been dismissed by Minkowski ([23]), since there are no emission lines to indicate that any interaction is in process. It is believed to be an early-type galaxy with some absorption patches which suggest the beginning of spiral formation. The Sydney radio measurements reveal a large source of about 1 deg diameter whose centroid is coincident with NGC 1316. The present conclusion is that for unknown reasons NGC 1316 gives strongly enhanced radio emission without showing any obvious optical abnormalities.

(b) *The Virgo galaxy NGC 4486 (M 87).* — The identification of the radio source 12N1A — Virgo A — with M 87 is regarded as well established. The galaxy is elliptical, and has a jet predominantly blue in colour, close to the nucleus. The light from the jet is polarised but the radio source which has a diameter of 5 min × 2½ mins shows no evidence of direct association with this feature either in size or position angle. An important comment made by Minkowski (23) on this case is that the galaxy would be very difficult to recognise as one containing a peculiarity if seen from another direction or at a much greater distance.

(iii) **Clusters.**

The present position with regard to the radio emission from clusters of galaxies has been summarised in Section 2 (c) and there is little further which can be added pending further investigations. The point which is important for present consideration is whether the radio emission from clusters is enhanced by means other than increased probability of collision (NGC 1275 in the Perseus cluster is typical). For example the Coma cluster is now under study (8) (9). This shows a radio emission which is greater than would be expected from the aggregate emission of the individual galaxies. The issue which remains to be settled is whether this emission results from a particular object such as interacting galaxies as in the case of the Perseus cluster, or whether there is some other general feature of galactic associations giving enhanced emission. Until these questions are resolved it is difficult to speculate as to the extent to which the unidentified radio sources may be associated with remote clusters.

(iv) **The angular diameter of the unidentified sources.**

Apart from the extended sources with galactic concentrations the information about the angular diameters is not very precise. For example in the Cambridge survey it was possible to conclude only that the 1906 small diameter sources had diameters less than 20 minutes of arc, and that the diameters of 50 of these were less than 10 minutes of arc. The data for the Sydney sources have not been published but it is understood that the diameters of the majority of sources in that survey are 1 minute of arc or less.

The most detailed information about the diameters of a small number of the sources at galactic latitudes greater than $\pm$ 5 degrees, has been obtained by Morris, Palmer and Thompson [26]. One of the sources included was the Perseus source 03N4A; the diameter was found to be 2.4 min $\pm$ 0.5 min, consistent with previous measurements [22] and with the NGC 1275 identification. Of 4 other unidentified sources only one was resolved at the longest baseline used (10,600 λ). The diameter was found to be 4 min 20 sec $\pm$ 2 min. The other three were unresolved indicating that the diameters were less than 12 seconds of arc. These are the smallest diameter sources which have yet been reported. The intensity of these sources was of the order 50×10^{-26}w.m^{-2}(c/s)$^{-1}$ giving a brightness temperature of at least 2×10^{7} deg K. The brightness temperature of the Cygnus source is 10^{8} deg K at the same frequency (158 Mc/s) and its apparent photographic magnitude +17.9. If the three sources under discussion are of similar type with the same ratio of radio to light output as Cygnus, then the apparent photographic magnitude of the associated object would be + 23. Within the framework of the assumptions about the nature of the sources these measurements therefore support the idea that such objects would be very difficult to identify optically.

9. GENERAL CONCLUSIONS

(*a*) The existence of several thousand radio sources of small diameter (probably of the order of 1 min or less) is now established. There is general agreement that these have an overall isotropic distribution, but the surveys of their spatial distribution are discordant. It seems unlikely that the results of the present work will require any special cosmology for their interpretation.

(*b*) The more general question as to whether the unidentified sources are the result of processes on a cosmical scale must still be regarded as undecided. The positive evidence in favour depends on the association of 4 sources with interacting galaxies, together with one rather uncertain case and two sources with peculiar nebulae having a high ratio of radio-optical emission. The circumstantial evidence that the bulk of the unidentified sources are of this type arises (i) from the difficulty of explaining the isotropy in terms of galactic sources, (ii) the failure to find any galactic identification for the small diameter sources, (iii) the upper limits of 12 seconds of arc set to the diameter of 3 of the sources which is compatible with remote collision type sources beyond the limits of optical identification.

(*c*) Future progress should resolve these issues. The important requirements are :

(i) The measurement of the angular diameters and intensities of the individual sources. In all the existing identifications the radio source extends considerably beyond the optical image of the nebulae, moreover the spatial extent of the radio emission varies only between about 10^4 to 10^5 p.s. in the various sources. Thus the determination of the brightness temperature may be a significant factor in assessing the class and distance of the object involved.

(ii) The direct measurement of the distance of even a few of the sources by using the 21 cm line absorption technique to determine the radio red shift.

(iii) The evolution of new methods for determining larger numbers of sources without the confusion introduced by the existing techniques. There is evidence that the confusion limit is at about one source per twenty beam widths. This means that there are severe difficulties in extending the limits of the existing Sydney survey to much weaker sources.

REFERENCES

[1] C. Hazard, *Occ. Notes R.A.S.*, **3**, 74 (1954).

[2] J.R. Shakeshaft, M. Ryle, J.E. Baldwin, B. Elsmore and J.H. Thomson, *Mem. R.A.S.*, **67**, 97 (1955).

[3] B.Y. Mills, *Aust. Jr. Phys.*, **8**, 368 (1955).

[4] W. Baade and R. Minkowski, *Astrophys. Jr.*, **119**, 206 (1954).

[5] W. Baade and R. Minkowski, *Astrophys. Jr.*, **119**, 215 (1954).

[6] R. Hanbury Brown and C. Hazard, *Phil. Mag.*, **43**, 137 (1952).

[7] J.E. Baldwin and D.W. Dewhirst, *Nature*, **173**, 164 (1954).

[8] C.L. Seeger, G. Westerhout and R.G. Conway, *Astrophys. Jr*, **126**, 585 (1957).

[9] R. Hanbury Brown (unpublished work, Jodrell Bank).

[10] M. Ryle and P.A.G. Scheuer, *Proc. Roy. Soc. A*, **230**, 448 (1955).

[11] B.Y. Mills and O.B. Slee, *Aust. Jr. Phys.*, **10**, 162 (1957).

[12] B.Y. Mills, *Aust. Jr. Sci. Res.*, **A 5**, 266 (1952).

[13] J.G. Bolton, G.J. Stanley and O.B. Slee, *Aust. Jr. Phys.*, **7**, 110 (1954).

[14] P.A.G. Scheuer, *Proc. Camb. Phil. Soc.*, **53**, 764 (1957).

[15] B.Y. Mills and A.G. Little, *Aust. Jr. Phys.*, **6**, 272 (1953).

[16] C. Hazard and D. Walsh (in publication).

[17] D.O. Edge, P.A.G. Scheuer and J.R. Shakeshaft, *Mon. Not. Roy. Ast. Soc.* (in publication).

[18] F.G. Smith, *Nature*, **168**, 962 (1951).

[19] A.E. Lilley and E.F. McClain, N.R.L. Report 4689, *Astrophys. Jr.*, **122**, 361 (1955).

[20] R.C. Jennison and M.K. Das Gupta, *Phil. Mag.*, **1**, 55 (1956).

[21] R.C. Jennison and V. Latham (in publication).

[22] J.E. Baldwin and B. Elsmore, *Nature*, **173**, 818 (1954).

[23] B. Minkowski, I.A.U. Symposium No. 4 (Jodrell Bank) Cambridge University Press, p. 107 (1957).

[24] G. de Vaucouleurs and M.V. Sheridan, I.A.U. Symposium No. 4 (Jodrell Bank) Cambridge University Press, p. 169 (1957).

[25] J.G. Bolton and O.B. Slee, I.A.U. Symposium No. 4 (Jodrell Bank) Cambridge University Press, p. 174 (1957).

[26] D. Morris, H.P. Palmer and A.R. Thomson, *Observatory*, **77**, 103 (1957).

Discussion of Lovell's Report

Dr. Baade. — I would like to congratulate professor Lovell and his collaborators on their beautiful results about the radio emission of clusters of galaxies.

The localisation of strong radio sources within clusters of galaxies provides additional evidence that the strongest radio sources (like Cygnus A and NGG 1275) must be due to collisions or interactions of galaxies.

Dr. Gold. — The curve of log —N — Log I may be taken to be a cosmological piece of evidence only if either there is more identification of sources or if the plot differs from the 3/2 law.

Otherwise one is entitled to believe these sources to be extra-galactic, but not at very great distances. Ryle originally suggested the great distances only because of the difficulties of accounting for the steeper law, by anything other than a cosmological effect.

The great extend (such as 8° for M 31) of radio galaxies suggests there is a lot of gas background between the galaxies. Also perhaps the excessive radiation from the Coma cluster would be an indication in that direction.

Dr. Oort. — I am not quite so pessimistic as Gold with regard to the possibility that counts of radio sources may eventually give information on the large-scale structure of the universe.

In the first place there is the possibility of using diameter measures to sort out sources of high brightness temperature, like the sources of which Lovell has just told us.

In the second place, I can well imagine that it will become possible from the nearer sources to determine the frequency curve of the ratio of radio to optical brightness. Once one would know this frequency curve one could certainly interpret the counts of radio sources.

As regards the possibility of radio sources of an entirely unknown type I would mention that Minkowski told me that some of the sources in high latitudes have larger diameters and might well belong to the galactic corona; they might possibly be « knots » in the magnetic field.

As regards gas in intergalactic space, I do not think the radio observations of the clusters of galaxies present evidency of its existence.

Dr. Bondi. — May I ask Lovell what his view is of the method that was put forward by Scheuer and claimed to be an independent analysis of the Cambridge data, wholly insensitive to any confusion that might affect the direct count?

Dr. Van de Hulst. — Do the Iodrell Bank experiments reported by Lovell show that side lobe effects are essential in producing spurious sources? If so, the problem of the statistics gets very complicated and an analysis as Scheuer has made can hardly be expected to have given entirely trustworthy results.

Dr. Mac Crea. — Would you expect spurious sources to lie upon a smooth curve such as that given by Ryle?

Dr. Bragg. — Could you please explain in somewhat more detail how the spurious interference fringes arise by the mutual interaction of separate sources.

Dr. Lovell. — In the comparison of the pencil beam and interferometer records to which I referred, the spurious sources on the interferometer were positioned between two adjacent sources which were prominent on the pencil beam records. The interferometric pattern which was interpreted incorrectly as a source arose from the beating pattern of these adjacent sources. Side lobe effects will, of course, provide spurious effects on both interferometer and pencil beam equipments, but the primary effects in this particular experiment did not arise from side lobes, but from the finite extension of the overall reception pattern. The statistical treatment of the interferometer records by Scheuer's method should include this effect of blending. However, it does not include the effect of source diameter on the intensity, or clustering. The influence of these effects may be important. For example, recent work indicates that the radio size of the galaxies may be considerably greater than the optical image. The particular case of the Perseus cluster where 75 per cent of the radio emission is concentrated within 1 minute of arc provides a good example of the possible uncertainties in any statistical treatment. With a given resolution the manner in which such a cluster is counted in the records will depend on its distance in quite a different sense from the normal intensity/distance relationships. Until much more is known about the distance and nature of the individual radio sources there must be considerable uncertainty about the interpretation of any statistical treatment of the results.

III. EVOLUTION OF GALAXIES AND STARS

The oral report of Dr Baade was followed by this discussion :

Dr Oppenheimer. — Are the E-galaxies also composed — like the central disk of the Andromeda nebula — in part of metal-rich stars?

Dr Morgan. — Yes.

Dr Oppenheimer. — Is there a correlation for disk stars, between metal content and age?

Dr Baade. — We do not know, because we cannot determine with the present criteria accurately enough, the age *difference* between a cluster like M 67 and a cluster like M 3.

Dr Morgan. — The principal luminosity contribution of the giant ellipticals is from metal-rich stars.

Dr Hoyle. — There are a few points that have emerged from recent calculations which relate to Dr Baade's remarks on the Hertzsprung-Russell diagram. The first is that although at first sight it might seem as if M 3 should be younger than M 67 (since it lies higher in the diagram) this does not turn out to be so.

The second relates to cosmological considerations on the ages of star-groups and galaxies. The ages for M 67 and M 3 turn out to be almost disturbingly long : about 7 billion years for M 67 and about 9 billion years for M 3. Also on this matter of ages, there are *no* homologous conditions that can be applied over the main sequence. Massive stars appear to have lifetimes almost three times longer than homology would suggest. This is due to the development of very large convective cores.

Dr Sandage. — I should like to ask to Dr Hoyle what the physical reason is for the inversion in the order of age he assignes to M 67 and M 3. The observed break-off point in M 67 seems fainter than in M 3 whereas the age assigned to M 67, by Hoyle, is younger than that for M 3.

Dr Hoyle. — At corresponding stages of evolution, the M 3 stars burn more hydrogen than the M 67 stars. This is due to different starting conditions, lower metal content and lower Helium content for M 3.

Dr Ambarzumian. — What kind of stars give the visible radiation of giant E-galaxies? Are they stars with strong metallic lines or with weak metallic lines? What is known in this connection of the dwarf E-galaxies (M 32 for instance)? Perhaps Dr Morgan can answer these questions.

Dr Morgan. — The spectrum of M 32 indicates either that its light comes from metal-poor stars, or that the average spectral type of these stars is somewhat earlier than in the case of the giant elliptical systems. I do not yet have sufficient observational data to decide this point.

Dr Oort. — Dr Baade has referred to the two factors, age and composition, that determine the character of a star. It is likely that these two factors are not independent. The composition of a star depends on that of the interstellar medium from which it is formed and this in turn depends on the age. Dr Baade has pointed out that there are quite old objects in the galactic system which display almost no metal content. We must conclude that most of the change of composition in the interstellar medium took place in a rather early stage of the galactic system. As was discussed at a conference in Rome last year, star formation has probably beengoing on at a very much more rapid rate in an early stage of the system.

Dr Lovell. — Dr Baade has deduced from the Leiden hydrogen line measurements of neutral hydrogen content that the star formation in certain spiral galaxies must be nearly at an end. Is it possible to exclude on spectroscopic or dynamical grounds that significant amounts of hydrogen might exist in molecular or other forms?

Dr van de Hulst. — There does not seem to be an optical or radio test for the detection of hydrogen molecules. So the presence of a large proportion of molecular hydrogen cannot be excluded on the basis of observations. Dynamical theory, so far, has also failed to give a satisfactory answer to your question.

I should also like to answer to Dr Hoyle : the 21 cm radiation is indeed relatively low in some dark regions, but a good distinction between the interpretation that the hydrogen is molecular, or that the gas is much cooler and has a stronger selfabsorption, cannot yet be made.

Dr Hoyle. — Is it not possible that much hydrogen might be concentrated in small clouds? Once such a cloud becomes self-absorbing for 21 cm radiation it is possible to pack in more hydrogen without altering the observational effects.

Dr Mac Crea. — Two years ago, in Berkeley, Dr Morton Roberts and I developed a semi-empirical theory of star formation. At that time, it puzzled us that we required a total amount of interstellar matter of only 1 or 2 percent of the mass of the galaxy, whereas most astronomers would have expected about 50 percent. It is very satisfactory that Dr van de Hulst has since then shown that the low value is given by observation. This may be an indication that there is no large amount of interstellar gas that cannot be observed.

Dr Oort. — In connection with the question about the abundance of molecular hydrogen, would it not be possible to set at least an upper limit by a discussion of the disintegration of molecules that would be caused by collisions of interstellar clouds? This must be a powerful dissociation mechanism. The mechanism of formation of the H_2 molecule may be thought of as taking place by the condensation of H on to solid particles and the evaporation of molecules from them.

Dr Lovell. — In connection with Dr Oort's remarks would it be possible to deduce from other radio arguments and measurements of line spectra e.g. of deuterium, the probable amount of molecular hydrogen?

Dr van de Hulst. — Is it not strange that a young cluster as Persei has had recent star formation and yet is in region remote (5° in latitude) from a spiral arm and in a region where there is no gas at all?

Dr Baade. — I did not relalize that *h* and chi Persei is another young cluster without detectable remnants of dust and gas. The case which has me puzzled for some time is the Canis majoris cluster which consists of O- and early B-stars and which seems to be similarly free from dust and gas. It may be significant that both clusters are rich in O- and B-stars of high luminosity. Perhaps the remaining gas and dust is rapidly driven out of such a cluster.

Dr Gold. — I should like to make two comments :

1. Gas galaxies with no stars in them are a possibility, and radio investigations may reveal them. If they exist, they may be responsible for stripping other galaxies of gas, as Dr Baade mentioned.

2. Is there any discrepancy of metal abundance as deduced spectroscopically and from the H.R. diagram of clusters?

Such a discrepancy would show the addition of gas to the surface where it would only affect the spectrum and not the structure. Such an observation would thus be of interest in deciding on the importance of the acquisition of gas by stars.

Dr Baade. — I am afraid that our present data are not accurate enough to establish significant discrepancies of the kind Dr Gold has in mind.

Radio Investigations of the Galactic System and Near-by Galaxies; Evidence for Magnetic Fields in Galaxies

by **H. C. van de HULST**

1. INTRODUCTION

In this report « The Galactic System » will mean our own galaxy, i.e. the stellar system one of whose stars is the Sun. « A Galaxy » will mean any other stellar system of roughly comparable size and mass. The masses of galaxies appear to range from 3×10^{11} to 5×10^{7} solar masses (Table 1). The Galactic System ranks among the larger and heavier ones.

An intermediate or late-type galaxy like our own has its main light and mass strongly concentrated to an equatorial plane and rotates around an axis perpendicular to this plane. We employ cylinder coordinates z = distance from plane, R = distance from the axis. The symbol r is reserved for the distance from the Sun. Other population groups in the system may extend to very high values of z and have a less flattened or nearly spherical distribution.

Omitting finer distinctions we call the strongly flattened system the *disk* and the less flattened or nearly spherical system the *halo*. The rough division of observational data that may be obtained by radio-astronomical methods, is as follows :

Disk. — *Neutral hydrogen* by 21-cm line, giving mass, distribution, rotation and other motions.

Ionized hydrogen by local patches (HII regions) in emission at decimeter waves and in absorption at meter waves and further by the general distribution of the thermal component of galactic radio emission.

Halo. — A radiation of unknown origin called *non-thermal* galactic radio emission, tentatively identified with *high-energy electrons* radiating in a magnetic field (synchrotron mechanism).

The division thus made is not quite strict. As will be shown below, neutral hydrogen may occur also in the halo and, conversely, it is fairly certain now, that the non-thermal radiation shows a fair concentration towards the disk.

2. THE NEUTRAL-HYDROGEN DISK POPULATION FOR $R > 3$ kpc

The most prominent discovery in this area is the existence of a large number of spiral arms, thus confirming speculations of at least a century ago, that the Galactic System might be a spiral nebula. Pieces of three spiral arms in our immediate vicinity were mapped by the Yerkes optical astronomers in 1952 ([1]). Maps extending over the entire disk were made by 21-cm observations in Holland for the northern part and in Australia for the southern part. A paper comprising the results from both hemispheres about « the Galactic System as a spiral nebula » has just been prepared for publication ([2]). The main conclusions from this report with some additional comments are cited below. For the full method, tables and graphs, see the five papers in B.A.N. No. 475 ([3]).

The mass of atomic hydrogen in the Galactic System is 0.15 $\times 10^{10}\, M_{\odot}$, or 1 to 2 per cent of its total mass. It is distributed in a disk that is both extremely thin and extremely flat. The thickness between the points where half the maximum density is reached is about 220 pc, between $R = 4$ and 9 kpc. Systematic deviations from one plane remain below 75 pc for $3 < R < 10$ kpc, which shows that there has been sufficient exchange of angular momentum between the inner and outer parts to make their vectors coincide in direction. Only the outermost spiral arms, beyond $R = 10$ kpc, show deviations up to 400 pc, curving up on several sides like the rim of a dry leaf.

The number densities of hydrogen atoms in the plane of symmetry range from about 2 cm^{-3} in the heaviest parts of spiral arms to something probably less than 0.1 cm^{-3} in the « empty » spaces between spiral arms. These numbers refer to the finest details that have been resolved in these surveys with small telescopes (7 meter diameter in Kootwijk, 11 meter in Sydney). A representative estimate of the region of space lumped together by the limited resolving power is $150 \times 150 \times 600$ pc.

The greatest uncertainty is in the coordinate along the line of sight. It will not be resolved by using bigger telescopes as it arises from the basic uncertainties in the conversion :

frequency shift $\rightarrow$ velocity in line of sight $\rightarrow$ distance.

So far, all reductions have been based on the assumption that the line-of-sight velocity consists of the effect of regular rotation with an angular velocity ω which is a function of R and random cloud motions of the order of 6 km/sec. There is some evidence of local deviations from these assumptions. They may quite well have distorted the present maps at several places. However, large-scale systematic deviations, which would, for instance, occur if the actual motion were a spiral motion consisting of rotation and expansion on the order of 10 km/sec, may be excluded for $R > 3$ kpc on the basis of the observations.

There is a peculiar inconsistency in this method. At one hand we know that spiral arms probably originate by effects of gas dynamics or magnetic forces and not by gravitational forces alone. On the other hand, most of our actual knowledge about the spiral arms is based on the 21-cm maps, reduced with the assumption of undisturbed circular orbits. It is evidently desirable to investigate to what extent magnetic forces may force parts of spiral arms to move at angular velocities considerably above the rotation speed at the given value of R. Unpublished calculations by D. Wentzel show that such effects may not be entirely negligible. For example, the different angle of the Orion arm with the galactic radius, as mapped by stellar associations or by atomic hydrogen might find an explanation from this effect. The situation is not yet clear.

If the density is averaged over even larger areas of the plane so as to wash out the spiral arms, a maximum density about 0.7 cm^{-3}

from 5 to 8 kpc is found. The density may drop to half this value inside ($R < 3$ kpc) and gradually drops to zero outside. The outermost trace of hydrogen is found at $R = 15$ to 20 kpc.

3. ATOMIC HYDROGEN IN OTHER GALAXIES

Atomic hydrogen emission from beyond the Galactic System has first been found in the Magellanic clouds ([4]). The angular resolving power of the Sydney telescope ($1^{\circ}.5$) was quite adequate for a first map of the densities and line-of-sight velocities in these two galaxies. The erection of the Harvard 18 meter and Dwingeloo 25 meter telescopes opened the possibility of observing also a dozen or more other galaxies contained in the « local group ». The beamwidth of the Dwingeloo telescope at 21 cm is $0^{\circ}.57$ between half-power points. In addition to directivity a high stability and low noise figure of the receiver is required for results of some precision, as the antenna temperatures often are only a few degrees Kelvin.

All measurements were made relative to a comparison field at a « safe » distance from the nebula. It has been questioned if a distance of 4° as used for M 31 was really safe. Recent comparisons between the comparison fields of M 31 and M 33 showed a difference of $+ 0.1 \pm 0.1$ °K in the frequency range corresponding to velocities $+ 150$ to 300 km/sec. These data set plausible upper limits to neutral hydrogen radiation from the haloes of M 31 and M 33 and to possible irregularities in any neutral-hydrogen halo of the Galactic System.

Table 1 summarizes the most relevant data obtained in Dwingeloo together with the Sydney data on the Magellanic clouds. The measurements on M 31 have been published in full ([5]); the other data are preliminary data or estimates supplied by Miss L. Volders on the basis of an incomplete reduction. The Harvard measurements ([6]), which tend to show larger extensions and higher antenna temperatures, would seem not numerous or not accurate enough to be quite comparable. At the present stage, in the middle of an extensive investigation no detailed references will be given. The main conclusions apparent from the present material are the following :

There are six systems with well-determined hydrogen masses. These masses range from 0.4 to 2.5×10^9 ⊙. The types are Irr,

TABLE 1

Extragalactic Systems observed for 21-cm emission at Dwingeloo (April 1958).

Three systems at top serve for comparison

GENERAL DATA				RADIO MEASUREMENTS				MASS		
System	Type and cosec. i	Adopted distance (kpc)	Vel. of centre rel. to LSR (km/sec)	Time of measuring (hours)	Radio size (°)	Max. T_{ant} (°K)	Max. n ⊥ plane of nebula ($cm^{-3}kpc$)	Hydr. mass 10^9 ☉	Total mass 10^9 ☉	Ratio hydrogen to total
Galactic system	Sb	—	—	—	—	125	0.3	1.5	100	0.015
LMC	Irr	50	+ 295	—	15	30		0.7	1	0.7
SMC	Irr	50	+ 172	—	9	30		0.5	1.3	0.4
M 31 = NGC 224	Sb; 4	500	— 296	320	5	7.2	0.3	2.5	268	0.009
M 32 = NGC 221	E 2	500	— 200	11		< 0.2		< 0.04	16	< 0.0025
NGC 205	SBo	500	— 236	9		< 0.5		< 0.1	20	< 0.005
M 33 = NGC 598	Sc; 1.8	460	— 176	190	1.5	6.3	0.6	0.45	13	0.035
M 51 = NGC 5194	Sc	1200	+ 449	2		0.5				
M 81 = NGC 3031	Sb; 1.6	2600	— 45	50	> 0.3	0.5			97	
M 101 = NGC 5457	Sc; 1.1	1100	+ 260	230	1.25	1.5		0.4		
NGC 253	Sc	1300	— 88	25		0.4		> 0.07		
NGC 4236	Sc	1000x	+ 33	100		0.4		> 0.09x^2		
IC 1613	Irr	460	— 242	90	> 0.5	1.0		> 0.02	0.05	0.4?

Sb, Sc. The masses are systematically too small if a strong self-absorption in the 21-cm line occurs. However, unless a good proportion of the mass is contained in very dense or very cool clouds, this should not change the figures very much. A small correction to M 31 and somewhat higher correction for the Galactic System have been applied in the reduction.

The total masses have been computed from models based on the observed rotation curves of M 31 and the Galactic System by M. Schmidt, of M 33 by Miss L. Volders. Estimates of the total masses of the other systems are taken from an unpublished Table of Oort. The estimates for the smaller systems are based on an assumed mass/luminosity ratio. In the last published comparison, the optical rotation curve of M 31 was of lower accuracy and systematically deviating from its radio rotation curve. It appears that more recent optical measurements made by Mayall at Mt. Wilson remove the discrepancy.

The accuracy of one measuring point with a bandwidth of 150 kc/s (taking one hour) is about 0.15 °K in antenna temperature. On the systems that now have a measurable radio size, this figure will not be improved by using a bigger radio telescope, unless there are strong details that are now below the limit of resolution. In receiver development, only a maser can be hoped to give a good factor improvement in this mean error; not much more than a factor 5 may be hoped for.

Radiation at the position and expected frequency of the companions of M 31 is just measurable but the uncertainty in estimating what M 31 itself gives at this position and frequency makes it impossible to state with certainty that radiation from M 32 and NGC 205 has been measured. The elliptic system can contain at most 1/4 % interstellar hydrogen.

The three systems at the bottom of Table 1 have not yet been explored far enough to find how far they extend. A definite lower limit to the hydrogen mass can already be given. If the estimated total mass of IC 1613 is correct, this irregular system appears to have about half its mass in the form of interstellar hydrogen, like the Magellanic clouds. The Sc nebula M 33 has relatively more hydrogen than the Sb nebulae, as would be expected, but does not nearly approach the percentage found in the irregulars. The radiation from M 81 is too weak and its frequency too close to that of galactic hydrogen to give a reliable mass estimate.

The best hopes for finding atomic hydrogen in even more distant objects was to find it as intergalactic hydrogen within the confines of clusters of Galaxies. Heeschen (7) indeed reported positive evidence of the 21-cm line from the Coma cluster, with maximum antenna temperatures of 1.8 °K. More recent measurements by C.A. Muller at Dwingeloo give a zero result with the mean error 0.1 °K. Also the reported absorption of the strongly displaced and broadened 21-cm line in the spectrum of Cyg A(8) should be confirmed by other measurements before becoming the subject of theoretical conclusions.

4. EXPANDING GAS NEAR THE GALACTIC CENTRE

The most surprising result so far obtained with the 25-metre radio telescope at Dwingeloo was the existence of expanding spiral arms tn the region near the galactic centre. This was noticed first (9) by a clear absorption line appearing at velocity — 53 km/sec in the spectrum of the source at the galactic centre (Sgr A). It was verified that the frequency of this absorption effect formed a continuous sequence with that of emission peaks at neighbouring longitudes. We thus must visualize a continuous mass of gas (briefly : an arm) that is passing between us and the centre. As a purely rotational motion gives zero velocity in the line of sight towards the centre, the 50 km/sec must be interpreted as expanding motion from the centre.

This arm is the most prominent feature of long wings of the line profiles visible at all longitudes within 25° from the centre direction. These long wings, initially described as turbulence in the central regions, may be due completely to such expansion. By the longitude dependence an outer limit of $R = 3$ kpc is set; the Australian observations hint at a turning longitude that puts the spiral arm just described close to this outer limit. The rotation velocity at this distance would be about 200 km/sec.

A long programme of observations by Mr. W. Rougoor is now in progress at Dwingeloo. Instrumental provisions by Mr. Muller have made the zero line quite straight and stable so that these extremely wide lines, with no certain drops to zero, yet allow measurements within $\pm$ 1 °K. The programme so far has been mainly confined to longitudes 318-334° for negative velocities (approach to us) and 322-354° for positive velocities (recession from us). The width of the layer in latitude is about 1°.5 = 200 pc.

A first striking result from the latitude dependence at various frequencies through the centre source is the presence of absorption at all negative velocities and its absence at the positive velocities. This confirms the assumption that the high-velocity wings are due to expansion. The depth of the absorption dip drops from about 40 °K in the arm at — 50 km/sec to 10 °K at — 100 km/sec and becomes < 1 °K at — 150 km/sec. The original authors estimate an optical depth $\tau = 0.5$ and an integrated number of atoms of 1.5×10^{21} cm^{-2} across the arm. This is smaller by a factor 3 than found in the heavy arms near the Sun.

The dynamical consequences of this expanding motion and its source of energy have not yet been explored. If we look at this phenomenon as a one-way motion, the travel time from the nucleus to $R = 3$ kpc is of the order of 3×10^7 years and the nucleus has to furnish new kinetic energy at a rate comparable to the energy needed to maintain cosmic rays in the Galactic System. No such source of energy in the nuclear region has yet been suggested. On the other hand, the gravitational effects may be quite important. If the present models of the mass distribution in the Galactic System are anywhere near correct, the velocity of escape at R = 2 kpc is of the order of 500 km/sec. So the observed hydrogen clouds, if subject to gravitational forces only, will move at most 1 or 2 kpc further out before falling back to the centre in an excentric orbit. The big puzzle then is how the gas can be kept in an invisible (ionized?) state during its return journey. This seems next to impossible so that perhaps a one-way outward motion maintained by enormous non-gravitational forces would seem the more plausible solution.

5. IONIZED HYDROGEN REGIONS AND THE THERMAL COMPONENT OF CONTINUOUS RADIO EMISSION

It was realized in 1940 that *thermal* emission from the ionized regions of interstellar gas (radiation by non-relativistic electrons passing through the Coulomb field of an ion) should be of observable magnitude. The conviction soon arose that some other emission mechanisms must be predominant, at least in the metre range of wave-lenghts. This component was called *non-thermal* and its conventional interpretation is now by the synchrotron mechanism (radiation by relativistic electrons describing helical orbits in a magnetic field). The correctness of this interpretation has not definitely been proven (see sections 6 and 7).

Criteria for separating the thermal and non-thermal components have been sought in their spectrum, in their distribution on the sky, and in their polarization. The polarization method has not yet been successful. At one time the tendency was to go by the distribution on the sky : if the distribution of thermal radiation is made to agree with our ideas on the thickness of the gas layer (about 200 pc) then the rest must be non-thermal. Or, if by a preset idea the non-thermal radiation is called an « isotropic » component without a strong increase towards the galactic plane, then the rest must be thermal. Neither assumption is justified, so at present it seems best to adopt the spectrum as the main criterion. Such a study ([10]), based on the comparison of surveys at different frequencies, is made difficult by the different beam widths and the possibility of scale errors of 1 or 2 db in the intensity. We quote in the following a number of results made available to us before publication by G. Westerhout.

A. *Individual sources.* — The brightness B and brightness temperature T_b of an area on the sky, or on an extended source, are related by :

$$B = 2kT_b\nu^2 c^{-2}$$

The exponent α in :

$$T \sim \nu^{-\alpha},\ B \sim \nu^{-(\alpha-2)}$$

is called the spectral index. A black body, e.g. a thermal source of large optical depth has $\alpha = 0$. A thermal source of small optical depth has very nearly $\alpha = 2.0$. The non-thermal sources $\alpha = 2.5$ to 3.0. Virtually all sources observed in the metre range are non-thermal. The great majority of sources observed at 30 cm to 3 cm are thermal and have positions within a few degrees from the galactic plane, characteristic of the disk population.

Thermal gas of small optical depth gives the brightness temperature :

$$T_b' = 0.37\, E\, (\nu/100 \text{ Mc})^{-2}$$

where ν is the frequency and E is the integral $\int n_e^2 dl$ expressed in $\text{cm}^{-6} \times \text{kpc}$. The radio estimates of E for the identified thermal sources are in fair agreement with the estimates from optical data; they are of the order of 10^4 for the well-observed ones.

As T_b'/T is the optical depth, τ, we find for $T = 10^{+4}$, $\nu = 15$ Mc, an optical depth of the order of 15. This means that the same nebulae should stand out as dark objects against the non-thermal background as soon as the latter's brightness temperature exceeds 10^4 °K. The contrast will be strongest for the nearer objects that have little non-thermal radiation in front. Shain has indeed collected a good deal of information about these regions from observations with the big 1,500 $\times$ 1,500 meter cross at 15 metre. No detailed data are available. The cross observations by Mills at 3.5 metre do not admit such an easy interpretation as the optical depths are smaller and the gas temperature of the ionized regions of the same order as the brightness temperature of the non-thermal background.

B. *The general background* (as distinct from recognizable individual sources). — Westerhout assumes that the radiation at 22 cm at latitude $b = \pm\, 5°$ is entirely non-thermal. A measurable thermal intensity would have meant a conspicuous absorption at the very long wave lengths (15 metre), which is clearly absent at these latitudes. The non-thermal spectral index $a = 2.7$ observed at these latitudes is adopted also in the galactic plane at the same longitudes (0 — 50° from the centre). A well-mixed medium with thermal emission and absorption and non-thermal emission gives :

$$T_b = \left(T_e + \frac{T_{nt}}{\tau}\right)(1 - e^{-\tau}).$$

This formula is applied to find the two unknowns τ (22 cm) and T_{nt} (22 cm) from the two measured values T_b (22 cm) and T_b (350 cm) in any given direction. It is assumed that $T_e = 10^4$. τ (350 cm) = 270. τ (22 cm), based on $\alpha = 2.0$, and T_{nt} (350 cm) = 1,900. $.T_{nt}$ (22 cm) based on $\alpha = 2.7$.

The results are that the thermal radiation is indeed limited to a narrow zone of latitudes; the intensity drops to half at $\pm$ 0.8°. The non-thermal intensity also drops sharply in the first degrees of latitude but levels off at higher b. On the face of it, a division may be made between a non-thermal component extending to very high latitudes and one extending to $\pm$ 4° and reaching half-intensity at $b = \pm$ 2°. This does not necessarily imply a physical distinction or different origin.

The smoothed data in Table 2 summarize the situation for 22 cm, where $\tau \ll 1$ in all direction. Here l' = longitude to centre, b' = latitude to true plane. The uncertainty in absolute and relative intensities is about 1 °K. An intensity of 1 or 1.5 °K spreads over the entire sky.

TABLE 2

Continuum at 22 cm.

l'	b'	Thermal $10^4\tau$	Non-thermal T_{nt}	Total T_b
0	0	170*	30*	200*
10	0	10	11	21
20	0	13	9	22
30	0	10	7	17
40	0	7	6	13
10	1	3	9	12
10	2	0	6	6
10	4	0	3	3
10	10	0	2	2
10	> 20	0	1	1

* Top line uncertain; source Sgr A, ref. [11].

When, finally, the thermal component is analyzed on the assumption of cylindrical symmetry, it is found that the region $R < 3$ kpc is almost empty of ionized hydrogen gas. The thermal emission per unit volume is strong in the ring near $R = 4$ kpc and gradually drops to 0.20 times the maximum value at 8 kpc.

The mass does not follow directly as we find $\overline{n^2}$ and not $\overline{n}$, where n = number of electrons and protons per cm^3. However, if we assume clouds all of number density N and empty interspace then $\overline{n^2} = \overline{n}.N$. The peak value of $\overline{n^2}$ is about 1. With $N = 5\ cm^{-3}$ this gives the following mass ratios :

	stars	neutral H	ionized H
whole Galaxy . .	50	1	0.06
near $R = 4$ . . .	—	1	0.3
near Sun	—	1	0.06

These estimates seem to be somewhat more firmly founded than the conventional estimate that about 10 per cent of the interstellar hydrogen gas would be ionized.

6. THE GALACTIC HALO; ORIGIN OF NON-THERMAL RADIATION

No other suggested interpretation of the non-thermal radiation from the galactic halo has to be taken seriously but the synchrotron mechanism. This means ([12]) that there have to be about 10^{-12} electrons per cm^3 moving with energies of the order of 2×10^9 eV in fields of the order of 2×10^{-6} gauss. Unlike the halo distributions of stellar objects, the emission per unit volume in the radio halo does not strongly increase towards the galactic centre. From an analysis of observations at 3.7 metre, it may be spherical or flattened in the ratio 0.5 : 1 at most. The radius is of the order of 15 kpc ([13]).

For comparison we may mention the best data now available for the Andromeda nebula, at $\lambda = 75$ cm ([14]). The flattening is about 0.4 : 1 (but 0.6 : 1 as projected on the sky) and the equatorial radius is 25 kpc. The volume emission of M 31 and the Galactic System is of the same order of magnitude.

The electrons responsible for this radiation should not be identified with the gaseous halo or corona. It is very probable that electrons and protons of much lower energies and in much higher numbers form the actual gas, which is the seat of the electric currents connected with the magnetic fields.

The main theoretical argument for the existence of such a gaseous corona is, that the gas in the ordinary clouds in and near the disk would expand too rapidly unless there is approximate pressure equilibrium with the outside gas. As of summer 1957 two rather different ideas existed. Spitzer ([15]) maintained a gas of $T = 10^6$ °K, $N = 5 \times 10^{-4}$ cm^{-3}, $z_{max} \gtrsim 8$ kpc, all fully ionized. A higher temperature would allow the gas to escape into intergalactic space, a lower temperature would make the corona too flat. Pikelner ([16]), ([17]), ([18]), however, defended a gas with $T = 10^4$ °K, $N = 10^{-2}$ cm^{-3} with possibly 20 — 80 per cent neutral hydrogen. It would be held up by turbulent motions; magnetic fields were envisaged to keep these motions from degrading into thermal motion.

Skipping the theory, we shall look if other observations can bring the decision. The following arguments have been advanced.

(*a*) Wide absorption lines of Ca^+ in certain stars; these seem irrelevant as Spitzer claims that they are stellar, not inter-stellar.

(*b*) Multiple narrow absorption lines of Ca^+ in high-latitude stars. Apparently, clouds still exist at $z = 1$ kpc and higher. Spitzer shows that these would have vanished by expansion in the course of their path up from the galactic disk, unless held together by outside gas pressure. Either Pikelner or Spitzer's corona might do this.

(*c*) Published high-latitude observations of the 21-cm line ([19]) show radiation at $b = 20^\circ - 40^\circ$ and v up to 50 km/sec, probably coming from details in the galactic disk. Lines with widths of the order of 30 km/sec or smaller are seen in all high-latitude fields observed in Dwingeloo. Faint wings with $v = 100 - 200$ km/sec have first been suspected ([19a]). Present estimates are that such wings are < 1 °K if existing at all. The north-pole line profile at Dwingeloo now has a total halfwidth of 8 km/sec, and wings extending from $- 60$ to $+ 40$ km/sec. Beyond that the radiation is below ½ °K. The comparison of two points mentioned in sec. 3 suggests that they are either fainter, or very uniform in the sky to within 0.1 °K. A corona by Pikelner's specifications would give about 1 °K and is not favored by these observations.

(*d*) The Dwingeloo observations put a similar upper limit, say 0.3 °K, on the brightness temperature of the 21-cm halo of the Andromeda nebula.

It may be concluded that the evidence is most favorable for Spitzer's corona (which is simply unobservable!). A corona of this type might quite well house the fields and relativistic electrons that are needed to explain the non-thermal radio emission. Not explained is, why this non-thermal emission should still show a marked concentration towards the galactic plane as set forth in the preceding section. There has to be a rather conspicuous change in some physical parameter between $z = 0.3$ kpc and 1.5 kpc (i.e., $b = 2°$ and 10° at the distance of the galactic centre). This range of z is well above the limits of the neutral and ionized hydrogen disk but still small compared with the galactic halo at large.

7. MAGNETIC FIELDS IN THE GALACTIC DISK AND HALO

No astrophysicist seems to doubt the existence of magnetic fields of the order of $(0.5 - 1) \times 10^{-5}$ oersted in the galactic disk and of slightly weaker fields in the halo. Yet the evidence is by no means very direct; Table 3 lists, in a rough order of directness of proof, the phenomena that might be considered as evidence for magnetic fields in astrophysical bodies. So far, for the interstellar medium the more direct proofs are hopes rather than facts.

Historically, the cosmic rays (F) have given the first reason for seriously considering galactic magnetic fields. A comprehensive review was given by Biermann ([20]), ([21]). Quantitative estimates give $H = 10^{-6}$ in order to prevent the fastest particles (10^{18} eV) from individually escaping and $H = 10^{-5}$ to keep in the pressure of the cosmic ray gas as a whole. Later work does not seem to have added fundamentally new points relevant to the direct interpretation of observational evidence, except for the existence of modulating effects in and near the Solar System.

What would happen if the magnetic pressure were insufficient? Probably the conductors that carry the currents connected with the magnetic field would be subject to an expanding force because of the tendency of the cosmic-ray gas to expand. All of these arguments

TABLE 3

Evidence for magnetic fields.

No.	Phenomenon	Familiar example	Interstellar gas
A	Zeeman effect	Sunspots	Not yet observed; perhaps feasible for 21-cm line.
B	Polarized radiation by synchrotron mechanism	Crab nebula	Strong non-thermal radiation observed but polarization not yet detected.
C	Forms of gaseous masses suggestive of lines of force	Polar rays of corona	No convincing examples.
D	Birefringence	Ionosphere, radio emission from Sun	Faraday effect = circular birefringence not yet detected.
	and		
E	Dichroism		Interstellar polarization = linear dichroism discovered 1949; many detailed data.
F	Other phenomena for which magnetic fields afford the simplest explanation	—	Acceleration of cosmic rays and maintenance of their intensity inside galactic system.

refer to the galactic halo rather than the disk and there is not much evidence against such an expansion. Another wild idea is that, if for some reason the main source of cosmic rays is situated in the galactic nucleus, the expanding gas motion near the nucleus might be caused by the pressure of these cosmic rays.

Interstellar polarization (E), i.e. the difference in extinction for light of the same wave length in two linear polarizations, gives like any optical method information mainly about the region of the galactic disk within a few kpc from the Sun. Several thousand stars have been measured. A first striking fact is that the polarization values are similar over fairly large areas. This means that a turbulence theory leading to magnetic eddies smaller than 100 pc does not fit the observations. Further, the electric vectors of the measured light tend to lie parallel to the Galactic equator. On the Davis-Greenstein theory ([22]) these vectors should coincide with the magnetic lines of force projected on the sky. Consequently, the magnetic lines of force tend to have a preference for directions parallel to the galactic plane. This preference would find a natural explanation in the stretching by differential rotation, no matter how the initial field looks.

Two refinements of this picture have been suggested. The first one is Chandrasekhar and Fermi's suggestion ([23]) that the spiral arms are tubes of force. This suggestion has some striking points of support but generally the present observational data on the arms and on the polarization do not confirm this hypothesis. The second one is that in certain local regions certain deviating directions occur, which can also be recognized by a careful inspection of the forms of emission nebulae or dark clouds. Shajn and others have extensively explored this point of view. Behr by measuring polarizations with an accuracy of 0.03 per cent, even claims to find a polarizing cloud at less than 50 pc from the Sun. These investigations are certainly interesting but not very relevant to the present discussion on large-scale fields. For full references see ([24]).

Although in earlier work the non-thermal radiation was sometimes identified with the isotropic background, we now have evidence (Table 2) that it shows a fair concentration towards the disk. In the disk, there are signs of somewhat enhanced intensity in the spiral arms ([2]). It may be asked which quantity changes so rapidly in space : the magnetic field or the number of relativistic electrons, or their energy spectrum. The only definite suggestion ([17]), ([29]) about this is that the energy spectrum does not change. The emission per unit volume then is proportional to $NH^{+1\cdot8}$, where N is the number of relativistic electrons per unit volume, H the magnetic field, and where the exponent follows directly from the spectral index of the non-thermal radio emission. Moreover, these authors assume $N \sim H$, so that the emission per unit volume becomes proportional to $H^{2\cdot8}$. On this interpretation, the observations show that the field in the halo is about half as strong as the field in the disk.

Measuring the polarization (B) would give a direct support of the hypothesis of synchrotron radiation. It would at the same time provide extremely valuable information on the orientation of the magnetic fields at large, provided the difficulty of the Faraday effect (D) can be overcome. This is possible by employing neighboring frequencies. With careful technique this experiment can be hoped to give positive results ([25]) but the attempts reported so far have been unsuccessful ([26]), ([27]).

Finally, there is a distinct hope that the Zeeman effect of the 21-cm line may be measured ([28]). The problem is to measure line shifts of 30 c/s (assuming $H = 10^{-5}$) when shifting between the two circular polarizations. This will take a lot of very careful technique but it should be possible with a big telescope in the directions in which very sharp lines are seen, notably on the absorption lines in Cas A and Sgr A.

Besides the observational evidence reviewed above, an impressive number of investigations have been devoted to some aspect of the theory of magnetic fields and interstellar gas dynamics (see, for instance, ([24])]. It seems best not to include a superficial review of these investigations in this report.

REFERENCES

(1) W.W. Morgan, S. Sharpless and D. Osterbrock, *A. J.*, **57**, 3 (1952); (cf. also *Sky and Telescope*, **11**, 134, 1952).

(2) J.H. Oort, F.J. Kerr and G. Westerhout, *M. N.*, in press (1958).

(3) C.A. Muller and G. Westerhout, *B.A.N.*, **13**, 151-195 (1957);
A. Ollongren and H.C. van de Hulst, *ibid.*, 196-200;
G. Westerhout, *ibid.*, 201-246;
M. Schmidt, *ibid.*, 247-268.
E. Raimond, *ibid.*, 269-273.

(4) F.J. Kerr and G. de Vaucouleurs, *Aust. J. Phys.*, **9**, 90 (1956).

(5) H.C. van de Hulst, E. Raimond and H. van Woerden, *B.A.N.*, **14**, 1-16 (1957);
M. Schmidt, *ibid.*, 17-19.

(6) D.S. Heeschen, *Ap. J.*, **126**, 471 (1957).

(7) D.S. Heeschen, *Ap. J.*, **124**, 660 (1956).

(8) A.E. Lilley and E.F. McClain, *Ap. J.*, **123**, 172 (1956).

(9) H. van Woerden, G.W. Rougoor and J.H. Oort, *C.R. Acad. Paris*, **244**, 1691-1695 (1957).

(10) R. Hanbury Brown and C. Hazard, *Phil. Mag.*, **44**, 939 (1953).

(11) G. Westerhout, *C.R. Acad. Paris*, **245**, 35 (1957).

(12) G.R. Burbidge, *Ap. J.*, **123**, 178 (1956).

(13) J. Baldwin, *M. N.*, **115**, 684 and 691 (1956).

(14) C.L. Seeger, G. Westerhout and R.G. Conway, *Ap. J.*, **126**, 585 (1957).

(15) L. Spitzer, *Ap. J.*, **124**, 20 (1956).

(16) S.B. Pikelner, *Izv. Krimsk. Obs.*, **10**, 74 (1953).

(17) S.B. Pikelner and I.S. Shklovsky, *Ast. Zh.*, **34**, 145 (1957).

(18) S.B. Pikelner, *Ast. Zh.*, **34**, 314 (1957).

(19) H.L. Helfer and H.E. Tatel, *Ap. J.*, **121**, 585 (1955).

(19a) H.C. van de Hulst, *Rev. Mod. Physics*, July 1958.

(20) L. Biermann, *Kosmische Strahlung* (W. Heisenberg, editor), 1953.

(21) L. Biermann, *Ann. Rev. Nucl. Sci.*, **2**, 335 (1953).

(22) L. Davis and J.L. Greenstein, *Ap. J.*, **114**, 206 (1951).

(23) S. Chandrasekhar and E. Fermi, *Ap. J.*, **118**, 113 (1953).

(24) H.C. van de Hulst, *Draft Report Commission* 34 (Interstellar Matter) *for I.A.U. Assembly Moscow* 1958.

(25) G.G. Getmanzev and V.A. Razin, *5th Conference on Cosmogony*, 496 (1956).

(26) V.A. Razin, *Radiotechnics and Electronics* (Moscow), **1**, 846 (1956).

(27) J.M. Thomson, *Nature*, **180**, 495 (1957).

(28) J.G. Bolton and J.P. Wild, *Ap. J.*, **125**, 296 (1957).

(29) V.L. Ginzburg, « Progress in Elementury Particle und Cosmic Ray Physics », **4**, 339, Amsterdam (North-Holland Publ. Cy) 1958.

Discussion of van de Hulst's Report

Dr Hoyle. — On the picture I presented, the galaxy should still continue to act as cooling centre, and I think that the piston effect of the external hot gas should be discernable. This amounts to about 250 km/sec, which is just of the order of the motions of the gas observed in the nucleus.

The general idea would be of the temperature rising as one goes outside the galactic disk, perhaps to Spitzer's value of 10^6 °K in the halo, and then to still higher values outside the halo.

Dr Gold. — A relative motion between gas and stars is suggested by the appearance of the galaxies where the dust is frequently seen to the inside of the bright stars in a spiral arm.

Andrew Young at Harvard has studied this problem and concludes that orbital motion of the gas somewhat faster than the gravitational orbital one is implied.

The new stars formed from the gas would then first drift in the outward direction and backwards relative to the gas. They would thus appear on the outside of a circumferential spiral arm, but to the inside — that is behind — a radial spiral arm.

Just this is observed, with a very clear changeover point, in the case of some spirals where the dust lane leads the stars in the spoke, but changes to the following side — that is the inside — for the circular arm.

Dr Lovell. — In connection with the evidence for magnetic fields, I think that last August, at the URSI meeting in the U.S.A., convincing evidence was produced of the polarisation of radio emission in the centimeter band from the Crab Nebula.

Dr van de Hulst. — This was the observation of polarisation in the radiospectrum of the Crab Nebula by Mayer, from the Naval Research Laboratory at Washington. I have not included it in Table 3, because I wished to restrict this table to the interstellar gas, as distinct from nebulae.

Dr Hoyle. — About the cosmic rays, the upper limit of the energy spectrum is now known to be so high that it is becoming somewhat difficult to suppose that the most energetic particles are contained within the galaxy, even if the magnetic field is chaotic in form. And of course, since the dE/E^3 law is maintained to the top of the spectrum, leakage at the top carries the implication of leakage over the whole spectrum.

Dr van de Hulst. — How high is the highest energy observed at present?

Dr Hoyle. — About 10^{19} eV.

Dr Oort. — But energies of this order must occur only very rarely.

Dr Oppenheimer. — The fact that events with $E = 10^{19}$ eV have been found with present equipment shows that they are not very rare compared to what was expected from the power law.

Dr Hoyle. — At an upper limit of 10^{19} eV the radius of curvature of a proton orbit is of order 1 kpc, about the thickness of the galaxy — this is for a field of 10^{-5} Gauss.

Dr Bondi. — How strong are the arguments against a component velocity of matter in the nuclear region normal to the plane of the galaxy?

Dr van de Hulst. — The velocity component itself cannot be measured, of course. But the distribution of the emission in the vertical direction is very narrow : a change of $1/2^{\circ}$ in latitude makes already a marked difference. This makes it practically certain that the motions are in the plane.

Dr Gorter. — Have you supposed the particles which scatter light to place themselves perpendicular or parallel to the magnetic field in the spiral arms, and would the experimental data clearly favour one of these possibilities?

Dr van de Hulst. — The interpretation of the interstellar polarisation data in terms of magnetic fields in my report, was based on the theory of Davis and Greenstein. The fact that magnetic fields of this type seem plausible gives some support to Davis and Greenstein's theory.

On the Evolution of Galaxies

by V. A. AMBARTSUMIAN

INTRODUCTION

Attempts to approach the solution of the question on the origin of galaxies were, until the present, based mainly on speculations connected with the remarkable fact of mutual recession of extra-galactic nebulae. In other words, these attempts have been made within the frames of the existing cosmological theories, which, as a rule, are based only on some integral and averaged characteristics of the world of galaxies surrounding us.

Although the study of the nearest galaxies and also the investigation of groups and clusters of galaxies have not been advanced far enough, nevertheless rich material has been gathered, which may be possible to rely on while considering the problems of the origin and evolution of galaxies. Among the data obtained from observations those concerning multiple galaxies, groups of galaxies and clusters of galaxies deserve special attention.

In this connection it is worth while to dwell briefly on the significance, which the study of multiple stars and stellar clusters have had for the problem of the origin and evolution of stars.

1. In the thirties of the present century the very existence of stellar clusters in our Galaxy, together with some statistical-mechanical considerations on the non-reversible character of the process of dissolution of clusters led to the conclusion that stars forming a cluster originate together. In other words, it was established that at least some stars in the Galaxy have been formed in groups (1).

2. Statistical data concerning double stars have led to the con-

clusion that the components of each double star have a common origin (2).

3. The mere existence of stellar associations has made it possible to come to a conclusion on the continuing process of star formation in the Galaxy (3). The discovery of the expansion of stellar associations has permitted to conclude that at least the majority of the stars of the disc population has also been formed within the stellar groups, now already disintegrated (4).

4. The study of the spectrum-luminosity diagram for stellar clusters has permitted to construct interesting schemes of the evolution of different stars. These schemes need further verification, but in any case their significance for the theory of stellar evolution is extremely great.

5. Marking out the Trapezium type multiple systems rendered it possible to establish the existence of particularly young multiple stars and thus to approach the very moment of the formation of stellar groups.

In this respect it seems to us that the situation in the world of galaxies is even more favourable. Multiple galaxies and groups of galaxies provide interesting material for the reasoning out on the group formation of galaxies. Moreover, the tendency of grouping in the world of galaxies is so strong that any study of galaxies is inevitably connected with the question of the nature of one group or another.

For instance, such nearby giant stellar systems as M 31, M 81, and M 101 are the centres of the highly interesting groups of galaxies. Our Galaxy itself has several companions of different nature.

Therefore, it is natural to think that the question of the origin of galaxies cannot be separated from the question of the origin of their groups and clusters.

Let us now turn to the fact that in multiple galaxies the periods of revolution reach billion years and more, while in clusters, the time necessary for one revolution around the centre of the cluster is to be measured by several billions of years. Whereas, the age of galaxies themselves is thought to reach also several billions of years. In such a case the multiple galaxies as well as clusters of galaxies in their present state must, even in configurations of their components, maintain traces of primary conditions of group formation. And this seems to mean, a possibility to approach at least the kinematics of those phenomena which have led to the formation of groups.

In the present report we begin the consideration of the question with the problem of multiple galaxies and cluster galaxies. The study of some multiple systems, however, has led us to the conclusion that an intimate relation exists between the mechanism of the formation of components and the ways of formation of characteristic details in the structure of galaxies. For the time being, it is difficult to understand, the exact character of this relation, but it seems to us that further study of this aspect of the question will open some prospects in solving the question on the origin of the observed structures of separate galaxies. Finally the investigation of radio-galaxies representing systems in which violent non-stable processes take place, has shown that in every such system we meet with traces of some sort of duplicity. Comparing this with other data concerning multiple galaxies we see that the narrower the double or the multiple system is the sharper the traces of non-stability become. All this emphasizes again the significance of the data concerning multiple galaxies and the tendency of galaxies to form groups for the problem of the origin and evolution of galaxies.

Confining this report to the above mentioned questions I hope, that actual data referring to the stellar population of galaxies and representing interest to the problem of evolution of galaxies, will be elucidated in other reports presented at this conference.

1. THE PHYSICAL MEANING OF CLUSTERING TENDENCY

After the works of Zwicky (5) and also of Shane and Scott (6) there is a sound basis to consider that the majority of galaxies are members of clusters or of groups of galaxies, while the number of isolated galaxies in the general metagalactic field is comparatively small. In this sense, it is even difficult to speak about any general homogeneous metagalactic field, as distinguished from the condensations of galaxies. In reality we have a metagalactic field mainly composed of different clusters and groupes, i.e. of inhomogeneities of different scales.

In this respect the situation in Metagalaxy strongly differs from that of the stellar systems where the general stellar field with a slowly changing density usually dominates, while clusters are comparatively rare and separate inhomogeneities occur in this field.

It may be concluded from statistical mechanics that clusters and groups must eventually dissolve ([1]). At the same time the dissolution will proceed differently and require different lengths of time, depending on whether the clusters and groups in consideration are in a steady (or in a quasi-steady) state with negative total energy, or in a state when among the cluster-members, there is a certain percentage which possesses positive energy and can immediately leave the clusters with great velocity.

In the second case, the dissolution must take place in time order of interval, which is necessary for the galaxy, entering the cluster, to intersect it from one end to the other, i.e. the time order of a hundred million or 1—2 billions of years.

In the first case, when the clusters possess negative energy, the dissolution must take place owing to the fact that some galaxies during mutual encounters receive positive energies and leave the cluster. In other words, in this case, a mechanism is, in effect, similar to that which takes place in steady-state stellar clusters. This mechanism, however, needs time of the order of a hundred billion and more years. Taking into account that the ages of galaxies are measured only by a few billions of years, the significance of this mechanism in most cases is not great.

Thus, we may say that either the clusters are to dissolve because of their non-steadiness if they possess in their components a significant number of members with positive energy, or they are in steady-state and their dissolution must proceed so slowly that it cannot have any essential significance.

The question is, which of the considered two versions takes place regarding any given cluster. In each concrete case this question must be solved on the basis of analysis of radial velocities and their comparison with the mass of clusters, which must be determined, if possible, in an independent way. Later we shall give some concrete examples. Isolated galaxies being few in number we may conclude that the majority of clusters are neither clusters of positive energy nor have a large number of members with velocities surpassing the escape velocity.

We have mentioned above that owing to mutual encounters of members the disintegration of a steady-state cluster may take place on account of the escape of galaxies receiving a large increase in kinetic energy. We may, of course, imagine a reverse process when an outer galaxy enters a cluster with great velocity and, giving up its energy, remains in the cluster. However, it is not difficult to show that in the present state of Metagalaxy such processes must take place with a frequency of much lower order of magnitude than the direct processes of escape of galaxies out of clusters. Meanwhile we have seen that even these direct processes take place so seldom, that they cannot have an essential significance for systems with negative energy. It follows that the process of capture may be absolutely neglected.

Conclusion.

In the present conditions of Metagalaxy the clusters and groups may either persist or desintegrate. But they cannot be enriched at the expense of galaxies which have originated independently from them.

2. DEVIATION FROM DISSOCIATIVE EQUILIBRIUM

Attention should be paid to the fact that among the members of clusters of galaxies known to us, we meet with double and multiple galaxies. Double and multiple galaxies are met with more often in loose clusters of the Virgo type. Apparently they are seldom met with in compact clusters of the Coma type. In such relatively poor groups as the Local group of galaxies double and multiple systems are comparatively frequent. Nevertheless, if the existence of subdwarf galaxies of the type of the objects in Sculptor and Fornax is considered, then, apparently, every one of the multiple galaxies transforms into a group consisting of approximately a dozen members. For example, our Galaxy with Magellanic clouds forms a triple system. But it is also surrounded with several subdwarf systems of the Sculptor type. The galaxy in Andromeda is a multiple system composed of five members. But probably there are also some systems of the Sculptor type near it. That is why it seems that we should rather speak of the groups that include our Galaxy and M 31 respectively. Let us remember, however, that when speaking about the multiplicity of stars we do not take into account the possible presence of planets, since the latter possess masses, insignificant compared with stars. When determining the multiplicity of galaxies it is advisable not to take into account the Sculptor type systems, just as the globular clusters, which have, apparently, masses a little less than the masses of galaxies of the Sculptor type are not taken into account.

In such a case we may admit that in our Local group, containing only a few individual galaxies (M 33, NGC 6822, IC 1613 and perhaps some others), there is one triple galaxy and one other galaxy with a higher multiplicity. We may put the question as to what the mathematical expectation of numbers of double and multiple galaxies amounts to in our Local group in case of dissociative equilibrium. It turns out that in case of dissociative equilibrium the mathematical expectation of the number of double galaxies in the Local group should be less than 0.05, and the mathematical expectation of the number of triple galaxies and galaxies of higher multiplicity even many times less.

Therefore, the fact that we have two systems of very high multiplicity in the Local group of galaxies is a very strong deviation from dissociative equilibrium. The situation is similar in many other groups and clusters. In some instances the degree of deviation from dissociative equilibrium is many times greater.

If double galaxies and multiple galaxies originate by way of mutual capture (in the course of triple encounters) or otherwise, from previously independent single galaxies, then at the primary stage of the evolution of clusters there should certainly have been a deviation in them from dissociative equilibrium. However, these deviations should have been in the opposite direction, i.e. the number of multiple galaxies should have been less than in case of dissociative equilibrium. Only after a long time the average number of multiple galaxies in clusters could have reached the theoretical value in accordance to dissociative equilibrium. The percentage of multiple galaxies with an accuracy determined by statistical fluctuations would, in this case, never surpass the indicated equilibrium value.

The fact, that the percentage of multiple systems is indeed much higher than this theoretical limit, proves our assumption that multiple galaxies are formed from individuals to be incorrect.

Conclusion.

Components of any multiple galaxies are formed together.

This conclusion is based on statistical considerations, and that is why it is valid for the overwhelming majority of multiple galaxies. Therefore, some exceptions are possible, and some insignificant minority of multiple galaxies could have been formed on account of mutual capture (by triple encounters or otherwise) from individual galaxies.

3. THE OBSERVED CONFIGURATIONS OF MULTIPLE GALAXIES

During the lifetime of galaxies (a few billions of years) perturbations in the state of multiple galaxies rising owing to nearby passages of outer galaxies should be insignificant. Therefore, it is possible to presume that these states still bear traces of primary conditions of formation of multiple systems. It is reasonable, therefore, to search for information about the mechanism of formation in statistical data, characterizing the bulk of double and multiple galaxies.

Unfortunately we have insufficient *quantitative data* of such kind. For example, it would be interesting to know the law of distribution of distances between the components of double galaxies. As regards the value of this distance for individual pairs, they themselves can hardly be bases for cosmogonic conclusions.

The matter is altogether different with multiple galaxies whose number of components is more than two.

Each such galaxy is characterized with some *space configuration* of its components. Considering even a small number of such configurations we may come to some conclusion on *the dominating type of configurations among multiple galaxies.*

It is true, we do not directly observe space configurations, but only their projections in the sky. However, the study of these projections permits us to make conclusions on the character of space configurations.

When studying the problem, considering comparatively young multiple stars, we divided all possible configurations into two main types : *configurations of the Trapezium type multiple systems and configurations of the usual type* (7). Let us remember the definition of those and others.

Under the *Trapezium* type multiple system we mean a multiple system where it is possible to find three such components *a*, *b*, *c*, so that all three distances *ab*, *bc* and *ac* are of the same order of magnitude. If it is impossible to find three such components in the multiple system, then it is called a system of *usual* type.

This definition needs some supplementation. For its application we must agree as to what we exactly mean when we say that the three distances are of the same order of magnitude.

It is convenient to consider the distances *ab*, *bc* and *ac*, as being of the same order when all three ratios $\frac{ab}{ac}$, $\frac{ab}{bc}$ and $\frac{ac}{bc}$ are confined within the limits between K_0 and $\frac{1}{K_0}$, where K_0 — is any number of the order of $\sqrt{10}$. If we want to make a stricter choice then the value of K_0 can be taken a little less than $\sqrt{10}$. For example, in some papers we took $K_0 = 2.5$. Such systems, wherein it is impossible to find three components for which the ratio of the greatest mutual distance to the smallest is less than 2.5, but where it is possible to find three components for which this ratio lies between 2.5 and 3, have been called by us systems of *intermediate* type.

Such a division of configurations of multiple systems into two basic types with the addition of an intermediate type, introduced only for stricter delimitation of the basic types, is also useful for purposes of extragalactic astronomy.

As it is known, among multiple stars the usual type systems are dominating. Only among multiples containing O type stars, we observe a large percentage of Trapezium type systems. The same is true but to a lesser degree for multiples containing Bo stars. As it is known, this peculiarity of O and Bo type stars is connected with their relatively young age. Since, however, the O and Bo stars constitute an insignificant percentage of all multiple stars, this does not change the fact that multiple stars, as a rule, have configurations of the usual type.

In the case of multiple galaxies the situation is quite different. If we take the multiple systems which figure in the published lists of double and multiple galaxies, we see that the percentage of configurations of the Trapezium type among them significantly surpasses that of the usual type systems.

For example, among the 132 multiple galaxies of Holmberg's ([8]) catalogue 87 have such configurations that must certainly belong to the Trapezium type. Only 27 systems are of the usual type while the rest 18 have configurations of the intermediate type ([9]).

A sharp contrast between the configurations of multiple galaxies and the configurations of multiple stars may also be illustrated by the following examples.

If we choose from the catalogue of the visual double stars of the whole sky the six multiples, whose main components possess the greatest visual brightness among all the main components of the multiple stars of the catalogue, we find that all of these six multiple stars possess configurations of the usual type.

If we now take six multiple galaxies from Holmberg's catalogue with the greatest apparent brightness of the main components, all of them are found to be of Trapezium type systems.

Further let us take the brightest star of high multiplicity. For example, among the known sextet-stars the Castor possesses the greatest visual brightness. Speaking about the multiplicity of this star we take into account that each of its three visual components is a spectroscopic binary. This is a system having a typical usual configuration. On the other hand, the object most prominent in apparent brightness among the sextet-galaxies is the multiple system NGC 6027, studied by Seyfert ([10]). It is a typical trapezium. Moreover in this system it is possible to choose in many different ways three galaxies, between which all the distances are of the same order of magnitude.

Here we shall not consider the question on the selectivity of the catalogue of multiple galaxies relative to the configuration of different types. We shall neither consider technical questions about the corrections which are necessary to make in the statistical data for the transition from type distribution, received in *projections*, to type distribution which should have been received if we had the possibility to conduct the statistics of *space* configurations. These questions are considered in the first approximation in an earlier paper of the author.

The obtained quantitative corrections do not change the qualitative results. Therefore, we have the following conclusion : *Most of the multiple galaxies possess configurations of the Trapezium type.*

4. THE CAUSE OF PREDOMINANCE OF TRAPEZIUM

The fact that the overwhelming majority of multiple stars have configurations of the usual type has been explained as follows. The configurations of the Trapezium type are, as a rule, unstable, even in case the total energy of multiple system is negative. When the motions in the system of the usual type can be approximated by a sum of a few keplerian (i.e. periodical) motions, the movements in the system of the Trapezium type are very complex and involved. During nearby encounters of two components, occurring in the course of time, one of them may acquire enough kinetic energy to leave the system. This same mechanism operates in open stellar clusters. Calculations show that for the dissolution of a system possessing a configuration of the Trapezium type, it is necessary, on the average that its components make a few revolutions. For most of the stars such interval of time is insignificant in comparison with their age. Therefore, the overwhelming majority of the systems of the Trapezium type, originated in Galaxy, have been dissolved. This explanation gives us, at the same time, the possibility to understand the exception, observed in the case of O and Bo stars. Many of these stars possess the age of the order of 10^6 years and considerably less than 10^7 years. Meanwhile the period of revolution, in observed multiple stars of the Trapezium type, must be of the order of 10^5 — 10^6 years. Therefore, the components of these trapeziums may have time to perform only few revolutions around their centres of gravity. That is why these multiple systems have failed to be dissolved.

For *multiple galaxies* the situation is the same as for O — B stars. The age of multiple galaxies is measured by a few billions of years, meanwhile the periods of revolutions in them reach the order of billion years. Consequently, the components of multiple galaxies may have time to make only a small number of revolutions. That is the reason why multiple galaxies having configurations of the Trapezium type have not been dissolved.

Since the mechanism of the disintegration of observed multiple galaxies acting in a selected manner only on the systems of the Trapezium type should not have had time to influence most of the systems, it seems that the present distribution of configurations into types reflects that primary distribution which depends upon the laws of the formation of multiple galaxies.

Conclusion.

The high percentage of configurations of the Trapezium type among multiple galaxies is in full agreement with the ratio of the age of galaxies to the periods of revolution in multiple systems.

5. MULTIPLE SYSTEMS WITH POSITIVE ENERGY

In the previous paragraph we spoke about the « revolutions » of components in a multiple galaxy. This implies that in our systems the components, at least in their primary stage of evolution, hold one another by gravitational forces. In other words, till the present, we presumed that all multiple galaxies are systems with negative total energy.

However, in order to determine the sign of energy of a given multiple system, beside the data on the configuration of components, reliable data are necessary on the masses and velocities of components. Unfortunately, the knowledge we have on the masses of double and multiple galaxies is obtained under the assumption that these systems possess negative energies, i.e. under the assumption which is necessary to examine.

To put a question on the possibility of existence of multiple systems with positive energy may seem superfluous, because, in the case of stars, all double and multiple systems, well studied until now, prove to have only negative energy. Let us, however, suppose a moment that some multiple stars are formed in Galaxy with negative energy while others with positive. Systems of positive energy must disintegrate in time order of 10^5 years. This time is very short if compared with the age of the overwhelming majority of stars. Only thus must we explain that multiple stars whose internal motions have so far been studied by us possess a negative energy. But as the multiple stars of the Trapezium type are young, it is impossible to state without further study that all of them possess negative energy. On the contrary some systems of the Trapezium type, met with in stellar associations, are likely to have positive energy. For example, the star ADS 13626 in the cluster IC 4996 (Cygnus association) has visible components, whose radial velocity difference is so great, that it cannot be explained by assuming a negative total energy.

Similar reasoning is true with regard to multiple galaxies, since the age of some multiple galaxies may be such (order of billion years) that the components could not recede far from one another, although they are in the process of mutual recession. However, the final solution of the question on the existence of multiple galaxies possessing positive energy is possible only on the basis of a critical study of the actual material which, indeed, is still very poor.

We shall give, at this point, some data in favour of the positive sign of total energy of some multiple galaxies.

(*a*) If we assume that all double and multiple galaxies possess negative total energy then by observing the radial-velocity differences of their components it will be possible to come to statistical conclusions about the average masses of galaxies. Separate examination of radial-velocity differences of double galaxies and of galaxies of higher multiplicity made under this assumption has brought us to the conclusion, that the average masses of galaxies in systems of higher multiplicity are approximately three times greater than the masses of components of double galaxies. But there is no reason to believe that the nature of galaxies in systems of various multiplicity is different. The only way out from this contradiction is to assume that among systems of higher multiplicity systems with positive energy are comparatively oftener met with. Admitting their energy to be negative we get an artificial increase of the probable mass value for galaxies included in those systems.

This evidence of the existence of multiple systems with positive energy has an indirect character. Therefore, in the following we shall introduce two direct facts.

(*b*) Let us consider the group of galaxies connected with M 81. It is composed of four bright galaxies : NGC 3031 (M 81), NGC 3034 (M 82), NGC 2976 and NGC 3077 and also of some faint galaxies. The apparent integral photographic magnitudes of the four bright galaxies given above, according to Holmberg's determination ([11]), are equal to 7.85, 9.20, 10.73, 10.57. If we do not want to admit too high values of mass/luminosity ratios, we must assume that the masses of all the members of the group, except those given above, are small and therefore we may consider the group as a wide quadruple system. Its configuration corresponds to the Trapezium type. That all the four given galaxies are probably members of one physical group is evident from the following considerations. Three of them (except M 82) have radial velocities close to one another. Their mean radial velocity corrected for the solar motion is equal to + 72 km/sec. But the galaxy M 82 has a radial velocity equal to + 410 km/sec. Therefore, doubt may arise whether it does or does not belong to the group. However, there is a very close physical similarity between galaxy M 82 and galaxy NGC 3077. They both belong to the class of irregular galaxies composed of the population of the second type and both possess nearly the same high surface brightness. Because the coincidence of the characteristics indicated above is very rarely met with among the relatively bright galaxies, it is very improbable that here we have an accidental projection of M 82 in the region of the sky occupied by the group. Thus, it is almost certain that all four galaxies are physically connected with one another. Then the differences in radial velocities must be explained by orbital motions.

It is natural to assume, at the beginning, that the brightest among the four galaxies, M 81, possesses the greatest mass. But its mass is determined from its rotation by Guido Münch ([12]). It is about $10^{11}\,M_{\odot}$. The radial velocity of M 82 differs from that of M 81 by 327 km/sec. The difference of space velocities may be much more. It is not difficult to calculate that such a difference in velocities may correspond only to hyperbolic motion if the sum of the masses of galaxies M 81 and M 82 is less than 3.10^{11} solar masses. So, if we assume an elliptical movement, the mass of the galaxy M 82 must, in any case, exceed $2.10^{11}\, M_{\odot}$. Thus the dominating role in the system must be placed by the galaxy M 82. If so, difficulty arises with the galaxy NGC 3077, whose radial velocity differs from the radial velocity of M 82 by 436 km/sec and whose projected distance from M 82 is approximately 55 thousand parsecs. In order to explain this difference of velocities it must be assumed that the minimum mass of M 82 is more than 10^{12} solar masses. Such an assumption will lead to an extraordinarily great value of $\frac{M}{L}$ for M 82 (of the order of 500). Taking into account that real relat ve velocities may have considerable angles with the lines of sight we arrive at a greater mass value for M 82. The only way out of this situation is to assume that the galaxy M 82 is simply moving away from the group connected with M 81 by a velocity significantly surpassing the escape-velocity. This means that one of the members of the group has received positive energy in the process of its formation.

(*c*) The discovery by Zwicky ([13]) of the group of three galaxies, consisting of IC 3481, IC 3483 and of the anonymous one found between them, is an interesting example. Their radial velocities are equal to + 7,011 km/sec, + 33 km/sec and 7,229 km/sec respectively. The galaxy IC 3483 is a riddle. If it is physically connected with the remaining two, which is evidenced by the filament joining all the three galaxies and by the proximity of the apparent magnitudes of IC 3481 and IC 3483, then we must directly conclude that we are dealing with a galaxy moving away from the group wherein it was formed.

But if IC 3483 is accidentally projected at the end of the filament and is actually a nearby galaxy in accordance with its radial velocity, then the absolute magnitude of this galaxy must be very low. If, for example, we suppose, that it is a member of the Virgo cluster, then we must attribute to this galaxy an absolute magnitude of about —14.5. Such a low absolute magnitude is indeed unusual for spiral galaxies. Therefore, it is probable that the first assumption is correct.

(*d*) Stephan's Quintet is undoubtedly a physical group. Examining the photographs of this group we particularly notice a close connection between the components NGC 7318*a* and NGC 7318*b* of this group. In spite of that the difference of the radial velocities of these two galaxies reaches about 1,000 km/sec. Since two other galaxies of this system, NGC 7317 and NGC 7319, have radial velocities differing from the radial velocity of NGC 7318*a* by not more than 100 km/sec, we conclude that the galaxy NGC 7318*b* is leaving the group with positive energy.

(*e*) A series of narrow double galaxies are met with where it is very difficult to treat the pair as optical and yet the difference of radial velocities is great. The pair NGC 2831 and NGC 2832 may serve as an example where the distance between the components is less than 30″, which corresponds in the projection to less than 4,000 ps, while the difference of radial velocities is approximately 1,800 km/sec ([14]). However, the pair under consideration is in a cluster of galaxies where the probability of casual projection may be comparatively high, while the difference of the radial velocities of the members sometimes reaches 2,000 km/sec and more. Nevertheless, it is surprising that two so closely projecting galaxies possess so great a velocity difference.

The above mentioned facts are difficult to explain if we suppose, that in every physical multiple system all the components are retained owing to the force of mutual attraction.

Conclusion.

Among multiple galaxies there are systems in which one or more components have velocities sufficient enough to leave the system.

6. ON THE SIGN OF ENERGY IN LARGE CLUSTERS OF GALAXIES

As it is known, in order to determine the average mass of galaxies the virial theorem is often applied to clusters of galaxies. According to this theorem, the mass of cluster is determined by :

$$M = \frac{2v^2R}{G}$$

where v^2 is the average square velocity relative to the centre of the mass of the cluster, and R is the radius of the cluster. The virial theorem is applicable only to steady state clusters possessing negative energy.

It is known, on the other hand, that the application of the above formula to some clusters of galaxies leads to such values of their masses that contradict the data obtained from the rotation of galaxies. Thus, for the cluster in Virgo we obtain a mass of the order of 1,500 M_r, where M_r is the mass of our Galaxy. This means that the average mass of a galaxy in the Virgo cluster is of the order of one M_r. However, only supergiant galaxies may possess masses of the order of one M_r. Meanwhile, we know that the Virgo cluster has only a few dozens of supergiant members. Most of the members of this cluster are dwarfs whose masses are between 0.01 M_r and 0.1 M_r. This discrepancy is fully explained, if we admit that the system in Virgo has positive total energy, i.e. it represents a dissolving cluster.

We have approximately the same situation in the Coma cluster. If we apply the virial theorem, we obtain the immense figure of the order of 10,000 M_r for its mass. In this case the average mass of cluster members surpasses one M_r, which again is in contradiction with the luminosities of these members. It has lately been found that the mass of neutral hydrogen in this cluster reaches the order of 1,000 M_r. However, this does not help to eliminate the discrepancy. Therefore, here gain, we possibly have a system possessing positive energy.

Conclusion.

In some large clusters of galaxies the velocity dispersion is so great that they must represent desintegrating systems.

7. RADIOGALAXIES IN PERSEUS AND CYGNUS

If we admit the above conclusions that components of any multiple galaxy originate together and that the galaxies in some clusters and groups are in a state of mutual recession, then it is natural to conclude, that every group, immediately after its formation, represented a narrower system than we now observe. In that case two hypotheses are possible : (*a*) galaxies of a given group or of a multiple system originate from a single amorphous mass whose diameter in the order of magnitude is not less than the diameter of an average galaxy (a few thousands of parsecs); (*b*) the primary nucleus of a galaxy for reasons unknown to us is divided into separate parts which give birth to independent galaxies that become components of the system. In this case the process of division must take place in small volume with a diameter measured by some parsecs or tens of parsecs.

The parts of the divided nucleus must recede from one another, in their primary stage, with velocities of the order of hundreds or even of a thousand kilometers a second. Otherwise, their mutual attraction cannot be overcome and a few galaxies with coinciding centre arise which eventually join again into one galaxy.

Let us consider the second hypothesis in some detail.

The division of the nucleus and subsequent mutual recession of the unstable products of division (new nuclei) in the already existing galaxy must cause very violent processes continuing for some tens of millions of years. The new nuclei, during their transition to a steady state, can be imagined to eject matter which by spreading forms shells consisting of stars and gas. Thus we obtain the picture of young galaxies moving through primary galaxies. These young galaxies are in a state of formation and have rapidly growing shells.

It is exactly such a picture of violent non-steady processes that we observe in radio galaxies Cygnus A and Perseus A. The presence of intense radio-emission should be looked upon as a sign of violent processes of collision of masses of interstellar matter as well as the result of ejection of high energy particles from the atmospheres of very young stars.

In both cases we observe an enormous velocity of mutual motions. Thus galaxy NGC 1275 (Perseus A) to all appearances consists of two galaxies moving relative to each other in such a way that the difference of radial velocities, determined by Minkovsky, reaches 3,000 km/sec ([15]).

As to radio galaxy Cygnus A we directly observe two nuclei within the same galaxy. No data are, as yet, available on the velocity of the relative motion of these nuclei. However, they cannot evidently be fixed relative to each other. Besides the enormously intense radio-emission, galaxy Cygnus A radiates emission lines of high intensity and of considerable width. All of this proves the intensive motions and processes of excitation within this galaxy.

Thus the second of the above-mentioned hypotheses is roughly in line with the data obtained on the radio galaxies Cygnus A and Perseus A. Such accordance should certainly not signify a final proof of the second hypothesis, the latter requiring further comparison with observations.

As regards to the first hypothesis it is difficult to point out at present the observational data which are in conformity with the conceptions on the origin of groups of galaxies from amorphous matter. Radio emission of the neutral hydrogen in 21 cm line coming from the clusters of galaxies in Coma, Corona Borealis and Hercules proves the existence of large masses of neutral hydrogen in these clusters ([16]). However, it is uncertain as yet as to what extent these masses are independent of the separate galaxies. It is also uncertain as to what degree the intergalactic matter, radiating in the optical part of the spectrum is connected with this neutral hydrogen. Therefore, for lack of adequate data one cannot develop the first hypothesis. Subsequently we shall dwell in more detail only on the second hypothesis, i.e. on the supposed fission of the nuclei of galaxies.

It should be pointed out that the discovery of radio galaxies has served as pretext for advancing the hypothesis about the collision of formerly independent objects. In view of the fact that all radio galaxies, i.e. galaxies emitting especially intensive radio-emission, are super-giants with an absolute magnitude of the order of — 20, we have to discard this hypothesis since the mutual collisions of the dwarf galaxies ought to be much more frequent. In this respect attention should be paid to radio galaxy Perseus A which is the brightest object in the Perseus cluster where it occupies a central position. Almost the same is the role of galaxy Cygnus A in the cluster of galaxies surrounding it.

In this connection the narrow double galaxy NGC 2831 — 2832, mentioned above, deserves special consideration. At least in the projection it forms a pair of mutually penetrating galaxies with an angular separation between the centre of less than 30″. The radial velocity difference of this double, as already pointed out, attains 1,800 km/sec. It is interesting to note that this double occupies a central position in the cluster surrounding it and is endowed with a luminosity much greater than that of any of the remaining members of the cluster. The bright component of this pair is a super-giant with an absolute photographic magnitude of about — 19.5. These features are indicative of the profound resemblance of this pair with NGC 1275 where the velocity difference reaches 3,000 km/sec.

In the case of NGC 2831 — 2832 we have to deal with a double where the process of the formation of single galaxies is completed. No intensive radio emission is observed.

Conclusion.

Radio galaxies Perseus A and Cygnus A are systems in which division of nuclei has taken place but the separation of galaxies is not yet complete.

8. RADIO GALAXY VIRGO A = NGC 4486 = M 87

Two structural peculiarities in the optical region distinguish this radio galaxy from other elliptical galaxies : 1) the presence of a jet with condensations that emit polarized radiation, and 2) the presence of a great number of globular clusters ([17]).

The fact that the jet comes from the centre leaves no doubt that here we have an ejection from the nucleus of the galaxy. On the other hand polarization of the radiation shows that the mechanism of luminosity is partly, if not wholly, similar to that of the Crab nebula. It follows that the sources of luminosity in the condensations consist not only of stars but also of diffuse matter which is in the same state as in the Crab nebula. In other words, a considerable amount of high energy electrons should be expected in these condensations. On the other hand the sources of radio emission are known to be concentrated continuously in the body of galaxy NGC 4486.

Two conjectures are likely to be made : (*a*) the high energy electrons were directly emitted from the nucleus of the galaxy, and (*b*) objects are ejected from the nucleus which are the sources of the electrons of very high energy, a considerable part of the synchrotrone radiation of which is concentrated in the optical region.

One cannot confine oneself to the first hypothesis since the concentration of optical radiation in a small volume of condensations would remain unexplained. Therefore, one should believe that the sources emitting electrons of high energy are concentrated within the condensations themselves. Observations of the objects of our Galaxy reveal that various non-steady stars (supernovae, stars of the T Tauri type and others as well) constitute powerful sources of high energy electrons. Therefore, it is very likely that a large number of similar non-steady objects are present in the condensations referred to. Thus we seem to get closer to an understanding of the nature of the condensations under consideration. They represent conglomerates of clouds of relativistic electrons, gaseous clouds and non-steady stars. Such conglomerates can hardly be assumed to exist in the nuclei of galaxies. It should be inferred, therefore, that the matter ejected from the nucleus evolves, over a short period, into such conglomerates. The emission line λ 3727 noticed in the region of the nucleus of NGC 4486 gives, apparently, some idea of the velocity of the ejection from the nucleus. Hence one can evaluate the order of the periods in the course of which such conversions take place. They turn out to be of the order of 3.10^6 years.

Conclusion.

Apart from the division of the nuclei of galaxies, processes of relatively small masses from the nuclei of galaxies may occur in nature. These ejected masses can, over a short period, turn into conglomerates made up of young non-steady stars, interstellar gas and clouds of high energy particles.

9. BLUE EJECTIONS FROM THE NUCLEI OF ELLIPTICAL GALAXIES

The galaxy NGC 4486 is not the only one where an ejection of matter from the nucleus is observed. We have paid attention to some other similar cases of which galaxy NGC 3561*a* is of especial interest ([18]). That galaxy is apparently spherical and has an outflow that looks like a jet. The jet ends in a condensation which is bright enough on the blue photograph and almost invisible on a red one. The colour index of the condensation in the international system is $-0^m.5$. The distance of galaxy NGC 3561*a* is unknown to us. However, a very cautious estimate based on a comparison of the apparent magnitude of the bleu condensation with that of brighter galaxies of the same cluster including 3561*a*, makes it possible to consider the absolute photographic magnitude of the condensation not less than — 14.5. This means that the blue condensation in question is not a common O-association. Judging from its absolute magnitude this ejection is, as a matter of fact, a dwarf galaxy that has evidently got detached from the nucleus of a giant galaxy. The extraordinary value of the colour index of ejection indicates that it is composed of population entirely peculiar to itself. That its blue colour can be accounted for by a short wave continuous emission is not ruled out either. No doubt that galaxy NGC 3561*a* deserves further study.

It is already known that the ejection observed in NGC 4486 is also, though to a small degree, bluer than the main galaxy. Therefore, it has been deemed expedient to try to search for similar blue objects in areas surrounding other elliptical galaxies. About a dozen and a half blue companions have been found out which, as a rule, are not bound by means of a jet to the elliptical galaxy and possess negative colour index. The luminosity of a substantial number of these objecfs exceeds by far that of the common stellar associations. They can be taken as separate galaxies.

This does not mean that the ejections from the central regions of elliptical galaxies cannot be yellow or even red. However, it is difficult to distinguish ejections with positive colour indices from weak galaxies of the distant background.

The ejections and the companions under consideration, unlike those of NGC 4486, are projected beyond the confines of the brighter parts of the corresponding galaxy. Sometimes they are over a distance of several radii from the main galaxy. These distant companions are probably older than the ejections in NGC 4486. Perhaps this is the reason why we do not observe the radio emission of these objects.

Conclusion.

In certain cases ejections from central parts of elliptical galaxies have very blue colour. This colour may be the result of preponderance of the great number of highly luminous blue stars or of the strong continuous violet emission. In both cases it cannot last very long. It is, therefore, highly probable that these blue ejections and companions are very young galaxies.

10. BRIDGES AND FILAMENTS JOINING GALAXIES

The valuable service of Zwicky ([13]) is connected with the investigation of many double and triple galaxies joined together by means of filaments or bridges of various thickness. Zwicky himself is inclined to believe that these bridges were formed as a result of the tidal interaction arising from the encounter of two galaxies. According to Zwicky bridges and filaments consist of stars ejected as a result of tidal action from the given galaxy. It is not hard to see that such an interpretation is at variance with the facts. As a matter of fact filaments joining two galaxies are sometimes very thin. Even if we assume that the tidal wave has been ejected as a stream from the surface of the given galaxy out of narrow localized region, and, therefore, should originally have been very thin, yet the resulting filament ought to go on widening in consequence of inherent velocity dispersion. The ratio of the thickness to be length at the end of the stream should be of the same order as that between the velocity dispersion of stars and the velocity of the stream.

The velocity of the stream, in its turn, must not exceed the relative velocities during the encounter of the galaxies. In most cases the relative velocities must have an order of magnitude not higher than 200 km/sec. Keenan's system, for example, where radial velocity difference is equal 22 km/sec may serve as an illustration. On the other hand, the velocity dispersion of stars in a small volume of a galaxy must be of the order of 30 km/sec. It follows, therefore, that the width of the filament at its end must be of the order of one-sixth of the length of the filament. Meanwhile the width of the filament in Keenan's system is many times less.

In most systems the joining bridge or filament forms a continuation of the spiral arms. Thus, the assumption that bridges and filaments are of tidal origin leads one, in fact, to the conclusion that the spiral arms are also the products of tidal interaction. It should be natural to apply this also to all the remaining spiral galaxies, i. e. to those that are not linked with other galaxies by means of bridges or filaments. Such a conclusion, however, might arouse strong objections. For instance, it is known that in dense clusters of galaxies (for example in Coma cluster) where tidal interactions are more frequent, there are very few spiral galaxies. On the contrary, they are in considerable numbers in loose groups and clusters.

Consequently the conception that tidal interactions account for the formation of filaments must be given up. In light of the hypothesis on the formation of galaxies after the division of a primary nucleus filaments should be regarded as the ast bond still linking together galaxies that have already separated and are at considerable distances from one another.

If the filament linking, for example, a pair of spiral galaxies, appears during the process of the division of a single primary nucleus the spiral structure of the galaxies formed should also be closely bound with the process of division. One can believe that the connection between duplicity and the spiral structure must also be present in cases when one of the components is not a spiral galaxy but belongs to some other class.

Finally, it should be pointed out that although the nature of the spiral structure in many cases is connected with the duplicity or multiplicity of galaxies further investigation is required to confirm the fact that *all* spiral structures are the result of such a division.

Conclusion.

Bridges and filaments between galaxies are not the result of tidal interactions. They can be supposed to arise during the process of mutual recession of two or more galaxies that have originated from the same nucleus.

11. TYPE M 51 GALAXIES

The presence of the companion NGC 5195 at the end of the spiral arm of the galaxy M 51 has always seemed to us to be a strong argument in support of the concepts put forth in the foregoing paragraph. That the spiral arm discontinues almost immediately beyond NGC 5195 furnishes, in our opinion, a serious argument against the conjecture of NGC 5195 accidentally projecting on the arm of the spiral galaxy NGC 5194. It is desirable, however, to find out another case in which the connection between the spiral structure and the presence of the companion should prove more convincing. Such a case has been found out by a student of mine — Iskoudarian — on the prints of the Palomar atlas. This refers to the double galaxy NGC 7752 — 7753. On the blue print the spiral arm is made up of three parallel filaments which disrupt simultaneously as soon as they reach the companion. Two of the three filaments are directed to the centre of the elliptical companion while the third filament running parallel to the former two comes very close to the periphery of the elliptical cluster where it takes a sharp turn to its centre. Large photographs can, of course, give a more accurate idea of the phenomenon and specify certain details. However, the existence of connection between the elliptical companion and the spiral arm is evident.

The similarity between the double galaxy under consideration and M 51 is emphasized by the fact that in both cases, the curvature of the spiral arm sharply decreases in the region adjacent to the companion. Thus, formations of the M 51 type should not be considered the result of mere projections. As it was pointed out by B.A. Vorontsov-Velyaminov ([23]), they represent a class of double galaxies in which the components are bound together by means of a powerful spiral arm and not by a thin filament. Perhaps this is partly connected with the fact that the distance between the two companions, at least at the given stage of the evolution of the group, is small. In the case of M 51 this distance is of the order of three thousand parsecs. As the distance between the components increases the bridge becomes considerably thinner.

Conclusion.

The existence of galaxies of the M 51 type confirms the suggestion on the connection between the process of the division of the original nucleus and the formation of spiral arms.

12. BIG CONDENSATIONS IN SPIRAL ARMS

Galaxies of the S*c* type and those with still more disrupted arms often contain bright condensations which are rich stellar associations. The associations of hot giants of an absolute magnitude of — 11 are already very bright objects. But in certain cases galaxies of the S*c* type contain condensations of still greater luminosity.

Condensations of an absolute magnitude of about — 14 can be compared with separate galaxies. In other words, such condensations can be looked upon as companions of the galaxy while the latter is to be regarded as a certain kind of multiple group. Thus, there is no clear-cut boundary between usual condensations in the arms and companion galaxies.

NGC 4861, NGC 2366 and others can serve as examples of galaxies with very bright condensations.

As it is known the galaxy IC 1613 which is a member of Local group has on its periphery a formation consisting of a considerable number of O associations. This formation is a kind of superassociation. A similar superassociation representing a whole constellation of O associations is observed on the periphery of spiral IC 2574. Such superassociations have sizes quite comparable with whole galaxies and we may consider them as companions of corresponding central galaxies.

The objects we have mentioned in this paragraph represent in some degree population I counterparts of the companion of spiral M 51. Evidently, these objects could have originated only as a result of the separation of a considerable and compact mass from a primary central nucleus. Apparently, it is impossible to explain the existence of superassociations of the above mentioned type when we assume that their stars have been formed from purely diffuse (gaseous) clouds. A diffuse cloud of such a large size, after separation from the central nuclei should, indeed, have dispersed over the whole volume of the galaxy owing ot the effect of differential rotation.

Conclusion.

Besides cases where a spiral arm links a given galaxy with companions consisting of population of the second type, there are other cases where a spiral arm ends in a companion which presents a large conglomerate of objects of population of the first type.

13. NATURE OF NUCLEI OF GALAXIES

Our knowledge of the nuclei of galaxies is very poor. Referring to the nuclei we mean very small formations a few parsecs in diameter with a very high surface brightness in the centrum of a galaxy ([19]). Lats year Dr. Baade was kind enough to show me a photograph of the nucleus of the galaxy M 31. This is, in fact, an amazing formation with an exceedingly bright surface luminosity. Unfortunately, in most galaxies, we are unable to pick out the nucleus from the whole central body of the stellar system.

We concluded above that the nuclei can be divided and also eject spiral arms or radial jets containing certain condensations. But the spontaneous division of a stellar system consisting only of stars is dynamically impossible. Therefore, if the nucleus were made up of only common stars we should have to give up the concepts put forth above in which the basic role in the genesis of galaxies and the formation of spiral arm is ascribed to nuclei. A grave difficulty arises from the fact that in the region of the nucleus the amount of interstellar hydrogen is small as compared to the density of the interstellar hydrogen in the outer parts of stellar systems such as the arms of our Galaxy. Meanwhile in some instances the outflow of the matter from the nucleus can be observed almost immediately. I mean not only the jet in NGC 4486 and in NGC 3561 but also the outflow from the central part of our Galaxy of interstellar hydrogen with considerable velocity as discovered by Dutch astronomers. Thus, it turns out that hydrogen flows out from where there is none. To form a fuller idea of the difficulties involved we must realize that the spiral arms of galaxies contain large masses of hydrogen and also, irrespective of any hypothesis, that there is a definite genetical bond between the arms and the central nuclei.

We must consider this as one of the greatest difficulties in astrophysics which may be solved provided a radical change in the conception on the nuclei of galaxies is adopted.

Apparently we must reject the idea that the nucleus of a galaxy is composed of common stars alone. We must admit that highly massive bodies are members of the nucleus which are capable not only of splitting into parts that move away at a great velocity but also of ejecting condensations of matter containing a mass many times exceeding that of the Sun.

The new bodies, resulting from splitting or ejection, move away from the volume of the original nucleus with velocities, which are sufficient to overcome the attractive force of the nucleus and eject considerable masses of gases as well as some denser condensations. After some time these condensations can come to a state of quasi-stability under the influence of their own gravitation, i. e. turn into stars.

Not all the transformations referred to above should end immediately after the formation of the spiral arm or of the new galaxy. In some cases these transformations are very likely to delay in consequence of the transition of a number of fragments into some kinds of meta-stable states and then only would they turn into stars and gas. It is this last transformation that we probably observe in the arms of our Galaxy as phenomena of the origin of stars and nebulae in stellar associations. This refers to both O- and T-associations. This point of view may raise some objections. One can say that it is difficult to offer, at present, a model of massive bodies with the above described properties. Even if we should not try to understand directly the concrete mechanism of the division of massive bodies located within the nuclei, nevertheless, difficulties can be met connected with the preservation of the various laws of conservation such as that of the rotational momentum. On the other hand, it is quite possible that the consideration of the common origin of two or more stellar systems might be instrumental in overcoming these difficulties.

The basic idea we want to emphasize lies in the fact that before starting the construction of theories of the origin of galaxies it is necessary to determine from observations the type and character of the processes leading to the formation of new stellar systems. Next, the problem of theoretical interpretation of the processes observed should arise.

Conclusion.

There are evidences in favour of the formation of new galaxies and arms at the expense of matter contained in nuclei of galaxies. These nuclei are very small in size and of high density. Since these processes of the genesis of stellar systems cannot occur at the expense of the common stellar popilation of a nucleus we should conclude that, large masses of prestellar matter are present in nuclei.

14. ON THE REPETITION OF THE PROCESSES OF FORMATION OF COMPONENTS AND ARMS

Many spiral galaxies have a complex structure which evidences the fact that the processes of ejections and outflows from their nuclei took place at different times and in various ways. For instance, the spiral arms of our Galaxy and their stellar population are concentrated in the fundamental plane of Galaxy, but the Magellanic Clouds and weak spiral arm connected with them, are located in a quite different plane. Thus, it seems that the cosmogonic process connected with the origin of the arms of our Galaxy occurred twice.

Although we have not at our disposal data on the special structure of other big spiral galaxies yet a review of the images of a large number of outer galaxies leads to the impression that the spiral structure is not always concentrated in only one plane. This is particularly true of galaxies with inner and outer spiral structures. In some cases their planes do not coincide. These phenomena show that after a spiral structure of galaxy has come into existence its nucleus and perhaps also the fragments moving away from it preserve their potentiality as centres of the active cosmogonic process. On the other hand, there are, no doubt, nuclei, which have lost this power. Finally, there exist galaxies without nuclei (such as the Sculptor type group) where there can even be no question of the formation of new structural elements. Such a gradation of the intensity of the cosmogonic activity of nuclei is, apparently, to some degree, connected with the values of masses and luminosities of galaxies. Supergiant galaxies must possess most active nuclei. Then it becomes comprehensible why all radio galaxies are supergiants. However, it is possible that there are objects with different activities among galaxies of a given mass.

15. ON THE ROLE OF INTERSTELLAR GAS

As the radio observations in 21 cm line of interstellar hydrogen show, the interstellar gas composes a considerable part of the masses of spirals and of irregular galaxies. Confronting this with the richness of these systems or associations, one usually concludes that young stars are being formed from interstellar gas.

However, the parallelism between the abundances of interstellar gas and O association permits two different interpretations : (*a*) The formation of stars from gas, and (*b*) The common origin of stars and of interstellar gas from some protostellar bodies. Therefore, of the greatest value are the data which enable us to choose between these two interpretations. Let us consider here some such data.

(*a*) The Double cluster association in Perseus is situated in a region of Galaxy essentially devoid of interstellar gas. This is confirmed by observations made with nebular spectographs as well as by 21 cm observations. At the same time, the association in Perseus is one of the richest in our Galaxy. It is especially rich in supergiants. The assumption that the formation of the association has immeiately caused the exhaustion of gas seems too artificial. Moreover, the presence of bright supergiants with the ages of the order of 10^6 years indicates that the star formation in this association is continuing at the present time. But this is incompatible with the hypothesis of star formation from gas.

(*b*) The mean density of interstellar gas in the Small Magellanic Cloud is not lower, but probably higher than the corresponding density in the Large Magellanic Cloud ([20]). At the same time, in the Large Magellanic Cloud the abundance of O associations and especially of the associations consisting of high luminosity stars is very much higher. It is impossible to suppose that in the Small Cloud the associations did not have sufficient time to originate from gas. In fact, the time necessary for the formation of associations must be at most of the order of 10^7 years. Meanwhile, the present distribution of gas in Small Cloud must have the duration of the order of 10^8 years. In addition, a number of O associations is observed in the Small Cloud, but they are comparatively poor in high luminosity stars.

(*c*) The observations show that the distribution of neutral hydrogen in Galaxy is better correlated with the distribution of classical Cepheids than with the distribution of O associations. The Small Magellanic Cloud contains particularly a large number of classical Cepheids. There is no doubt, therefore, that the origin of classical Cepheids is, in some way, connected with interstellar gas. If stars originate from gas, then this must be interpreted as an evidence that the transformation of gas into stars has been going on in the Small Cloud for a long time. This makes the contradiction, mentioned in the foregoing point, still sharper.

(*d*) G. Münch has given attention to the fact that in M 13 and in other globular clusters of our Galaxy some high luminosity blue stars are present. But on great distances from the galactic plane the density of interstellar gas must be exceedingly low while the dispersion of turbulent velocities must be very high.

The data mentioned above contradict the formation of associations from gas. At the same time, we do not insist that these facts directly confirm the hypothesis of common origin of stars and gas from protostars of unknown nature. On the other hand, the parallelism between the abundances of gas and association is a strong support in favour of a genetic connection between them. Therefore, the hypothesis of common origin of stars and interstellar gas is the only possible solution.

Conclusion.

The facts connected with interstellar gas and associations are speaking in favour of common formation of stars and gas from protostars rather than of formation of stars from gas.

16. ON THE ORIGIN OF POPULATION II STARS

There are observational indications that globular clusters are moving in Galaxy in highly elongated orbits. This may serve as a direct evidence that they were some time ejected from the nucleus of our Galaxy with velocities of the order of some hundreds of km/sec. This shows that globular clusters have been formed not in a thin diffuse cloud, but in a volume of sufficiently high density.

Probably, further investigation of globular clusters will give us possibility to approach nearer to the solution of the problem of the origin of II type population. In this report we shall confine ourselves to only some remarks concerning this question.

(*a*) All globular clusters observed in our Galaxy must be systems of negative total energy. Had some of them been of positive energy, we should have observed clusters consisting of the same population but having much larger volume and much smaller density. We observe no such clusters.

(*b*) On the other hand, let us suppose that globular clusters have been formed out of masses ejected from the nucleus of Galaxy and that the transformation of these masses into stars proceeds right after the ejection immediately near the centre of Galaxy. In this case, the clusters of positive energy would expand to a diameter of about 1,000 ps. before reaching the distance of 10,000 ps. from the centre of Galaxy. Such loose stellar clouds will be difficult to discover on the general background of our stellar system. Therefore, the possibility that in the central regions of our Galaxy the formation of globular clusters of both signs of energy is continuing, is not excluded.

Prof. Parenago and others have shown that a small number of population II stars have velocities exceeding the velocity of escape from Galaxy. This means that the age of these stars does not exceed 10^8 years. Therefore, we cannot agree that all stars of spherical subsystems are « old », i. e. have ages of the order of some billions of years.

(*c*) The very distant globular clusters (intergalactic tramps) discovered on the photographs of the Palomar Sky Atlas are of considerable interest. Some of them are found to be on distances of about 125 thousands ps. If they had originated in the nucleus of our Galaxy and had kinetic energy insufficient to escape from the gravitational field of galaxy, the time to reach such distances should have been of the order of a billion years. We consider it improbable that they have originated in other galaxies, but, if so, much longer time would have been necessary to reach their present position. Consequently, all these clusters represent groups of old stars. According to G.R. Burbidge and E.M. Burbidge the colour luminosity diagrams of these groups are considerably different from those of nearer clusters ([21]).

(*d*) We know one case when a peculiar elliptic galaxy (NGC 4486) is particularly rich in globular clusters. Giant spirals often have hundreds of globular clusters within them. We may add that there is no indication that S*b* spirals are much poorer in globular clusters than conventional elliptic nebulae.

As regards late spirals and irregulars they also contain a considerable number of globular clusters. Thus, according to Kron the Small Magellanic Cloud contains not less than ten globular clusters. The Large Cloud, apparently, contains several dozens of them. Therefore, if one compares the number of globular clusters per unit of integral luminosity, or per unit of mass of galaxies, the Clouds will, in no case, appear poorer than our Galaxy or M 31. Possibly, they will appear richer, in this respect, than the majority of giant ellipticals. If we now take into account that the greater part of the luminosity of Clouds is produced by population of « flat subsystems », it may occur that the number of globular clusters per unit of luminosity produced by population II stars in Magellanic Clouds is substantially greater than the corresponding number in early type spirals and ellipticals. This will mean a very essential difference in the internal composition of populations of spherical subsystems in corresponding galaxies.

In other words, the percentage of population II stars contained in globular clusters will be higher in the Magellanic Clouds than in the early type spirals and ellipticals (possibly except NGC 4486).

Nevertheless, it is clear that the number of globulars per unit mass is changing slowly with the type of galaxies. In contrast to this the number of associations is changing rapidly. Therefore, we are justified to conclude that *the mechanisms of star formation in flat and spherical subsystems are not only different, but, to a considerable degree, independent of each other*. This conclusion is in complete accordance with the views developed by Kukarkin ([22]).

17. ON DOUBLE SPIRALS

We have admitted above that the formation of spiral arms is connected with the formation of a double galaxy. Then, the following possibility of verification of the hypothesis on the division of primary nucleus arises. Since, the primary nucleus, having a small volume, cannot have the rotational momentum comparable with the rotational momenta of spiral galaxies, we may assume that during the formation of two spirals the sum of the momenta will remain small. This condition is easily satisfied when the spirals formed have momenta of opposite directions. In this case, we must expect that the directions of wending of spiral arms should be opposite to each other, that is, the angle between these directions will be near to 180°.

When comparing this conclusion with observations we must take into account the following circumstances : 1) A pair of spirals chosen for comparison must be isolated. If three galaxies of comparable sizes have been formed, the momentum obtained by the third galaxy could compensate the total momentum of the pair under consideration; 2) We must be sure that each observed pair is a physical double. 3) If the inclinations of two spirals to the plane of projection are near to 90°, the small deviation of the real angle between momenta of two galaxies from 180° may lead to the consequence that these spirals will appear as wound in the same direction. Such pairs, therefore, should not be taken into account. 4) The direction of winding of spiral arms must be clearly expressed.

We have picked up 20 pairs of comparatively bright spirals on the charts of Palomar Sky Atlas (which at our observatory is still incomplete) to some degree satisfying the above conditions. Only in three pairs of the 20 we observe the same direction of arms. But, in these cases we are not sure that all the requirements enumerated above are strictly fulfilled. On the other hand, among the remaining 17 cases there are some pairs for which the above conditions are fulfilled almost rigorously. The pairs NGC 2207 and IC 2163, NGC 4618 — 4625, NGC 5394 — 5395 belong to this group. From these data we cannot finally conclude whether the rule of opposite directions of arms in isolated pairs is fulfilled strictly or it is obeyed only in the great majority of cases. But in any case, this rule in some form is valid.

As it was mentioned above the pairs under consideration must be isolated, though galaxies of very small mass and extent may be present in the vicinity of the pair. It is interesting that in this sense the galaxies M 31 and M 33 can also be considered as an isolated pair. It is a matter of fact that they show opposite direction of spiral arms.

If, however, we do not confine ourselves to pairs isolated in space, we can consider the case when spirals are connected by means of a bridge. In Wild's triple system the pairs connected by bridges clearly show the opposite direction of arms. In the well known cluster of galaxies of Hercules we have a remarkable twin of joined spirals which obeys the same rule. It seems that this question deserves further attention.

Conclusion.

In the overwhelming majority of cases physical pairs of spiral galaxies have opposite directions of spiral arms.

REFERENCES

([1]) V.A. Ambartsoumian, Outchenie zapiski LGU, nº 22, *Seria matematitcheskikh naouk* (astronomia), vyp. **4**, 19-22 (1938).

([2]) V.A. Ambartsoumian, *Astronomitcheskii Journal*, **14**, 207-219, (1937).

([3]) V.A. Ambartsoumian, *Evolioutsia zvezd i astrophisika*, Erevan (1947).

([4]) V.A. Ambartsoumian, *Astronomitcheskii Journal*, **26**, 3, (1949). A. Blaaw, BAN, **11**, 405 (1952).

([5]) F. Zwicky, PASP, **50**, 218 (1938); **64**, 247 (1952).

([6]) J. Neyman, E.L., Scott and C.D. Shane, *Ap. J.*, **117**, 92 (1953).

([7]) V.A. Ambartsoumian, *DAN Armianskoi SSR*, **13**, 129 (1951); *Soobchtchenia Biourakanskoi Observatorii*, nº 15 (1954).

([8]) E. Holmberg, *Annals of the Observatory Lund*, nº 6 (1937).

([9]) V.A. Ambartsoumian, *Izvestia AN Armianskoi SSR*, seria phisiko-matematitcheskikh naouk, **9**, 23 (1956).

([10]) C. Seyfert, PASP, **63**, 72 (1951).

([11]) E. Holmberg, *Meddelande Lund Observatory*, Ser. II, nº 136 (1958).

([12]) Guido Münch, *A.J.*, **62**, 28 (1957); « Report of the Director, Mount Wilson and Palomar Observatories » (1955-1956), p. 49.

([13]) F. Zwicky, *Ergebnisse d. exakt Naturwissenschaften*, Bd. **29**, 344 (1956).

([14]) M. L. Humason, N.U., Mayall and A.R. Sandage, *A.J.*, **61**, 97 (1956).

([15]) R. Minkowski, « Report to the meeting of the AS Pacific », Pasadena, June 21, 22 (1955).

([16]) D.S. Heeschen, *Ap. J.*, **124**, 660 (1956); *PASP*, **69**, 350 (1957).

([17]) W. Baade and R. Minkowski, *Ap. J.*, 119, 215 (1954).

([18]) V.A. Ambartsoumian and R.K. Chakhbazian, *DAN Armianskoi SSR*, **25**, 185 (1957).

([19]) W. Baade, *IAU symposium*, nº 5, p. 1, Cambridge (1958).

([20]) W. Buscombe, S.C.B. Gascoigne and G. de Vaucouleurs, « Problems of the Magellanic Clouds », *Austr. Journal of Sc.*, Supplement (1955).

([21]) C.R. Burbidge and E.M. Burbidge, *Handbuch der Physik*, vol. 51, in preparation.

22) B.V. Kukarkin. « Issledovanie stroenia i razvitia zvezdnykh sistem » M.-L. (1949).

([23]) B.A. Vorontsov-Veliaminov, *Astr. Journal*, **34**, 8 (1957).

Discussion of Ambartsumian's Report

Dr. Morgan. — The meaning of the terms « Population I » and « Population II » may have to be reexamined if we observe in the universe galaxies in very different stages of evolution. In those of very early stage, we may now be observing the formation of globular clusters or, rather, objects that would be considered globular clusters in a later stage of their evolutionary development. It might be difficult to assign such objects uniquely to either of the two population groups.

Dr. Oort. — I agree with Dr. Ambartsumian that there are phenomena which seem to indicate that something so enigmatic is going on that one is led to suspect that at least in some cases stars do not simply originate by condensation of interstellar gas. I think in this connection in particular of the extreme youth of the Orion nebula and the high velocities of some young stars apparently ejected from the Orion region. However, I feel some doubt whether there are sufficiently compelling phenomena in the world of galaxies to justify the adoption of such a revolutionary idea as the fission of galaxies. I do not yet see that there is much objection to the view that in most of the radio sources connected with interacting galaxies the interaction is an actual collision. Dr. Ambartsumian has pointed out that the galaxies connected with such radio sources are always objects of exceptional brightness. But this does not speak against the collision idea. It might only mean that in order to get a radio source we need not only a collision, but a collision between giant galaxies, possibly because these giant galaxies possess some characteristic that is essential for the creation of an intense radio source (e.g. a sufficiently large « halo »).

As regards the possibility that some clusters of galaxies may have positive energies like the expanding stellar associations in our galaxy, I agree that this cannot perhaps be quite disproved by the available data. However, if the Virgo cluster had positive energy it could not be older than about 1/10th of the age of the universe. Measurement of the colours of the globular clusters in M 87 might well decide between so young an age and the age of about 10×10^9 years which I would consider to be more propable. And, probably, the fact that the colours of all elliptical galaxies that we know are identical also indicates that the Virgo galaxies cannot be so much younger.

Dr. Gold. — The expanding associations of young stars and associated great masses of gas certainly require an explanation in terms of some process with which we are not yet familiar.

A speculation some of us have been persuing is a nuclear explosion following the gravitational condensation of very large masses of stars and dust. If masses of gas of the order of 10,000 solar masses were to form gravitational systems it can be estimated that the later phases would involve a very rapid Helmholz contraction. The central density and pressure may reach values that imply the release of a large amount of nuclear energy. The system may then blow itself apart, and in this phase there may be enough instability in the flow to leave fragments of stellar dimensions as self-gravitating objects.

This speculation is made particularly attractive for the following further reasons :

1) Some stars in expanding associations have velocities in excess of those that could be attributed to any gravitational interactions.

2) The kinetic energy of gas masses in the galaxy seems somewhat in excess of that which can be accounted for by the stirring up that O stars will cause.

3) The amount of helium in the galaxy seems to be high for the total luminosity and the age estimated.

4) The distribution due to other processes through the galaxy of the elements below iron in atomic weight is possibly not quite adequate. The development new proposed may actually assist both the nuclear and the distribution processes.

From a theoretical point of view one may say that it is easier to have very massive condensations rather than those of stellar masses from the original gas. The presence of dust may play an important part in providing radiating facilities in the early stages to assure that the contraction proceeds far enough to withstand further disturbance by the turbulent motions in the vicinity.

The brightness of such an explosion may not be too great, as it is not certain that radiation from the interior explosion would be able to penetrate the great depth of outer layers sufficiently quickly. In any case the phenomenon is rather rare, perhaps 1,000 times rarer than supernorvae, and it would thus not be easy to notice it directly in other galaxies.

The disposal of the angular momentum is the greatest difficulty; but long strung-out magnetic fields along which the initial contraction takes place may be able to take the angular momentum out in the early stages.

One may even wonder whether explosions on a still larger scale could not occur, and this would bring the speculations closer to those of Dr. Ambartsumian. Perhaps even entire E galaxies are generated in single events.

Dr. Bondi. — I have started calculations on this model in the stage when it has overcome the serious initial difficulties pointed out by Gold but before the nuclear reactions start. The early contraction would appear to be rapidly accelerating, whereas later the motion would continue at more or less constant speed, supplying by gravitational contraction the energy required for the heating and the luminosity of the body. Nuclear reactions starting in such conditions of rapid inflow would be bound to lead to a very severe explosion.

Dr. Mac Crea. — The theory of star formation to which I referred this morning is essentially the same as that just described by Gold. The suggestion that some release of nuclear energy is required in order to produce an expanding cluster was, I believe, due originally to Gold. Here I would mention only two points. The order of pressure to be met with in hot regions (H II) of interstellar matter can initiate the collapse of a mass of several hundred solar masses, but not of just a few solar masses. Secondly, I think it is true that if in a temporary large « star » the amount of energy released is about that which would be given by the nuclear energy in a mass equal to the Chandrasekhar limit, then this would be sufficient to account for the kinetic energy of the finally resulting group of ordinary stars.

On quite a different problem, we may recall that A. Salam has said that nuclear physicists would not be surprised if nuclear interactions of a certain order of « weakness » (in a technical sense) were to give rise to the creation of new matter as required by steady-state cosmology. This in turn suggests that the creation of new matter might be correlated with condensations of already existing matters. Previously, this idea was not acceptable for the steady-state theory because it would apparently not lead to fresh condensations. However, if the new matter appears with a high kinetic temperature, so that it can disperse rapidly, this objection might no longer hold good. I have not worked out any consequences, but I wonder if such ideas could have any connexion with those presented by Ambartsumian.

Dr. Oppenheimer. — I do agree with Salam that we would not be astonished if, at very low rate, Baryon number was not conserved. I think these are things we know of too little to say where, when or how they occur.

Dr. Oort. — Gold's suggestion of the explosion of condensing masses of gas that are too large to form a star is an interesting one. I should like to draw his attention to the fact that many elliptical galaxies have dense cores of gas. They have recently been studied by Osterbrock on Palomar. The masses are of the order of 10^6 times the solar mass or higher. In some cases these cores show a fast rotation, in other cases they do not show rotation. It might be imagined that the jet in M 87 would be part of the result of an explosion of a nuclear « superstar » of the type suggested.

Dr. Hoyle. — The explosion that Gold spoke about is presumably a rather gentle affair, not one in which there is an enormous release of nuclear energy — as in the case of a supernova. For this reason it seems as if the problem might be one quasi-stability and so that it was worth while looking at the stable solutions for very large masses. The interesting result appeared that the stable configurations are entirely convective. Hence any nuclear products would be uniformly distributed throughout the material. This means that if a large mass were to break into an 0 association, the resulting stars would probably be similar in their initial composition.

Dr. Schatzman. — You mentionned the fact that the age of the galaxies is several times the period of revolution of multiple galaxies. Does it mean that after a few billion years galaxies disappear as such?

I want to come back to the question of angular momenta. In case of star formation, it seems difficult to imagine that a star of maximum angular momentum could be formed from a denser state of matter.

It might be suggested for example that during contraction angular momentum is being stored in the internal angular momentum of matter. But that seems very difficult as the Planck constant is very small; or, to put it in other words, the quantum numbers corresponding to the rotation of a star are of the orders of 10^{20} so that the orbital angular momentum is very great. Could the physicists imagine a way of storing angular momentum in matter?

Origin of the Elements in Stars

by **F. HOYLE**

1. INTRODUCTORY REMARKS ON STELLAR EVOLUTION

Stars are known to be forming at the present time from the interstellar gas, and there is strong evidence that stars have been formed within our galaxy at all epochs since its origin.

The reverse process also occurs whereby stars return material to the interstellar gas. Easily the most striking phenomenon of this kind is the supernova, an explosion in which the main body of a star is shattered and blown apart with speeds that range from 1,000 to 5,000 km per sec. On the basis of rather fragmentary evidence it can be estimated that the number of such explosions that have occurred during the lifetime of the Galaxy probably exceeds 10^7 and may be as high as 10^8. With each supernova distributing a mass of the order of the Sun, it follows that this particular process has returned from 10^7 to $10^8{}_{\odot}$ of material from the stars to the interstellar gas.

Other processes whereby stars return material to the interstellar gas have been proposed. Of these, probably the most important is mass loss by giants. A giant is a star with quite dense inner regions (a central density of 10^5 gm. per cm^3 would be typical), but with a low density envelope of very large diameter, in some cases with diameter several hundred times the Sun. This giant structure occurs as a well marked phase of stellar evolution; its development has been traced by precise calculation.

Because of the large diameter of the giant envelope it is comparatively easy for material to escape to infinity from the surface of a giant. Thus to escape from the surface of a typical giant the necessary dynamical energy is only about one percent of that which is required to escape from the Sun. This arises of course because of the low value of the gravitational potential at the surface of a giant. Observations by Deutsch show that this process really does take place. The process is comparatively quiescent, not explosive as in the case of a supernova.

Estimates of the effectiveness of mass loss from giants are even more uncertain than in the case of supernovae. It seems likely however that the total mass returned to the interstellar medium in this way is very considerable, perhaps of the order $10^9{}_\odot$ or even more, taken over the whole lifetime of the Galaxy.

The present mass of the interstellar medium is also of order $10^9{}_\odot$, from which it follows that the circulation of material between stars and the medium must be considered an important process. Even in the early history of the Galaxy, when the total mass of the interstellar gas must have been considerably larger than it is at present, the degree of circulation must still have been important, amounting to a total turn over of perhaps 10 percent.

The point of view of this report is that all elements but hydrogen are synthesized within the stars. At the time of formation of the Galaxy the « first » stars contained little except hydrogen (N.B. In steady state cosmology it is not to be expected that the « first » stars would be entirely hydrogen, since there would be some slight contamination of the galactic material, due to the elements distributed by stars in previously existing galaxies). Subsequent generations of stars acquired progressively higher and higher concentrations of helium, carbon, iron, ... as more and more stars returned these products of synthesis to the interstellar gas.

This outlook is in good agreement with observation. Stars that by other criteria are thought to be old do have low concentrations of the heavier elements. In particular, Greenstein and his collaborators have investigated cases in which the concentrations are less than 1 percent of those found in the Sun.

In the present connection it is to be noted that a star possesses a self-governing mechanism whereby its internal temperature is adjusted so that the outflow of energy through the star is balanced by nuclear energy generation. The temperature required to give this adjustment depends on the particular nuclear fuel available. Hydrogen requires a lower temperature than helium; helium requires a lower temperature than carbon, and so on, the increasing temperature sequence ending at iron since energy generation by fusion processes ends there. If hydrogen is present the temperature is adjusted to hydrogen as a fuel, and is comparatively low. But if hydrogen becomes exhausted as stellar evolution proceeds, the temperature rises until helium becomes effective as a fuel. When helium becomes exhausted the temperature rises still further until the next nuclear fuel is brought into operation, and so on. The temperature rise in each case is brought about by the conversion of gravitational energy into thermal energy.

In this way, one set of reactions after another becomes important, the sequence always being accompanied by rising temperature. Since penetrations of Coulomb barriers occur more readily as the temperature rises it can be anticipated that the sequence will be one in which reactions take place between nuclei with greater and greater nuclear charges. As it becomes possible to penetrate larger and larger Coulomb barriers the nuclei will evolve towards configurations of increasing stability, so that the nuclei will become heavier until maximum stability is reached. The packing fraction curve shows that maximum stability is achieved at iron and nickel.

The temperature is not everywhere the same inside a star, so that the nuclear evolution becomes most advanced where the temperature is highest, at the centre. Nuclear evolution is least important near the surface and indeed may not have proceeded at all in the cool outer regions. It follows that a stellar explosion does not lead to the ejection of material of one definite composition, but instead a whole range of compositions may be expected.

2. THE SIMPLE TEMPERATURE SEQUENCE

The remainder of the present report is compiled from a recent article ([1]) by E.M. Burbidge, G.R. Burbidge, W.A. Fowler and F. Hoyle. An extensive abstract of this article has already been given by the same authors ([2]). The latter forms the basis of the following discussion.

As long as extremely high temperatures in excess of 5×10^9 dg.K. are not under consideration, the general tendency of nuclear reactions inside stars is to increase the average binding energy per nucleon. For a given temperature and density and for a given time scale of operation of the nuclear processes, the increase of binding that takes place is usually limited by Coulomb effects, but subject to this limitation the binding becomes as large as possible. That is to say, energy is degraded as fast as is consistent with Coulomb barrier effects, mitigated in some cases by resonance penetration. Since barrier effects become less severe as the temperature increases, it follows that the binding energies increase with temperature. This will become clear from the following examples.

At temperatures from about 10^7 to 5×10^7 dg.K. in main-sequence stars hydrogen is transformed to helium, with an average binding energy of 7.07 Mev per nucleon. It may be noted that the proton-proton sequence of reactions makes possible the production of helium starting only with hydrogen.

At temperatures from 10^8 to 2×10^8 dg.K. in giants and super-giants, He^4 is transformed principally to C^{12}, 0^{16}, and Ne^{20} with an average binding energy of 7.98 Mev per nucleon. The important roles of the ground state of Be^8 and of the second excited state of C^{12} in expediting the process of helium fusion, $3He^4 \rightarrow C^{12}$, have recently been clarified (3), and it is now clear that the longstanding difficulties in element synthesis at mass 5 and mass 8 are bypassed in this process. At temperatures of the order 10^9 dg.K., Mg^{24}, Si^{28}, S^{32}, A^{36}, and Ca^{40} are formed from the carbon, oxygen, and neon, the average binding thus rising to 8.55 Mev per nucleon, while at temperatures from 2×10^9 to 5×10^9 dg.K., Fe^{56} and neighbouring nuclei are synthesized, yielding an average binding energy of 8.79 Mev per nucleon. No higher binding than this exists, so that further simple heating of material will not synthesize in quantity elements of appreciably greater atomic weight.

The situation, then, is that a thermal « cooking » of pure hydrogen yields principally He^4 and the alpha particle nuclei with $A = 4n$, $Z = 2n$, $n = 3, 4, 5, 6, 7, 8, 9$ and 10 (C^{12} to Ca^{40}), together nuclei centred around Fe^{56}. These are the most abundant nuclei. Moreover the relative abundances that have been calculated for these nuclei, and particularly for the score or so of isotopes of titanium, vanadium, chromium, manganese, iron, cobalt, and nickel show good agreement with observed abundances. Typical results are indicated by the values in the following table for the case of chromium :

Isotope	Binding energy per nucleon (Mev)	Log. abundance relative to Cr^{52}	
		Calculated	Observed
Cr^{50}	8.706	— 1.89	— 1.27
Cr^{52}	8.776	0.00	0.00
Cr^{53}	8.760	— 0.85	— 0.94
Cr^{54}	8.778	— 1.78	— 1.50

Results such as these would appear to give strong support to the view that the elements under consideration were synthesized inside stars and that they became subsequently distributed in space, either by slow emission from late-type giants or by catastrophic explosion, as in supernovae.

3. FEATURES OF THE ABUNDANCE CURVE

The first attempt to construct an abundance curve was made by Goldschmidt (4). An improved curve was given by Brown (5), and more recently by Suess and Urey (6). These curves are derived mainly from terrestrial, meteoritic and solar date (7), and in some cases from other astronomical sources.

The simplest statement that can be made about the abundances is that when a mean curve is drawn through the points, the curve declines approximately exponentially from hydrogen to molybdenum and thereafter remains roughly constant. With abundances measured in numbers of atoms (not by mass) the decline from H to Mo is about 10^{10}.

It is worth noting that although from a nuclear point of view hydrogen may be considered the least stable of all the elements, it is nevertheless by a considerable margin the most abundant. This appears to be a strong reason for regarding hydrogen as the starting point of element synthesis.

The abundances show many important features that cannot be expressed in terms of a mean curve. These may be briefly mentioned :

(i) Lithium, beryllium, and boron fall far below the mean curve, their abundances being low by a factor of order 10^7.

(ii) The alpha-particle nuclei 0^{16}, Ne^{20}, Mg^{24}, ... Ca^{40}, Ti^{48} have appreciably higher abundances than their immediate neighbours (in general by a factor of order 10).

(iii) A strongly marked peak centred around Fe^{56} rises above the mean curve by a factor of about 10^4. The peak is narrow, ranging from atomic weight 50 to 62.

(iv) There are peaks among the heavier elements (atomic weights > 70), rising by a factor 10 above the general level. These peaks occur in pairs, viz at $A = 80, 90$; $A = 130, 138$; $A = 196, 208$.

(v) Proton rich heavy nuclei are extremely rare.

The abundances of the elements provide the main observational data to be explained by any theory of the origin of the elements. Abundances are available for nearly 300 nuclei, forming a rich body of information and setting a stringent test of theory.

Speaking very broadly the abundance curve shows the form to be expected if hydrogen represents the starting point of element synthesis, and if synthesis has not yet proceeded to any really extensive degree. We have seen that the general effect of thermal cooking is to convert hydrogen into nuclei centred around Fe^{56}, the conversion taking place by steps through intermediate elements such as carbon, oxygen, neon, ..., calcium. Thus the nuclei centred around Fe^{56} form a reservoir into which matter pours. As time goes on, more and more of the matter becomes degraded from hydrogen to iron, and the « iron peak » rises higher and higher above the mean curve.

These considerations do not, however, explain the detailed features of the abundance curve. The remainder of the present report will be given over to these detailed features [although for a really thorough going account the reader must be referred to the article already cited ([1])].

4. LITHIUM, BERYLLIUM AND BORON

It is a highly satisfactory feature of the observed abundances that lithium, beryllium, and boron are found to be exceedingly rare compared with neighbouring elements, for these particular light elements cannot be produced in any large quantity in stellar interiors. The possibility that they can be produced in hot spots on stellar surfaces has been considered ([8]).

Observations in radio astronomy and of cosmic rays make it abundantly clear that high energy particles are produced in solar flares. It is to be expected that similar but far more intense processes occur on magnetic stars, where all electromagnetic effects are likely to be very greatly enhanced over those that take place at the surface of an ordinary star like the Sun. Hence it does not seem unreasonable to suppose that on certain stars protons are accelerated in profusion, and that neutrons, alpha particles, lithium, beryllium, and boron are produced by spallation processes. The neutrons will diffuse from the hot spots into quiescient regions and will be primarily captured by hydrogen to form deuterium, with the emission of 3.23 Mev radiation — which may eventually turn out to be detectable.

5. THE TEMPERATURE SEQUENCE WHEN THERE IS CONTAMINATION

We have already seen that except possibly for « first » stars, stellar material does not consist entirely of pure hydrogen — the hydrogen is contaminated by small concentrations of heavier elements. In the Sun the total concentration by mass of elements heavier than hydrogen and helium is about 1 per cent. In older stars the concentration may be substantially less than this — concentrations as low as 0.01 percent having been established by observation. Yet although these concentrations are low, contamination by heavier elements turns out to be highly important when we come to consider the details of the abundance curve. Indeed the operation of the well-known carbon-nitrogen cycle depends on the presence of C^{12} as a contaminant. A similar neon-sodium cycle depends on the presence of Ne^{20}. In these cycles the nuclei C^{13}, N^{14}, N^{15}, O^{17}, Ne^{21}, Ne^{22}, and Na^{23} are produced as by-products during the conversion of hydrogen to helium. Eventually O^{18} and F^{19} are produced in helium reactions, so that all nuclei from carbon to sodium are accounted for.

The formation of Ne^{21} by the processes Ne^{20} (p, γ) Na^{21} (β) Ne^{21} is of special interest. At a later stage, after hydrogen exhaustion has occurred, free neutrons are generated by Ne^{21} (α, n) Mg^{24}. These neutrons are of great importance in building the elements beyond the iron peak. This question will now be considered in some detail.

The free neutrons are partly added to the light elements with $A = 4n$, producing the remaining isotopes of these elements, and partly added to Fe^{56} and other nuclei of the iron peak. Because Fe^{56} is present in low abundance compared to Neon, the number of neutrons made available per iron nucleus is of order 10 to 10^2, which is sufficient to build iron into the heaviest elements. The neutrons are produced in a medium that is mainly composed of alpha particles, which do not capture neutrons. If He^5 were stable, its production by neutron capture in He^4 would consume all available neutrons and heavy element synthesis would not be possible. This is worth emphasis, since in theories of primordial synthesis the break in the neutron chain at He^4 has been an insuperable stumbling block. In contrast, it is the saving factor in stellar neutron synthesis.

It is necessary to distinguish two conditions under which neutron capture can take place, a slow (s) process and a rapid (r) process. Suess and Urey (6) and Coryell (9) have already pointed out that the peaks of the abundance curves at stable nuclei with filled neutron shells ($A = 90$, $N = 50$; $A = 139$, $N = 82$; $A = 208$, $N = 126$) strongly indicate the operation of the s-process in element synthesis and that nearby peaks at $A = 82$, 130 and 194 require the operation of the r-process. The observed presence (1) of technetium in the atmospheres of S-stars may be taken as a demonstration that the building of heavy elements by a process of neutron addition does actually take place in the stars. Likewise the 55-day half life of light curves of supernovae of Type I would seem to demand the presence of Cf^{254} and this appears to provide a demonstration of the operation of the *r* process ([11]).

Burbidge *et al* ([1]) have shown that, with a small group of exceptions, it is possible to assign all the heavier nuclear species ($A > 62$) to either the *s* process or the *r* process. This separation allows the theory to be subject to quite stringent quantitative tests. Before coming to these tests it is worth interpolating a few remarks about the exceptional nuclei, which for the most part are proton rich isotopes of very low abundance. These nuclei have probably been formed in the following way.

Just as adulteration of the hydrogen of a star by elements up to Fe^{56} is to be expected, so adulteration of the hydrogen by the heavier elements that are built in the *r* and *s*-processes is to be expected. Such adulteration leads to the interesting possibility that on rare occasions, perhaps in supernovae, protons become added to the *r* and *s* products. The process is rare because a high temperature is necessary if protons are to penetrate the Coulomb barrier of nuclei of large charge. At a temperature of 2.5×10^9 dg.K. however, (p, γ) reactions occur in a time of order 10 seconds, even on the heaviest nuclei. It seems likely that proton-rich isotopes of the heavy elements were built in this fashion. These isotopes are characterized by the important property that they cannot be built directly by either the *r* or *s* process, and that their abundances bear an approximately constant ratio of 10^{-2} to those of neighbouring isotopes that are built by the *r* and *s* processes.

The effects of these various processes can be estimated quantitatively throughout the heavy element region, and it turns out that the abundance curve is explicable in terms of beta decay rates and of the cross-sections for neutron and proton capture processes. In the past such an analysis could not be made, for it was not realised that the abundance curve represents a superposition of contributions from several processes. Peaks due to neutron shell and quadrupole effects are not as pronounced in the over-all abundance curve as they are in the contributing curves.

It seems clear that both the *s* and the *r*-processes operated under conditions of « steady streaming ». For instance, in the case of the *s*-process the products obtained on multiplying abundances by the appropriate (n, γ) cross-sections turn out to be remarkably constant from one nucleus to another, as would be expected on the basis of steady streaming. There is one notable discrepancy, in the case of the element lead. The *s* process should terminate in cycling among the lead and bismuth isotopes at the onset of alpha-activity. If steady streaming has occurred there should be a consequent building up of the abundance of lead. Indeed the theoretical estimate turns out to be substantially higher than the abundance given by Suess and Urey (6), the factor of discrepancy being approximately 10 to 10^2.

An interesting outcome of the quantitative calculations for the *r*-process is that the original production ratio of U^{235} to U^{238} can be estimated to have a value close to 1.5. The present-day terrestrial ratio is approximately 1/138. A simple calculation shows that if terrestrial uranium was produced in just one star, probably in a supernova, then the event must have occurred nearly 7×10^9 years ago. An equally interesting calculation can be performed on a somewhat different basis. Thus we might assume that, instead of being produced by just one star, terrestrial uranium had the general average composition for the whole Galaxy *at the time the solar system was formed.* If further we take 5×10^9 years for the time that has elapsed since the formation of the solar condensation, and if uranium were produced at a uniform rate from the time of origin of the Galaxy, then the age of the Galaxy itself must be close to 9×10^9 years.

REFERENCES

(1) E.M. Burbidge, G.R. Burbidge, W.A. Fowler and F. Hoyle, *Rev. Mod. Phys.*, **29**, 547 (1957).
(2) F. Hoyle, W.A. Fowler, G.R. Burbidge and E.M. Burbidge, *Science*, **124**, 611 (1956).
(3) W.A. Fowler, *et al*, *Bull. Am. Phys. Soc. Ser.* 2, **1**, 191 (1956).
(4) V.M. Goldschimdt, *Skrifter Norske Videnskaps, Acad. Oslo I. Mat.- Naturv. Kl. No. 4* (1937).
(5) H. Brown, *Rev. Mod. Phys*, **21**, 625 (1949).
(6) H.E. Suess and H.C. Urey, *Rev. Mod. Phys.*, **28**, 53 (1956).
(7) Goldberg, Aller and Muller, preprint.
(8) W.A. Fowler, G.R. Burbidge and E.M. Burbidge, *Astro. J.*, **122**, 271 (1955) and *Astro. J. Suppl.*, **2**, 167 (1955).
(9) C.D. Coreyll, *Lab. for Nuclear Sci. Ann. Rept.* (1956).
(10) P.W. Merrill, *Science*, **115**, 484 (1952).
(11) G.R. Burbidge, E.M. Burbidge, R.F. Christy, W.A. Fowler and F. Hoyle, *Phys. Rev.*, **103**, 1145 (1956).

Discussion of Hoyle's Report

Dr. Oppenheimer. — What are the events which lead up to the supernova explosion ?

Dr. Hoyle. — At very high temperatures in excess of 5×10^9 deg. K., nuclear reactions become endothermic, not exothermic. This refers to a stage of evolution beyond those that I described in my paper.

If the central regions of a star rise to such a very high temperature, then gravitation is unable both to supply energy to maintain the hydrostatic pressure balance within the star, and also to make good the endothermic requirement. It seems that the hydrostatic balance becomes lost, and that the core of the star implodes.

Such an implosion leads to a rapid infall of the outer parts of the star, which undergo adiabatic compression. The resulting rise of temperature is adequate to trigger off nuclear fuels remaining in the outlying material, and hence to produce an explosion.

Dr. Oppenheimer. — Why does the 55-day period persist so long, after most of the Cf^{254} is exhausted ?

Dr. Hoyle. — It is certainly surprising that the Cf^{254} continues to dominate the light curve for more than about 400 days. Some slight upturn of the curve is to be expected. But we ought to ask Dr. Baade about that.

Dr. Baade. — I am afraid that systematic errors may have crept into the faintest part of the light curve of the SN in I.C. 4182 because the image of the supernova was imbedded in unresolved nebulosity of I.C. 4182. If I had realised at the time of the observations the importance of the lower part of the light curve I would certainly have taken special precautions.

Dr. Bondi. — Could the observed tail (after 400 days) be depressed by observation due to dust produced by the explosion itself ?

Dr. Baade. — Since there were slow changes in the intrinsic color of the supernova it would be very difficult to distinguish between obscuration effects and changes in intrinsic color.

Dr. Morgan. — How far would the enrichment of heavy elements be expected to proceed ? That is, is there a possibility that the heavy elements might have a considerably higher relative abundance in the most evolved systems than in our galaxy ?

Dr. Hoyle. — I think there is a distinct possibility of this being the case. So long as the heavy elements are being distributed into a large reservoir of gas their concentration is not likely to rise very greatly. But at the end of the evolution, when most of the gas is exhausted, the continued addition of heavy elements into the greatly depleted reservoir of gas could lead to rather high concentrations being generated.

Dr. Schatzman. — If supernovae of type I have all the same period of decay of 55 days, the observations of supernova would, according to Milford, permit us to find out if there is actual recession or not. If there is recession, there would be a red shift in light curves of supernova type I; if the red shift is due to some other phenomena, the period of decay would appear to be the same, whatever is the distance.

Dr. Baade. — The effect is too small to be actually observed.

Dr. Bondi. — Could you tell us little about the relative amounts of heavy elements due to the supernovae and red giant type of process ?

Dr. Hoyle. — The supernovae modify the composition of the gas to the extent of about 1/4 percent in the form of common metals, iron, etc. The red giants modify it, certainly to the extent of about 1 percent in the form of carbon and oxygen, and possibly to much more than this in the form of helium.

Dr. Baade. — I believe that only the supernova II should be considered in enriching the interstellar medium (and not the supernovae I which are old stars and may not have existed at the early time of star formation). Supernovae II are undoubtedly young stars since we have observed them only in the spiral arms of galaxies. Their numbers should have therefore been much higher in the earlier history of galaxies when the rate of star formation was much higher than it is today. Supernovae II should be excellent " spreaders " of metals because their velocities of expansion are of the order of 7 000 km/sec and the ejected masses are certainly much larger than those of the supernovae I.

Dr. Gorter. —

1. What would be the ratio of the neutron densities required for the fast and the slow road for the synthesis of heavy nuclei ?

2. Would a middle road be physically possible and are there indications that it has been used by nature ?

Dr. Hoyle. — The neutron flux is greater in the fast case by a factor of order 10^{10}. The time scale for the slow road is about 10^2 seconds. There is some indication of an intermediate road on a time scale of 10^2 to 10^3 years. From an astrophysical point of view, this may possibly be associated with supergiants rather than with ordinary giants.

Dr. Oppenheimer. — Can one understand in terms of nuclear content the differences between supernovae of type I and type II ?

Dr. Baade. — We know only that supernovae I are old (population II) stars while supernovae II must be young (population I) stars.

Dr. Heckmann. — As the discrepancy between the steady-state theory and the nonstationary models has always been in our minds during the discussions, I would be glad if Hoyle could tell us which of the processes he outlined are necessarily connected with the steady-state theory and which are not.

Dr. Hoyle. — These considerations are consistent with both types of cosmology, provided any superdense state of matter that may occur in non-stationary cosmology satisfies the requirement that matter emerges from the superdense state essentially as hydrogen.

Dr. Ledoux. —

1. I would like to make a few comments on questions of stability connected with supernovae explosions. If we have a true nuclear statistical equilibrium in the core, one could define a generalized value of the effective ratio of specific heats in that region. This would facilitate greatly a quantitative discussion since we know that, to have dynamical instability, the average value of Γ over the whole star must be smaller than 4/3. Of course, in the core, Γ will take a very small value close to unity but, still, the region where the iron group has reached a state of statistical equilibrium has to represent an appreciable fraction of the total mass to make $\overline{\Gamma} < 4/3$. Yet, when there is such a core, I will admit that it is very likely that dynamical instability could occur.

2. On the other hand, in the earliest stages of element formation (thermonuclear transformations of H, He, C, etc.), I fail to see how instability of the core of the star could arise. And still, this seems necessary to spread the material of the core which has been appreciably affected by nuclear reactions into interstellar space. This instability is also necessary to produce a temperature rise in the external layers capable of starting *violent* hydrogen reactions. It seems somewhat dubious to me that the observed regular shedding of material from the external layers of giants is real evidence for such deep seated instability.

3. What is the order of magnitude of the effective temperature exponent in the law of energy generation in the intermediate stage when C, O, Ne, etc., are reacting to built up Mg, Si, Ca, etc. ?

Dr. Hoyle. — There is no suggestion that the observed regular shedding from giants should be associated with any catastrophic instability of the star. This was only invoked for the supernovae.

The temperature exponents are probably in the region of 20.

Dr. Ambartsumian. — Dr. Baade has told us that supernovae II appear usually among the population I. Can he say anything more exact about the positions of supernovae II in the spiral arms ? Are they exploding in the concentrations (associations) which we observe in spiral arms, of do they explode also in the gaps between the associations ?

If we could have some answer on this question, the age of exploding stars could be estimated with more certainty.

Dr. Hoyle. — I think this is another question, that I prefer to pass on to Dr. Baade...

Dr. Baade. — I would like to look up my plates in order to answer Ambartzumian's question.

Dr. Oort. — Since we must assume that the production of the bulk of the heavy elements in the Galactic System has taken place in the early stage of its evolution an estimate of the total amount produced on the basis of the frequency of supernovae at the present time must be considered as doubtful. Certainly we need not worry about factors of 3 or 10 between the predicted and observed abundances of heavy elements.

Dr. Hoyle. — I would certainly agree with Dr. Oort's comment.

Dr. Shapley. — Dr. Baade, what proportion of the supernovae so far known are of type II ?

Dr. Baade. — Since the absolute photographic magnitude of the supernovae II at maximum is about 2 magnitudes fainter than that of the supernovae I, the number of supernovae II found in the first M. Wilson-Palomar search for supernovae was multiplied with the proper volume factor to make them comparable with the observed number of supernovae I.

Dr. van de Hulst. — Could Dr. Hoyle tell us about the amount of deuterium to be expected in the interstellar gas ? Would this be formed by fragmentation of heavier nuclei in the cosmic radiation or are there other processes ? The answer would have considerable interest also in connection with the possibility of observing the radio line of deuterium at 93 cm.

Dr. Hoyle. — Deuterium is probably mainly formed on stellar surfaces, by neutron capture by protons. The rate of formation of deuterium would therefore seem to be closely connected with electromagnetic processes occurring on stellar surfaces (neutrons being released in such processes).

It appears that such processes may be quite common, even in stars like the Sun. But even so it is difficult to build sufficient deuterium to give an universal abundance as high as that which is found in the Earth.

Dr. Mac Crea. — I should like to ask about nuclear abundances in cosmic rays and whether these abundances can be related to the production of nuclei in stars ?

Dr. Hoyle. — I think that the nuclear abundances found in cosmic rays will depend most crucially on the injection process. There is no reason why this process should accelerate all nuclei equally in proportion to their cosmic abundance. It is quite possible that heavy nuclei are preferentially accelerated.

Dr Oppenheimer. — Nuclei like Li, Be, relatively abundant in primary cosmic rays, would be secondaries in interstellar collisions.

Dr. Fierz. — May I ask, if Aten in his guess of the age of elements uses any considerations different from those Hoyle used ?

Dr. Hoyle. — The considerations are exactly the same in principle, except that his uranium isotope ratio was not calculated by allowing for the different primaeval abundances of the various progenitors of these isotopes.

Dr. Oppenheimer. — May I make a final comment ?

Dr. Heckmann expressed misgivings about the arbitrary alterations of the basic equations of relativity involved in the " steady-state theory ". In these I concur. But, by providing an incentive for understanding the present state of the cosmos in terms of processes that can now be in progress, this theory has led to the beautiful work reported yesterday by Hoyle on element synthesis. Even if the hypothesis is, as I believe, quite wrong, it has thus led to great progress in our understanding.

Some Spectroscopic Phenomena associated with the Stellar Population of Galaxies

by **W. W. MORGAN**

The integrated spectral types in the violet region of the bright, inner parts of galaxies range from classes B to K. At one extreme (B type), occur certain highly irregular systems having a generally chaotic appearance ; at the other (K type), are systems showing pronounced nuclear concentration of luminosity in an amorphous bulge and symmetry in general appearance. Systems having spectral types intermediate between these two extremes possess in general intermediate characteristics with regard to form; however, in what follows we shall consider principally two categories in spectral type located near the two extremes described above. We shall be concerned then, with the following categories : (1) Systems having violet spectral types of classes B or A (certain classes of irregulars, and spirals having minor nuclear concentration of luminosity); and (2) systems whose inner parts have violet spectral types of late G or K (spirals having a high degree of nuclear concentration of luminosity, giant ellipticals, some SO systems).

The above two categories are cleanly separated — both with regard to form and to spectral type. If the form is used as the principal criterion, a certain scatter in the correlation with spectral type is observed; that is, some systems in the first *form* category may have violet spectral types as late as F; systems in the second form category may have violet types as early as the neighborhood of GO; however, even with this amount of scatter, there is probably no actual overlapping as far as the violet spectral type is concerned.

This implies that there must be a systematic difference in the kinds of stars contributing mostly to the violet light of systems in the two categories. Symbolic HR diagrams for these two categories of galaxies have been published by Morgan and Mayall. The spectra of the B and A galaxies are strongly suggestive of the expected appearance of integrated spectra of certain open clusters in our galaxy. For instance, the spectra of such B-type galaxies as NGC 3395, 3396, 3991 and 3995 resemble rather closely the integrated spectrum of the inner part of the Orion Nebula cluster and other clusterings of extremely early-type stars in our galaxy — both with regard to the absorption spectrum and to the emission lines.

Systems similar to the above four are to be considered as extreme examples with regard to their stellar population; however, they are not particularly rare objects in the nearer parts of the universe; and it seems quite possible that the majority of the exceedingly blue galaxies reported by Haro may belong to this category. In the case of the B group, therefore, there are strong indications that we may be observing conditions on very large scale which are similar to those existing in the restricted regions of O-associations in the spiral arms of our galaxy; we may describe this as a « young-star rich » population.

Also included in the first category are the systems whose violet spectra contain hydrogen absorption lines of outstanding strength — similar to those observed in main-sequence B8 — A5 stars. A number of irregular galaxies are in this category — for example, NGC 4490 and 4631 — together with some spirals having insignificant central concentration of light; in general, the last-named tend to have systematically slightly later spectral types than the irregulars above mentioned. For this group of A systems, the integrated violet spectra approximate the appearance of the integrated spectra of open clusters such as the α Persei cluster and the Pleiades; in these clusters, main-sequence stars are observed to a bright limit of B3 and B6, respectively.

In the second category (ellipticals similar to NGC 4472, the inner parts of spiral systems similar to NGC 224, inner parts of barred spirals similar to NGC 1398, and some SO systems), the spectral type in the violet region is vastly different, and resembles closely the spectrum of giant K stars and the integrated spectrum of the « late-type » globular cluster NGC 6356. The absence of sensible traces of stars of spectral types earlier than G suggests that a fairly close qualitative resemblance exists between the HR diagrams of the inner parts of these galaxies and those of the « evolved » galactic cluster M67. The overwhelming relative numbers of B and A stars in systems of the first catagory are absent from those of the second; even in the favorably situated violet spectral region, little or no trace of them is found; from analogy with the star clusters listed, we may label galaxies of the second category as « young-star deficient ».

Now the significance of the categories « young-star rich » and « young-star deficient » must be examined further. In the « young-star rich » systems we consider that the formation of massive stars is proceeding at the present time on a very large scale; in the case of the « young-star deficient » category, the large-scale formation of massive stars seems to have ceased at a remote period in the past. One critical difference, therefore, between the two categories is in the present rate of star formation.

The above by itself does not demonstrate that the « young-star rich » systems are less evolved than the « young-star poor » systems; to investigate this point, the problem of the relative abundances of stars in the secondary stages of evolution must be examined. That is, is there evidence for the existence of sufficient numbers of evolved yellow giants to support the assumption that star formation in the « young-star rich » systems has been proceeding at its present rate for billions of years?

The problem now arises of how the existence of these two extreme categories of galaxies is to be interpreted. On the one hand, we have to consider the possibility that « young-star rich » systems evolve into « young-star deficient » systems; that is, the rate of star formation decreases with time. In this case we should consider the possibility that the spectral sequence of galaxies (B $\rightarrow$ A $\rightarrow$ F $\rightarrow$ G $\rightarrow$ K) is also an evolutionary sequence. On the other hand, we must also consider the alternative possibility that the spectral sequence of galaxies is actually composed of end points of various differing evolutionary tracks. In the latter case, we would interpret the population differences between the first and second categories as resulting from differing initial conditions pertaining to the early stages in the formation of galaxies. In the former case, B and A type galaxies might be considered actually younger than those of type K;

in the latter, they might all be considered to be of approximately the same age.

There are certain interesting features in the space distribution of the « young-star rich » and the « young-star deficient » systems. In the region of the universe defined approximately by the galaxies in the Shapley-Ames Catalogue, large-scale regional fluctuations in the relative numbers of members of the two categories are found. For example, large relative numbers of « young-star rich » systems are found in the range $10^h - 13^h$ right ascension and between declinations $+30^{\circ}$ and $+60^{\circ}$. These galaxies as a group have low red-shift velocities and must be located in a neighboring region of space to that occupied by the Virgo Cloud — in fact, it seems possible that the two regions may form a single physical complex. The average stellar population characteristics differ sharply, however, between the Virgo Cloud and the concentration of galaxies to the north and west : The inner part of the Virgo Cloud contains a high percentage of « young-star deficient » galaxies, while the extension to the north shows a much higher percentage of « young-star rich » systems.

A large amount of the spectroscopic material discussed here was obtained by N. U. Mayall. I am greatly indebted to him for its use.

Discussion of Morgan's Report

Dr. Oort. — In connection with the differences in distribution of young-star poor and young-star rich clusters which Dr Morgan indicated, I should like to point out that it seems quite possible that the « turbulent » motions in the primordial medium differed for the central point and the outer regions of the Virgo cloud. Such differences might well have caused a preference for young-star poor galaxies in the central parts.

Dr. Baade. — I am quite certain that the « blue-star rich » systems of Morgan are Sc-spirals (with a sprinkling of Irr. systems) and the « blue-star deficient » systems, elliptical galaxies (with a sprinkling of Sa and Sb-spirals).

These correlations should stand out clearly if the spirals are classified according to the sizes of the central spherical systems (bulges).

Dr. Sandage. — There is one observational datum that the mean density of matter in galaxies is different between Sc-galaxies and E-galaxies. The density of matter in the E-galaxies is believed to be a factor of 20 higher than in Sc.

If the rate of star formation is some power of the gas density, then this rate must have been much higher in the past, in systems with high initial gas density. These are the systems which are now classed as E- or « old-star systems ». Because the *initial* rate of star formation was high in E-galaxies, they have now used up all of their available gas and no stars are now being formed. But in Sc-galaxies, the initial density was low, therefore dN/dt is small, and stars are still being formed today.

Dr. Ambarsumian. — In connection with the interesting communication by Morgan, I should like to make two comments :

1. We cannot identify the ages of « young-star rich » galaxies with the ages of blue stars and O-associations in them. It is very probable that each such galaxy has had a number of successive generations of hot stars and O-associations.

2. Before concluding that the star formation in « young-star deficient » galaxies of Morgan is finished, we shall have to have definite observational evidence that between the population II stars we have no young stars at all.

Dr Morgan. — 1. If a large number of successive stages of O-associations had occurred in « young-star rich » systems, we would expect to observe a bright stratum of « evolved » stars from these earlier epochs. Such bright strata do not seem to be present.

2. As far as my knowledge goes, there are no combinations of Population II stars with gas or dust in our galaxy. The general conclusion is usually drawn that Population II stars are very old.

Dr Baade. — The Magellanic Clouds undoubtedly belong to Morgan's « blue-star rich » systems. We known today that they contain also Population II objects (old stars). This indicates that star formation started in them at the same time as in the elliptical galaxies.

Dr Mac Crea. — Are the « young-star deficient » system also deficient in interstellar gas?

Dr Morgan. — Yes, usually — as far as our present information goes.

IV. GENERAL DISCUSSION

Dr. Bondi. — In our meetings we have first discussed the theories and then the observations. I think it would be of advantage if we now discussed the best ways of drawing theory and observation together. What are the observational tests that appear feasible in the immediate future and that would be recognized as decisive by the theorists? What are the consequences of the theories that deserve further examination with a view to subjecting them to observational tests?

It appears to me that the chief decision is as to whether we live in an evolutionary or in a steady state universe. Are there any tests that can take the place of the now abandoned Stebbins-Whitford effect? How many other crucial features are there? We have heard of the decisive importance of the random velocities of the galaxies. Is there any hope of measuring them?

To start the discussion, may I ask the observers whether they might be able to observe a variation of the ratio of the frequency of different types of galaxy with distance, or whether we must expect that selection effects would hide any such variation?

Dr. Sandage. — Concerning the possibility of finding a change in percentage of Sc to E galaxies as one goes back in time (out in space), it should be said that we cannot classify consistently the galaxy type to distances larger than about one billion years (if the Hubble constant is 75 kmsec per 10^6 psc). This time is probably too short to see evolutionary trends in percentages of Sc to E nebulae.

Dr. Baade. — I would like to state more bluntly today what was implied already in my report. The observational data are in favour of the evolutionary picture and not in favour of the steady state picture.

In my report, I pointed out the intimate relationship between the type of a galaxy and its physical content (stellar populations and gas). I also pointed out that we have convincing evidence that in all nearby galaxies (the members of the local group) star formation begun at about the same time, some 6×10^9 years ago. Together with recent data about the amounts of gas still present in the different types of galaxies this led to the conclusion that star formation started in all galaxies at about the same time but that its rate was different for the different nebular types. This is what has been termed the evolutionary picture of the universe. All galaxies are of essentially the same age.

On the steady state picture on the other hand, we should find interspersed among the galaxies both very old and very young systems. The question is : would we be able to recognize such systems?

Let us take as an example the E-galaxies because they are the best defined group. In their youth and at the height of star formation the present E-galaxies must have been truly magnificent stellar systems with their wealth of supergiant stars and bright H II regions. It is of course no argument against the steady state theory that we do not observe such youthful E-galaxies in our nearer neighbourhood. But even at very large distances such E-galaxies could not escape being noticed since exceptionally bright supergiants, O and B star associations, H II regions and globular clusters studded with supergiants would produce fluctuations in the general intensity distribution (a speckled appearance) which would draw immediately the attention of the observer. For a number of years now I have been looking for E-galaxies of this description on the best of my plates taken at the 100 inch and 200 inch telescopes but I have found none.

Dr. Sandage. — I should again like to mention an observational point which may have bearing on the evolution of galaxies. It appears to be true that the density of matter is higher in E-galaxies than in Sc-galaxies. This separation into galaxy types on the basis of mass density is probably *not* due to chance, but rather it would papear that the galaxy type is what it is today because of the original density of mass.

If the rate of star formation depends upon the mass density, then those galaxies with high initial density will have run their course in star formation, using up their initial supply of gas, and are now incapable of forming new stars. Although the E-galaxies and the Sc-galaxies may be (and in my opinion are) the same number of years old, they are in different stages of their evolution.

Because the mass density is low in galaxies of type Sc, these galaxies are just beginning their evolutionary history, whereas the E-galaxies have nearly finished their production of stars.

Mr. Hoyle. — Since Baade has mentioned his evidence against the steady-state theory, I would remark that in the latter theory one may have to say that newly condensed galaxies are never recognisable as elliptical galaxies, but as other types.

If one says that the spirals have ages of order 4×10^9 years, Sb and Sa have ages of order 8×10^9 years, and E-galaxies of order 1.2×10^{10} years, then the expansion of the universe supplies an explanation of the relative frequency of these types, in particular of the fact that E-galaxies seem less frequent than spirals.

Mr. Baade. — But we have convincing evidence that in the local group (comprising E-galaxies, Sb and Sc spirals and Magellanic cloud type galaxies) star formation started at the same time.

Mr. Gold. — I do not think that we have a sufficiently good understanding of the process of formation of a galaxy to know what the formation of an E-galaxy must be like. We are not even sure that the gas to make all the stars must have been there in the first place — some may be acquired during the evolution and thereby the star density may be caused to grow gradually. Without a real theory that accounts for the formation of galaxies, I do not think that this point can be settled.

Mr Baade. — I think we have good reason to believe — because the Universe is still so young — that the E-galaxies in their present form reflect to a high degree the forms of the E-galaxies in the gaseous state before star formation took place.

Mr. Hoyle. — The maximum dimensions of irregularities in the steady-state theory are of order 50 mps. Within a region of this dimension correlations could occur. I would like to ask Dr. Baade whetner his survey extended to a greater distance than this.

Mr. Baade. — Yes, indeed!

Mr. Gold. — The absence of heavy elements and thus of dust may drastically change the evolution processes.

Mr. Mac Crea. — In addition to the spatial considerations mentioned by Hoyle, there is also the temporal one. If a galaxy has to be less than 10^6 to 10^7 years old in order to give the characteristics sought by Baade, since the average age of all galaxies is probably about 10^{10} years, we should expect only in 10^3 or 10^4 to show these characteristics. Combined with Hoyle's considerations, the probability of finding a very young system may therefore be very small indeed.

Mr. Bondi. — I feel very strongly (I have said so plainly elsewhere) that it is of little use inferring backwards from observation. The sound procedure is to use observation to check a theory. When we have a theory of evolution of galaxies, then it will become possible to check it (and the steady-state theory with it) against the observations Baade has mentioned, but until we have such a theory, their interpretation is very much in doubt.

Mr. van de Hulst. — The agreement between quantities of the order of 10^{39} has been mentioned in two reports, those of Drs. Hoyle and Klein. If I understand Hoyle's report correctly, the choice is between :

1. The agreement is accidental;
2. The steady-state theory;
3. The relativity theory combined with a gradual change of the natural constants.

Could any of the physicists here make further comments on this problem?

Mr. Oppenheimer. — Klein gave us a very simple suggestion as to how the ratio of electrical to gravitational forces should be related to the number of particles in the observed regions of the Universe. This seems at least an interesting alternative.

Mr. Ambartsumian. — Let us suppose that the population II stars which we observe now have evolved from population I stars. Then the suggestion of Dr. Baade that giant ellipticals should have exceedingly high surface brightnesses at the initial stages of their life seems at first unavoidable.

But even remaining in the frame of a gas-star scheme, we have some possibilities to avoid this difficulty.

As Prof. Kukarkin suggests, we observe in our galaxy young groups of stars of different kinds. Between them we have pure T associations, which contain no blue giants.

Perhaps the distribution of luminosities of newly formed stars depends on the initial conditions. It may be that in central parts of giant systems, the stars of lower masses are more often formed.

But if we adopt the point of view described in my paper yesterday, then we may think about the existence in population II of young stars of a special kind (different from population I stars). In this case, we may suppose that the formation of stars in giant ellipticals continues at the present time.

Thus my suggestion is that it is possible to overcome the difficulty Dr. Baade has mentioned.

Mr. Gold. — 21 cm Radio Astronomy holds out the promise with new and more sensitive receivers to obtain a curve of neutral hydrogen amounts against velocity, reaching out to perhaps 50 % of the speed of light. Such a curve obtained for clusters of galaxies would be a straightforward cosmological item of information.

Mr. van de Hulst. — For the immediate future my hopes about the masers are not as strong as those of Dr. Gold. In my report I estimated a gain of a factor 3 only within the next 3 years.

Mr. Lovell. — Although I agree entirely with Gold's remark about the desirability of extending the range of measurements of neutral hydrogen and of the 21 cm line shifts, I think it would be a mistake to underestimate the difficulties of these things in the next five years. It is important that cosmologists should consider further the kind of tests of cosmological nature which are possible now : for example, I think that Hoyle's suggestion made during this conference about the measurement of angular diameters may be very important. These measurements of angular diameters may well be decisive. Therefore further theoretical work on Hoyle's suggestion would be very desirable.

Mr. Bondi. — It might be of advantage to spend time on getting a distance scale by such work as Lilley's, rather than on counting and defining sources.

Mr. Lovell. — I agree, and also it is important to search for a few more identifications of objects apart from normal nebulae. At present the whole of the cosmological interpretation of the radio sources is based on a half dozen identifications of the Cygnus type.

Mr. van de Hulst. — Also the intensity of even the nearest sources at 20-30 cm is extremely small. For instance, in the survey of Westerhout, reported yesterday, the bright source NGC 1275 was only barely detectable.

Mr. Bondi. — Can Lilley's work be extended so as to measure the displacement of the 21 cm line for a few more radio sources?

Mr. Lovell. — There is a very good hope of doing this in the near future, but the technical problems are considerable because of the large and unknown frequency shifts which are involved.

Mr. Gold. — The discovery of any sort of recognisable spectral feature at long wave-lengths would be very important, so that one can place sources approximately in depth. Perhaps a search ought to be directed accordingly.

Mr. Sandage. — The question of an observational determination of the mean random motion of galaxies outside the tight clusters or groups is very difficult. The only way which one can now think to find this answer is to obtain a distance independent of the red shift and then to find deviations of the observed red shifts from the systematic red shift. As of the present moment we do not know with sufficient accuracy how to determine distances to remote galaxies where significant red shifts occur.

Dr. Oort. — Bondi has stressed the importance of giving utmost attention to observations that will yield new information on the major problems discussed. We will all be in full agreement on this; radio-astronomers in particular have been and are very much aware of the importance of making observations that bear on these problems. It is all a question of what is *possible* observationally. Gold has mentioned hydrogen-line measures on clusters of galaxies. It is, however, still uncertain whether there is a measurable amount of interstellar hydrogen in these clusters. The observations that seem most promising at present are counts of weaker radio sources, if possible combined with estimates of diameters. If we could identify by radio measurements sources of the type of Cygnus A (or Hydra A) even when they are faint, it would already now be possible to observe these at distances corresponding to recessional velocities close to the velocity of light. It seems likely that such observations would give important new information on the universe.

I should like to see predictions for various models of the universe of numbers of such radio sources as a function of apparent brightness down to very faint objects.

Dr. Hoyle. — The curves that I showed can be interpreted for radio counts. In spite of my use of bolometric apparent magnitudes, the same curves can be used provided the energy spectrum is of a given type, and provided the observations are carried out over a fixed frequency range.

Dr. van de Hulst. — Any smooth cosmological theory clearly cannot make a useful prediction because the 21 cm line will be drawn out to a continuous spectrum. It all depends, therefore, on the irregularities about which it would seem hard to make a theoretical prediction.

CONTENTS

(Every Report was followed by a Discussion).