The Time Scale of Our Universe

By E. J. Öpik

Armagh Observatory
Armagh, Northern Ireland

INTRODUCTION AND HISTORY

The average scientist of half a century ago did not ponder much the question of the beginning and age of the universe. For lack of observational approach this problem remained outside the realm of exact science. It was generally felt that the universe should have neither beginning nor end, a viewpoint which was more influenced by opposition to former mythological or religious ideas of creation than by impartial reasoning.

Indeed, the second law of thermodynamics was well established at that time. According to this law, the universe is steadily running down toward equalization of the energy content of its parts. The ultimate state is that of universally constant temperature, "Wärmetod" or thermal death, where, in the absence of temperature differences, no exchange of energy, no relative motion except that of molecules could take place. Organic life, the metabolism of which consists in exchange of energy, could not exist then, even were the temperature favorable for life—which, in all probability, it would not be. The mere fact that temperature differences exist, that suns shine and planets carry life in the face of the immensities of cold space (into which heat energy is lost in the form of radiation), would point to the youthfulness of our world, to a beginning a finite interval of time ago.

Scientists of the beginning of this century preferred to ignore this writing on the wall. There were some reasons or, rather, pretexts which seemed to justify this eluding of the fundamental problem. The second law of thermodynamics, or that of increasing "entropy," determines only the direction, not the speed with which equalization is approached. The speed, depending on a number of unknown

processes, being itself unknown, no definite calculations of the time intervals involved could be made.

When going back in time, the second law leads to ever-increasing energy concentrations in the past; an unlimited past would lead to infinite energy concentrations, a concept which is physically unacceptable. Certainly some uneasiness was felt in this respect by those who did not want to draw the logical conclusion of a finite age for the present universe. But, then, there was Maxwell's demon, an imaginary intelligent being who could, at will, regulate molecular processes and thus do away with the law of entropy. This sufficed to show that the law is not absolute. The law is only of a statistical nature, exceptions being always possible although more or less improbable. Further, its validity for unlimited intervals of space and time was questioned. A perhaps not very justifiable complacency about the beginnings and ends of the world was thus sustained.

During the second quarter of this century a great change in the scientific outlook in this respect took place. The recession of the extragalactic nebulae, coupled with the finite age of the radioactive elements, suggested that there was a beginning a few thousand million years ago, the same for the galaxies and for the atomic material of which our planet is built. Following the above-mentioned phenomena back in time, moments could be reached beyond which the recession of nebulae and the decay of radioactive isotopes could not continue in the same manner as they do now. The two time limits were not found to be equal although they were of the same order of magnitude; but, within the uncertainties of theory and observation, they could be adjusted to each other. The idea of a finite age for the universe emerged. A stage, some 3,000 million years ago, was visualized at which the universe was closely packed together, when the temperature and density were high enough to invert the radioactive processes and to cause the building up of the heavy unstable isotopes at a rate equal to, or faster than, their total rate of decay (spontaneous + induced) in these conditions.

One view considered this stage merely the remotest phase of evolution of our world, beyond which extrapolation from the present state is not possible. It was not meant to be necessarily an absolute beginning—more likely it was not. The concept of age is thus reduced to that of a time scale, or a time interval during which the properties of the universe have radically changed. This definition appears to be somewhat vague; but it would imply nothing short of a complete absence, at the early stage, of all the classes of celestial bodies which are familiar to us now. In such a form the definition is stringent enough. Therefore, even if we could assign an upper limit of age to all existing stars, this would be only a subordinate time scale—that of stellar evo-
lution—unless we could prove the total absence of any stars before that date, and not only of those existing at present.

A more drastic view preferred the concept of an absolute beginning, perhaps identifiable with an act of creation. The definition of the time scale remained the same as before, but additional meaning was attached to it as that of the absolute age of the universe. The initial stage, a singularity from which the universe started expanding, was the limit of extrapolation not only from the present, but from any state of the universe, however close to the initial stage.

The difference between the two viewpoints is a matter of principle, and not of how the initial state of the universe is pictured. Although Eddington's primeval nebula, assumed to have preceded the present expanding state of the universe, could have existed indefinitely, it could equally well have been the first created object, called into being in a peculiar state of almost exact equilibrium between gravitational attraction and the hypothetical force of cosmic repulsion. On the other hand, Lemaître's primeval atom, "the egg from which the universe hatched," is most simply interpreted as the result of an act of creation; yet it could also have been the final outcome of collapse of a previous universe, oscillating indefinitely in alternating expansion and contraction. The choice between the two, continuous existence or creation, will remain a matter for esthetic judgment, not for positive science defined as theory verified by observation.

There is no proof in purely esthetic matters. This does not mean that esthetic methods of approach to scientific problems are worthless. On the contrary, scientific theories are created by intuition, or by an essentially esthetic process. However, without the flesh and bone of experiment such theories remain mere shadows of possibilities.

To remain on solid ground, in the following we will pay little attention to esthetic considerations, however important these might appear from the standpoint of philosophy or religion. We will, further, be guided by the principle of minimum hypothesis, or economy of thought, which requires that new laws of nature must not be used for the explanation of phenomena which can be accounted for by known laws. This is a safeguard against becoming lost in the blind alleys of guesswork. The chances are small that a theory not supported by facts would prove to be correct.

As already mentioned, the fundamental fact requiring a short time scale was, and remains, the red shift of the extragalactic nebulae. With the existing laws of nature this phenomenon is explained in the most straightforward way as recession. There are yet no facts known which would contradict this explanation. The Hubble-Humason law (1928) of proportionality of the red shift, or velocity of recession to distance, led to a time scale for the universe equal to a few thousand
million years. The uncertainty of the estimate depended upon how the rate of expansion of the universe, as revealed by recession, was assumed to vary with time. Nevertheless, various models of the expanding universe, based on different assumptions (de Sitter, Einstein, Friedmann, Lemaître, Eddington), gave figures within the same order of magnitude for the time of rapid change.

The new “short” time scale of some 3,000 million years was like a bombshell amidst the complacent “permanentists.” At that time the pundits as well as the rank and file accepted a thousand times longer time scale for the stellar content of our galaxy alone, no mention being made of the universe as a whole. This “long” time scale, a multiple of a million million years, was mainly the outcome of mathematical investigations by Jeans into the statistics of stellar motions and the distribution of the orbits of double stars. Jeans assumed a state of “statistical equilibrium,” or that the present motions of the stars are essentially influenced by their mutual gravitational action in past close encounters. A close encounter would mean the passage of another star through our planetary system at a distance—say, between Mercury and Neptune. Such a passage would leave both suns physically intact (although its effect on planetary orbits might be disastrous), yet their motions would be changed in much the same manner as those of two gas molecules after an elastic collision. Jeans actually applied the kinetic theory of gases to the stellar universe. Because of the great distances separating individual stars, close encounters can happen only about once in several million million years, a figure which can be arrived at by elementary calculation if the average velocity and distance between the stars is known. The “long” time scale was thus not a result of Jeans’ elaborate mathematical theories, which were undoubtedly correct, but follow merely from his basic assumption of statistical equilibrium, implying that each star during its lifetime had a fair chance of undergoing several close encounters with other stars. In trying to prove his basic assumption, Jeans selected only certain statistical data which, superficially, seemed to agree with it, and, strangely enough, overlooked numerous more important criteria which contradicted his assumption. Thus, while carefully considering the effects of encounters on close binary stars, he disregarded the wide pairs and star clusters upon which the effects, according to his own theory, should have been thousands of times stronger. Indeed, with the long time scale these objects should have ceased to exist long ago, in contradiction to observation which reveals numberless wide double stars and loosely bound clusters in the sky. The evidence against statistical, “gas-kinetic” equilibrium is overwhelming, and there is no foundation whatever for the “long” time scale in our stellar universe. The battle of “short” versus “long”
time scale is definitely won by the former, although the latter did not yield without a struggle.

In the course of the controversy, arguments based on subordinate time scales were produced. These subordinate time scales—of the earth, the radioactive elements, stellar evolution, stability of binaries and star clusters—all fell below a not too large multiple of one thousand million years; as, moreover, some were obvious overestimates, they were considered as supporting the short time scale of the universe itself. An early, apparently the first, synoptic account of the evidence relating to the age of the universe concludes as follows (1) 2: “the combined evidence presented by meteorites, by statistical data relating to wide double stars, by the distribution of stellar luminosities in globular clusters . . . , and by the observed recession of spiral nebulae . . . points to an age of the stellar universe of the same order of magnitude as the currently accepted age of the solar system: not much more than 3,000 million years.” In this account stress was laid on radioactive age determinations of meteorites by Paneth (whose results were later greatly changed), and on the abundance of lead isotopes in the earth’s crust as testifying to the age of the radioactive elements (results which have been corroborated since). Subsequent synoptic reviews invariably arrived at practically similar conclusions, formulated sometimes more, sometimes less cautiously, although, with the changing aspect of our knowledge and different personal approach, the emphasis was on different phenomena: radioactivity and the age of the earth, and stellar evolution with a hydrogen-helium source of energy (2); galactic dynamics (3); the stability of star clusters and binaries (4, 5); the red shift of nebulae and the radioactive age of the earth’s crust (6).

The survival of the idea of the short time scale over two decades of intense astronomical and physical research is in itself a measure of its worth; it serves now as a generally accepted working basis in widely different fields of study.

In the following an attempt is made to draw an up-to-date balance for the problem of the time scale or age of the universe.

THE AGE OF THE EARTH

The continental shields of northeastern Europe, Canada, South and Central Africa, and others, where mountain building ceased at an early age of our planet’s history, represent the oldest undisturbed portions of the earth’s crust. All these regions are lowlands or plateaus devoid of mountain chains, and are free from earthquakes which are the sign of continuing upheavals. The age of these old formations

---

1 Numbers in parentheses are references to the literature cited at end of text.
is expected to come nearest to that of the earth’s crust or the earth itself.

The most suitable method of age determination of rocks consists in a comparison of the abundance of radioactive isotopes, such as those of uranium, with the abundance of their end products, e. g., lead isotopes. Pure minerals in the form of crystals are chosen for samples; there must be a guaranty that no exchange of substance between the sample and its surroundings has taken place, in which respect only individual crystals can be considered as reliable. Knowing the rate of decay of the radioactive substance, the time during which decay has been going on can be calculated from the amount of end product accumulated. The determinations are accurate to within 8 to 10 percent, much more accurate than the other astrophysical age estimates referred to below.

Radioactive age determinations yielded, indeed, high values for some mineral samples from the shields. Pegmatites from northern Karelia, in the so-called Baltic shield, gave an average age of 1,950 million years according to the lead method, but without isotope analysis (7). In the Canadian shield, lead isotope determinations for pegmatites from southeastern Manitoba resulted in an average age of 2,100 million years (8); for the same an average of 2,240 million years was derived by the radioactive rubidium-strontium method (7), in good agreement with the former value. At that remote epoch, unlike the present conditions, volcanic activity was prominent in the Canadian shield. The Manitoba pegmatites are associated with granitic intrusions into older rocks which reveal traces of a long previous geological history, and whose age can be estimated at 2,550 million years (12, 22). Recent very consistent lead-isotope age determinations (9) have yielded still higher values of age for some samples from the continental shields:

<table>
<thead>
<tr>
<th>LOCALITY OF SAMPLE</th>
<th>AGE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Millions of years</td>
</tr>
<tr>
<td>Ivigtut, Greenland</td>
<td>1,830</td>
</tr>
<tr>
<td>Yellowknife Area, N. W. T., Canada</td>
<td>2,140</td>
</tr>
<tr>
<td>Horseshoe Island, Great Slave Lake, Canada</td>
<td>2,180</td>
</tr>
<tr>
<td>Phoenix Mine, Norseman, W. Australia</td>
<td>2,190</td>
</tr>
<tr>
<td>Borderline Mine, Busia, E. Province, Uganda</td>
<td>2,220</td>
</tr>
<tr>
<td>Risks Mine, Kenya</td>
<td>2,220</td>
</tr>
<tr>
<td>Copperhead Mine, W. Australia</td>
<td>2,300</td>
</tr>
<tr>
<td>Inguladhal, Mysore, India</td>
<td>2,300</td>
</tr>
<tr>
<td>Sioux Lookout, Ontario</td>
<td>2,310</td>
</tr>
<tr>
<td>Steel Rock Lake, Ontario</td>
<td>2,360</td>
</tr>
<tr>
<td>Rosetta Mine, South Africa</td>
<td>2,860</td>
</tr>
<tr>
<td>Sierra Leone, Br. W. Africa (13)</td>
<td>2,930</td>
</tr>
</tbody>
</table>

Pegmatites from the Rhodesian shield, near Bulawayo, yield an age of 2,640 million years (10), yet the surrounding rocks—sediments
and lavas—are still older; and, what is more remarkable, remains of primitive plants—algal structures—are found there in graphitic limestone (Macgregor 1940, 1941). This provides "indubitable evidence that life has existed for at least 2,600 million years and probably for considerably longer than 2,700 million years" (10). Similarly, signs of organic remains are either found or strongly suspected in the rocks of Lake Superior and Manitoba, which are 2,000 to 2,500 million years old (11, 12).

Thus, direct measurements set the minimum age of the earth's crust at 2,900 million years, the oldest specimens being found in Africa. This confirms also a long-maintained belief that Africa was the first continent to be formed.

It is but natural to expect that the oldest rocks have not yet found their way into man's laboratories, and that the age of the earth's crust is greater than the presently known oldest sample.

An ingenious method, based on data for rocks of widely different ages from 25 to 1,330 million years as determined by Nier, led Holmes (14) to the calculation of the true age of the earth, or the time during which radiogenic lead has been produced in its materials. The method is one of extrapolation, consisting in the study of the observed relative isotopic abundances within one age group and their theoretical variation with time. The moment when the isotopic ratios, calculated backward for various age groups, become equal is the "beginning." Holmes found in such a manner 3,350 million years for the age of the earth. The oldest analyzed sample used in his calculation was 2,000 million years younger than this figure—a gap which might have caused considerable error in the extrapolation, as was pointed out by Jeffreys (15).

Recent data as quoted above push the directly observed age limit much farther back in time, and nearer the beginning. Applying the method of extrapolation to modern data, the probable age of the earth results as 3,500 million years (16), in excellent agreement with Holmes' former figure, but more reliable, the range of extrapolation being now only a few hundred million years.

The figure of 3.5 thousand million years can at present be accepted as a close approximation to the age of the earth—the time elapsed since its elements were uniformly mixed, probably in a molten state. The same figure, or one perhaps only slightly greater, can be considered the age of the solar system; the formation of the planets and of the earth's crust must have taken relatively short intervals of time (17).

The uranium and thorium content of iron meteorites is so small that their lead can be assumed to be of primeval isotopic composition, no radiogenic lead having been added in the course of time. If this
is so, the present average isotopic composition of lead, uranium, and thorium in the earth's crust indicates an age of 4,500 million years (18, 19), in remarkable agreement with a similarly determined age of stony meteorites (20). This would be the time elapsed since the separation of the iron from the silicate phase, which may have taken place in a diffuse state of matter and may have preceded the formation of the planets.

THE AGE OF THE ELEMENTS

Time intervals can be calculated only for radioactive elements with a known rate of decay. According to the well-known laws of radioactive decay, the amount of these elements decreases exponentially with time; calculating their amounts for distant epochs in the past, one inevitably arrives at time limits beyond which the calculated abundances of radioactive isotopes become unreasonably large—greater than those of the presently observed end products, or even greater than the total amount of matter in the universe. Clearly, the radioactive elements can only be of finite age. Now, the rate of decay of radioactive elements is not influenced by external conditions if the temperature remains below 1,000 million degrees and the density below, say, one million times that of water. Neither in the interior of normal (dwarf or "main-sequence") stars, nor in interstellar clouds from which suns and planetary systems are believed to have sprung, do such extreme conditions exist. We may well say that the state of matter in the observable universe requires radioactive decay to proceed relentlessly. As this could happen only for a finite interval of time, it would mean that the observable agglomerates of matter in the universe could also have existed for only a limited time.

Thus, at a remote epoch a building up of the radioactive isotopes must have taken place, in addition to their spontaneous or forced decay. Now, conditions leading to the formation of the heavy radioactive isotopes will throw the rest of the lighter elements into a melting pot, too—will cause their rapid building up and disintegration; this is a trivial consequence of the theory of nuclear structure. The age of the radioactive isotopes is thus almost synonymous with the age of the elements.

According to a method proposed by Russell (21), a maximum age for the elements can be derived from the relative terrestrial abundances of a radioactive isotope and its end product. It yields a maximum age, because some of the end product must have been created nonradiogenically in the initial "melting pot," when all the elements came into being under extreme conditions of temperature and density. Of the different isotopes, that leading to the lowest estimate of age is to be taken. The upper limit of age of the terrestrial elements thus found
equals 5 to 6 thousand million years, less than the double of the age of the earth (1, 16, 22, 23). The sharpest margin results from the uranium 235–lead 207 ratio. The closeness of the order of magnitude of the upper limit to the age of the earth is significant and makes it likely that the true age of the elements does not differ much from that of the earth—a figure of about 4,500 million years appearing to be plausible.

In these estimates there is some uncertainty from the unknown composition of the earth’s interior, which, however, is hardly significant in view of the exponential law of variation of abundance ratios with time. Even a large error in the present ratio will not affect the order of magnitude of the resulting age. The mere presence of radioactive substances is a proof of the temporal origin of the terrestrial elements.

Except for meteorites, there are no data available as to the abundance of radioactive isotopes outside the earth; the above-mentioned time limit refers therefore strictly only to the sample of matter represented by our globe.

Although the relative abundances of the elements (excluding the lightest, which have escaped from small bodies like our planet) in the earth’s crust and in the atmospheres of the sun and most stars are very similar, this does not necessarily mean a simultaneous origin for their elements. Only a similar mode of origin is implied.

Several attempts have been made to explain, with more or less success, the origin and relative abundances of the elements by equilibrium conditions inside superdense stars (Klein, Beskow, and Treffenberg; Hoyle; van Albade). Supernova explosions inject the mixture into space, whence it condenses again into new-born stars (24). Observations of the Crab Nebula (25), a former supernova, suggest that the product of explosion—the amorphous core of the nebula—is poor in hydrogen, whereas its hydrogen fringe appears to be interstellar gas pushed ahead of the expanding core. We may thus have a double origin for the elements: hydrogen already present in space with an unknown original content of other elements; and the heavier elements enriching the mixture through supernova explosions.

Old stars—those of “Population II”—seem to show, indeed, a smaller metal content than those believed to be more recently formed (26), suggesting a gradual change in the composition of the medium of which stars are built. If this is so, we need not go to the beginnings of the universe to account for the radioactive isotopes on earth: they may be the products of supernova explosions that preceded the formation of the solar system.

Nevertheless, serious doubts with respect to the latter conclusion are justified. The theory of stellar structure would admit the building
up of the lighter metal nuclei in superdense stellar cores (27, 28), as well as during the hydrogen explosion of a star that has become hydro-
dynamically unstable (24). From this, however, it is a long way to the extreme conditions at which uranium and similar elements of high atomic number can begin to be produced. Although light metals can, indeed, be currently supplied by the above-described mechanism, it is doubtful whether the heavy radioactive isotopes could originate in stellar interiors. More likely, these isotopes have come into being in a more powerful “explosion” which involved the whole universe, namely, that which happened at an early stage of its expansion. In that case the age of at least the heavier terrestrial elements would still be synonymous with the age of the world.

This leads us to another group of theories which explain the observed abundances of all elements, including the heaviest, by their building up from a nonequilibrium, extremely hot mixture (chiefly neutron gas) at an early stage of an exploding universe (Alpher, Bethe, Gamow).

It is possible that the lighter elements (say, those lighter than iron) have originated from two different processes—during the primordial explosion, and currently in stellar interiors—whereas the heavy isotopes were all created at the “beginning of the world”; in such a case, as shown above, the “radioactive” age of the universe, or the time elapsed since the big explosion, is about 4.5 thousand million years. However, unless the possibility of formation of the heavy elements in superdense stellar interiors can be definitely disproved, a certain ambiguity will remain attached to the meaning of this figure.

**METEORITES**

The pioneer work of Paneth 20 years ago raised hopes that radioactive age determinations of meteorites, based on their helium content, might yield a clue to the age of the solar system at least, or even to that of the whole universe (1). Unfortunately, the meteorites did not come up to original expectations. Paneth’s struggle with this problem, which is not concluded yet, led over disappointments and disclaimers of former results; e.g., he announced that all his determinations prior to 1940 were technically unreliable. Paneth’s researches are a remarkable example of a gallant fight for the truth, without bias toward his former work, some of which he rejected as soon as it was found that it did not comply with his own high standards.

The leakage of helium from meteorites to space was one of the many difficulties, and for this reason stony meteorites proved unreliable, so that only data referring to iron meteorites could be fully trusted. From refined analysis of the helium content of the latter Paneth found the ages of meteorites to lie between 100,000 and 9 thousand million
years. The higher values represented a puzzle as, for example, they considerably exceeded the upper limit of age for the earth and the solar system as set by the abundances of radioactive isotopes (cf. preceding section).

Now came the latest act of the drama. Bauer (29) and Huntley (30) pointed out that part of the helium in meteorites must have been produced by nuclear transmutations, caused by cosmic rays during millions of centuries. This suggestion has now become an established fact, as otherwise the presence in meteorites of the isotope He\textsuperscript{3} in considerable amounts (18 to 32 percent of He\textsuperscript{4}) cannot be explained: radioactive disintegration leads to He\textsuperscript{4} only, not to He\textsuperscript{3}. On the other hand, cosmic rays produce both isotopes in the approximate proportion of 10 He\textsuperscript{4} to 3 He\textsuperscript{3} atoms (Le Couteur). This ratio being given, an analysis by the mass spectrograph leads to the determination of the amount of purely radiogenic He\textsuperscript{4}, which is very much less than the total amount of helium. As a result, the estimated ages of meteorites are greatly reduced and, from the provisional data available, hardly attain 1,000 million years (31). This is much less than the well-established age of the earth and the solar system; therefore, the method is of no avail in estimating the age of the universe. It has been suggested that the meteorites lost their original helium when passing near the sun and melting in its heat; their orbits are sometimes likely to become highly eccentric from perturbations at close approaches to the planets, in which case near passages to the sun become possible. However, unpublished calculations by the writer show that such happenings are very rare, and that the explanation is invalid.

Urey (32) pointed out that iron meteorites are unlikely to contain enough radioactive elements to account for measurable amounts of radiogenic helium. The correlation between the total amount of helium and its isotopic ratio in iron meteorites is highly remarkable (31). In the opinion of the author of this review the simplest explanation of Paneth's results could be that all the helium is produced by cosmic rays, the absolute amount and isotopic ratio depending upon the original thickness of the protective layer, subsequently lost through ablation in our atmosphere. The time of separation of the stone and iron of meteorites, as determined from the isotopic composition of lead, is consistently found to be 4,500 millions years (20). This may refer to a preplanetary stage. Potassium-argon-40 ages of stony meteorites are found to be 1,900 to 3,800 million years (33) and 4,700 to 4,800 million years (34). Evidently there has been little or no escape of argon from stony meteorites. The argon ages would date from the moment of last solidification, thus probably from a planetary or postplanetary stage. Also, these high argon ages of stone seem to indicate again that the helium ages of iron inclusions, often connected with stone, are unreliable.
Meteorites point to an age of the solar system, or its parent nebula, close to 4,500 million years.

THE AGES OF THE STARS

At present there is little doubt that main-sequence ("dwarf") stars depend upon the conversion of hydrogen into helium for their energy source. The correlation of radius and mass, indicating central temperatures of precisely the range required by the corresponding slow nuclear reactions, can hardly be interpreted in a different manner. This knowledge is so well founded that it furnishes a reliable basis for the calculation of time rates of stellar evolution.

To cover radiation losses to space, the sun has to spend an amount of hydrogen very nearly equal to 1 percent of its mass in 1,000 million years. Sirius, a typical star of spectrum AO quite common in the galaxy, emits 13 times more energy per unit mass than the sun, consuming thus 13 percent hydrogen by weight in 1,000 million years. With 60 percent hydrogen originally, the store of energy would last 4,600 million years. There is probably not much mixing in stars outside their central regions (35, 36, 37, 38); therefore, only about 25 percent of the fuel is available (from the central regions where the temperature is high enough for nuclear reactions to proceed at a non-negligible rate), and the lifetime of Sirius becomes 1,150 million years. It may then become a giant (35), and ultimately collapse—possibly by throwing off its outer shell in a supernova explosion, leaving behind a remnant which ultimately becomes a white dwarf. The success in calculating "composite" models of red giants (39, 40), as well as Trumpler's classification of star clusters, lends support to this concept of stellar evolution. The more massive B stars will have a lifetime of a few hundred million years only. This being much shorter than the lifetime of the galaxy, which cannot be younger than the earth, it is concluded that the early-type stars are currently replaced by new stars condensing out of diffuse matter (35). Where diffuse matter is no longer available, early-type stars are absent and only giants of the corresponding luminosity remain, as is actually observed in globular clusters. Using Baade's terminology (41), Population II of the globular clusters, the galactic center, and the general galactic background, consists of aging members born at a remote epoch; whereas Population I, connected with the diffuse matter and spiral structure of the galaxy, contains young early-type stars steadily coming into being and dying, in addition to the background of less massive young and old stars, some of the latter existing from the very beginning of the galaxy (35, 42).

The absence of normal B and A stars from the globular clusters sets their age, as well as that of the galaxy, at more than 1,500 million years. The energy source of the giants remains a puzzle. If we take their
persistent appearance in globular clusters as an indication of their longevity, a more powerful source of energy must be assumed for their maintenance (35)—either gravitation of their superdense cores, or annihilation of matter. On the other hand, these giants may represent short-lived objects in "statistical equilibrium" with the rest of the stellar population—those which blow up or collapse being replaced by others becoming giants. This latter concept would agree with the calculated red-giant models (39, 40) which are supposed not to draw on unknown sources of energy and are short-lived, their luminosities being abnormally high as compared with their masses. The giants of the globular clusters, as well as the short-period variables which should represent a phase preceding the giant stage, would then correspond to stars of more or less similar mass for which the exhaustion of hydrogen has reached a critical limit (35). Taking the observed luminosities with Schwarzschild's models, the limiting mass would be from 3.0 to 2.0 solar mass, indicating for the clusters an age between 800 and 2,500 million years.

The fork-shaped H-R (Hertzsprung-Russell) diagram of the globular clusters represents apparently the result of aging, in contrast to the continually rejuvenated Population I of our galactic surroundings (the difference in metal content having only a secondary effect). The globular clusters, which are all well outside the galactic plane and are not sharing in galactic rotation, will necessarily oscillate on both sides of the galactic plane, the period of oscillation being less than 100 million years (Oort). Thus, they must have repeatedly gone through the galactic plane. While passing for the first time through the plane, they must have been stripped of all their diffuse matter—which could have been but loosely bound by a gravitational potential of only 1/1000th that of the galaxy—through collision with the diffuse matter near the galactic plane; the mechanism is similar to that visualized by Spitzer and Baade (43) for collisions of galaxies. This would have prevented the subsequent formation of new stars in them. The stellar population of the globular clusters must therefore consist of members of almost the same age, which came into being when the galaxy was formed, and represents thus one of the oldest time indicators. The lower branch of their H-R fork appears to join the H-R diagram of Population I at absolute bolometric magnitude +2 (41); this should be the luminosity of old stars which have now arrived at the end of their career as dwarfs.

The evolution of dwarf stars, without much mixing of their substance, amounts to chemical changes around their central cores, where hydrogen is converted into helium; the composition of the outer regions remains unchanged. Öpik (27) has followed the evolution of such stellar models by numerical integrations. From these calcula-
tions it can be estimated that stars which have nearly exhausted their central store of hydrogen yet remain dwarfs should be about 0.5 mag (or by 60 percent) brighter than "normal" dwarfs of equal mass. If we take this into account, it is estimated that the above-mentioned "ultimate" dwarfs in globular clusters, about 10 times brighter than our sun, should have a mass of $1.7\odot$. The total duration of the dwarf stage at this mass would be around 4,000 million years.

The numerical value of this estimate may be considerably in error; yet, qualitatively there is little doubt about the soundness of the interpretation which ascribes to the stellar population of the globular clusters the same age as that of the galaxy itself. By essentially the same method, but on the basis of more recent observational data, Sandage (44) finds an age of about 5,000 million years for the globular clusters. We may take the average of the two estimates, 4.5 thousand million years, as the probable age of the globular clusters, as well as of our galaxy.

Among the many data concordantly pointing to an age of the stellar universe of a few thousand million years, there is one which seemingly strikes a note of discord—some uneasiness may be felt about the high frequency of white dwarfs. If they are remnants of supernovae, which appear only once in a few hundred years, they would have required perhaps 100,000 million years to accumulate. However, at the beginnings of the galaxy, at the time when Population II was formed, star formation must have proceeded at a faster rate than now. The frequency of supernovae, directly related to the frequency of formation of massive stars, may then have been much higher (26). Further, the possibility of white dwarfs being formed in another way cannot be ruled out. Doubts as to the time scale cannot be maintained on such slender evidence.

Besides, a direct estimate of the age of individual white dwarfs can also be made, and this turns out to be in agreement with the other estimates. The energy source of white dwarfs can consist only in the thermal agitation of atomic nuclei (45) or upon explicit heat—like a kettle of hot water gradually cooling. The time of cooling, until the present state is reached, or the age of a white dwarf can be easily calculated when $A$, the mean atomic weight of this material, is known. Considering that all hydrogen must previously have been converted into helium, and that, before the "degenerate" stage of a white dwarf is reached, triple collisions at temperatures of a few hundred million degrees will convert all the helium into carbon, and then into lighter metals such as magnesium, we find that $A=24$ can be assumed, and Mestel's highest values for the ages of white dwarfs become equal to 4,000 million years. This may be near the age of Population II and the galaxy, in thrilling agreement with other estimates (27, 28).
The dynamical stability of clusters has been investigated repeatedly, with the result that most galactic clusters will dissolve, either under the tidal action of the galactic center or through encounters with field stars or other members of the cluster, in time intervals of the order of 1,000 million years (46, 47). Although this statement refers to the future and, theoretically, is compatible with an unlimited past, the probability of simultaneous occurrence of a great number of old clusters which just now have come to the verge of disruption is very small. We may expect an average cluster to be observed in the middle of its lifetime, and assume, therefore, that the age of most clusters is some 1,000 million years or less. Yet, most of them contain early-type stars which cannot be very old. Consideration of the dynamical stability of clusters confirms thus the youth of their members, and adds another argument in favor of the theory that stars are being born continually. Apart from that, no new criterion of age for the galaxy is forthcoming—clusters which are older than their stellar content cannot be observed.

The situation is similar with wide double stars. The distribution of the distances between their components (48, 49) indicates that equipartition of energy cannot have taken place (50), and that the binaries could not have been subjected to encounters with field stars for longer than, say, 5,000 million years (4). On the other hand, the statistical material from which this conclusion is drawn is based chiefly on the relatively luminous A-type binaries which, according to the preceding, cannot have lived to so great an age, anyway.

Thus, conclusions as to age based on the dynamical stability of clusters and double stars are overruled by the shorter lifetime of their components, and can be used only to reaffirm the short time scale of stellar evolution.

The observations by V. M. Slipher, Hubble, and Humason, if interpreted in a straightforward manner, indicate a recession of the extragalactic nebulae proportional to distance, or an expansion at a uniform rate of the visible portion of the universe.

Recent developments have shown, in a manner that leaves practically no doubt, that Hubble’s scale of distances should be at least doubled. The distances of the nearest nebulae were determined by Hubble from the period-luminosity relation of the long-period cepheids. The zero point of this relation depended upon space absorption in low galactic latitudes, and was known to be inaccurate, but, for lack of better data, it was accepted and used during the past quarter of a century as a basis for work on the structure of the universe. Some cosmological theories
actually depended upon the particular value of the zero point and the resulting scale of distances. The unexpectedly large correction in the scale is a shock to all theories involving the so-called cosmological constant. We need not express regret that these theories were created—they were fully justified by the esthetic value alone—but, from the standpoint of economy of thought, the cosmological constant (equivalent to a repulsion) must be suspended from active duty for the time being and put in cold storage until new observational facts sound the trumpet for its revival. It is rather doubtful whether this ever will happen.

The zero point of cepheid luminosities affects only the distances of extragalactic nebulae. Within the galaxy, including the globular clusters, a more reliable criterion of distance is offered by the known luminosities of the short-period cepheids, the so-called cluster-type or RR Lyrae variables. The average luminosity of these Population II high-velocity objects does not depend so much upon space absorption, and is well determined. They were too faint to be observed in the nebulae by Hubble. In the Magellanic Clouds, whose estimated distances depended also upon long-period cepheids, persistent Harvard Observatory searches failed to reveal cluster-type variables, a circumstance sometimes interpreted even as indicating the actual absence of these objects.

Now, as last, numerous cluster-type variables have been found in the Magellanic Clouds (51), but about 1.3 mag (or 3.3 times) fainter than expected from the magnitudes of the long-period cepheids. Thus, the long-period cepheids are 1.3 mag brighter and all distances based on them 1.8 times greater than was formerly assumed. The apparent diameters and integrated luminosities of globular clusters in external galaxies call for a similar correction (52), and independent support for these conclusions is forthcoming from other sources (Baade).

This, however, is not the whole story. The recession constant of the nebulae depends entirely on the more distant objects, for obvious reasons; yet in these no variable stars could be observed. Their distances were linked to the cepheid scale of the nearer galaxies through intermediate criteria—the magnitudes of the brightest stars and of the nebulae themselves. Both criteria are of a statistical nature and not only involve various photometric errors, but also depend upon the true dispersion (variety) of the magnitudes of the objects used as standards; the dispersions, and therefore the distances, seem to have been underestimated by Hubble. A comprehensive survey of the problem has been given by Behr (53). He concludes that those of Hubble’s intrinsic luminosities of the nebulae which are not based on variable stars should be increased by 1.7 mag. Behr was not aware of the need for adjustment of the cepheid scale of the nearest nebulae, and this
correction, evidently, must be added to that found by him. The total correction amounts thus to $1.7 + 1.3 = 3.0$ mag, or an increase in the distances of nebulae (except the nearest, which are based on cepheids) in a ratio of 4 to 1. The constant of recession, or the rate of increase of velocity with distance as based on observed red shifts, now becomes 145 km./sec. per magaparsec (3.25 million light-years), only one-quarter of the formerly assumed value. The expansionistic time scales are increased fourfold, and even the shortest will yield more than the lower limit—the age of the earth.

The retention of the cosmological constant by Eddington and Lemaitre was justified by the need to extend the time scale; the slow phase of expansion, when gravitational attraction and cosmic repulsion nearly balanced each other, allowed this to be done almost indefinitely. Now, with the increased distances, cosmic repulsion becomes a superstructure of a purely esthetic nature, serving no practical purpose. Besides, Einstein, the originator of the concept, has disavowed the cosmological constant ever since, in spite of the then favorable numerical aspect of the problem.

Without the cosmological constant, the Friedmann-Einstein cosmological models furnish a working hypothesis best suited to deal with the expanding universe. These models are very similar to an ordinary gravitating sphere in uniform expansion. Gravitation, working against expansion, is slowing it down. When the velocity of expansion is below a certain limit, the expansion will be ultimately stopped by gravitation, and contraction will start; when the velocity of expansion equals or exceeds the limit (velocity of escape), gravitation will be unable to stop it and the sphere will disperse into space, expansion never ceasing. According to the general theory of relativity, and without cosmological repulsion, a similar state of affairs in the expanding universe prevails. The first case, when expansion is ultimately stopped by gravitation, would correspond to positive curvature of space, or to closed space and a relapse of the universe, after maximum expansion, into the original state of high density (atom or nebula). The second case would correspond to zero or negative curvature, to open and infinite space, and to a one-way development of the universe by perpetual expansion.

For an expansion constant of 145 km./sec. per megaparsec the line between the two cases is set by a certain limiting value of the average density of matter in space (i. e., if all the matter of the universe were spread uniformly over its entire volume, instead of being concentrated into galaxies, stars, and atoms), equal to $3.9 \times 10^{-29}$ gm./cm.$^3$. The volume of the earth filled with matter of so low a density would contain only a mass of 42 milligrams.

The probable value of the average density of matter in space can be estimated in the following way. There are in the universe, on the
average, 12 nebulae per cubic megaparsec (55). The average mass per nebula, including intergalactic matter, can be estimated from the internal motions in clusters of galaxies according to the "virial theorem" (mean kinetic energy per unit mass proportional to the potential of gravitation); this, of course, depends upon the assumption that the clusters are held together by gravitation. The assumption can nowadays hardly be subjected to doubt, considering that otherwise, with the velocities observed, the clusters would have dispersed long ago; on the contrary, they are gathered so closely together that numerous interpenetrations or "collisions" of the member galaxies of a cluster must have happened during the lifetime of the universe (43). Repeated collisions must have led to "statistical equilibrium" in the distribution of velocities of the member galaxies; the similarity between the radial density distribution of nebulae in these clusters and that of an isothermal gas sphere (56) supports this assumption and the validity of the virial theorem. For the Virgo cluster a mass of 500,000 million suns per nebula results with Hubble's scale of distances (5), and four times as much with the corrected scale. These data lead to a world density of $2.5 \times 10^{-29}$ gm./cm.$^3$ or 64 percent of the critical density. If the result is taken literally, this would mean negative curvature, an open and infinite space into which the universe is irreversibly expanding.

However, the calculations are not exact enough to warrant unreserved acceptance of such a conclusion. The estimate has come astonishingly close to the critical density, and therefore, within the limits of uncertainty in the data, the alternative case of closed space and limited expansion followed by collapse is also possible. Indeed, Zwicky (57) finds considerable amounts of matter in the space between the galaxies, and favors a world density about 25 times that of our estimate, which would bring it far above the critical value. However, Zwicky's value is a very rough estimate, not based on the virial theorem. Our estimate of 2 million million suns per nebula would ascribe 90 percent of the mass to intergalactic matter (that between the galaxies) and only 10 percent to the galaxies themselves; this figure seems to be more realistic than Zwicky's, which would set the percentages at 99.5 and 0.5, respectively.

It is, perhaps, permissible to speculate on the closeness of the world density to its critical value, and to suggest an intrinsic reason for this near equality of the kinetic energy of expansion and the absolute value of the gravitational potential. The reason should be sought in the past history of the world. For example, an oscillating universe whose maximum world radius greatly exceeds the present value would lead to the above-mentioned near equality except when close to the phase of greatest expansion (which should be far ahead of present
time). In that case the time of expansion from the state of greatest density until today is insensitive to the precise value of world density, and depends only upon the rate of expansion; it is practically equal to that of uncurved (Euclidean) space and, with the revised value of the expansion constant, becomes

\[ t = 4,500 \text{ million years.} \]

The figure is surprisingly close to the other estimates, although a considerable uncertainty is involved, the extreme admissible values being, perhaps, from 3 to 6 thousand million years.

This would represent the age of the universe in a restricted sense, or the time elapsed since it was in a highly condensed state. This state cannot yet be described. Lemaitre's primeval atom is one of the possibilities. The theory of the origin of the elements, as shown above, does not provide a clue. The same is true of the cosmic rays, which appear to be of stellar origin and whose connection with the prestellar stage of the universe seems to be improbable (58, 59).

**SPACE REDDENING OF THE GALAXIES**

This phenomenon, announced by Stebbins and Whitford (60), and consisting in an increase of the color index of distant galaxies, not accounted for by the red shift, led to far-reaching speculations on observable effects of stellar evolution. The effect seems to be restricted to elliptical nebulae (purely Population II), whereas spirals (mixed populations) do not show reddening (61). The distant nebulae are observed at an earlier stage of evolution (on account of light time), and it has been suggested that the effect could be accounted for by the red giants of Population II disappearing with time (blowing up or collapsing), which would tend to make the population bluer. However, a multicolor study of the spectral-energy distribution of a distant elliptical nebula has shown that “the result is definitely not that expected from the death of red giants” (62). The effect continues to the greatest distances (63).

Vaucouleurs (64) suggested that the effect is due to the depression in the ultraviolet produced by absorption lines and bands. With the red shift the ultraviolet depression is displaced into the blue, making the blue-red color index redder. At least part of the effect can be accounted for in such a manner (65).

As to the spirals, they are known to contain a considerable amount of nebulosity in emission (66); this, especially that due to hydrogen, will fill the ultraviolet depression, counterbalancing the absorption. The absence of the reddening for spirals is thus explained without invoking stellar evolution. Over the time intervals involved, evolution may well affect individual stars, but considerable effects upon the entire Population II are unlikely.
In any case, the present evidence is inconclusive as to the reality of the space-reddening effect. Some residual reddening from absorption by intergalactic matter may exist. The question may be decided by two-color observations in yellow and red, avoiding the unreliable violet-blue spectral regions.

**ALTERNATIVE HYPOTHESES**

It has been repeatedly stressed that the nebular red shift may not indicate recession, and alternative suggestions have been made recently (Freundlich, Shelton). It is difficult to imagine a collisional process of reddening without simultaneous blurring of the nebular images (67). Further, the nonexpanding universe will be unstable and will end in collapse; or in expansion, if the cosmological repulsion is introduced. Thus, the present state would be exceptional, the normal state being one of Doppler shifts corresponding to real approach or recession. It does not seem advisable to sacrifice the solid concept of recession to a piling up of *ad hoc* new laws and improbable states.

Continuous creation of matter under various aspects (Kapp, Jordan, Bondi, Gold, Hoyle) is another alternative which would dispense with a finite age for the universe. It requires the retention of the cosmological constant (repulsion), or a pulsating variety of it (Kapp). For reasons similar to those given above these theories can at present be assessed only from the standpoint of their esthetic value. It is not easy to image observational criteria for them which cannot be explained away.

Perhaps the distribution of masses of the galaxies can provide the least objectionable proof. In Hoyle's expanding universe galaxies will continually grow by accretion, especially large ones with gaseous envelopes firmly bound by gravitation; the envelopes will act as nets catching atoms from intergalactic space, or incorporating whole gaseous envelopes of smaller galaxies which happen to be in their way (43). They will grow almost indefinitely with time. Their frequency per unit volume in Hoyle's universe will vary inversely as the cube of age, thus more or less as the cube of mass, too; when selected by apparent magnitude, there will be no upper limit of mass and almost no correlation of distance with magnitude. The available evidence implies a frequency of nebulae in space decreasing with the 4/3 power of mass (68, 69), a definite upper limit of mass (66), and a correlation of distance with apparent magnitude. What evidence there is, is definitely negative.

**CONCLUSION**

The rate of irreversible processes in different physical complexes—the radioactive elements, the earth and the solar system, the stars, stellar systems, the galaxy, the observable portion of the extragalactic
universe—is such as to suggest an age not exceeding 6,000 million years for the universe in its present form and content. The extragalactic nebulae, with our galaxy and its backbone of Population II, may have been formed some 4,500 million years ago, the sun as a star of Population I coming into being perhaps later.

Cosmological repulsion is a theoretical superstructure which is not necessarily required by the existing observational evidence. The same is true of the continuous creation of matter and the alternative interpretations of the nebular red shift; these are mere possibilities, serving the purely esthetic purpose of denying the universe a temporal origin.

The observed velocities of recession exceed one-fifth of the velocity of light, the energy corresponding to a packing fraction (fraction of mass converted into kinetic energy) of 0.02 per nucleon (proton or neutron). Nothing short of an explosion from the densest-known state of matter—nuclear fluid—could be advocated as the cause. Our knowledge of the present density of matter in the universe is insufficient to decide between the two possibilities: that of open space, in which case the whole universe is an irreversible process of temporal origin, and that of closed space, in which the universe may return to its initial state, implying oscillations—the collapsing universe rebounding from the elastic forces of the nuclear fluid at a state of maximum compression, to begin a new phase of expansion.

It may appear at first sight that, at an advanced stage of collapse, when all individual bodies have melted into a uniform gaseous mass, the gaseous universe may be prevented from further collapsing by the elastic forces of the gas itself, like an oscillating gaseous star of which the cepheids are examples. However, it is likely that, with the enormous kinetic energy of contraction, the universe will first pass quickly through the stage of building up of heavy elements from hydrogen and helium, most of the hydrogen remaining unconverted before the next stage, that of nuclear dissociation and formation of neutron gas, begins—electrons being squeezed into and absorbed by the positively charged atomic nuclei. This is the reverse of the process by which Gamow and others visualized the origin of the elements after the explosion of the primeval atom. Formation of neutron gas absorbs enormous amounts of energy, and this, so to speak, blows the bottom off the resistance of the gas to compression. In such a case, the so-called ratio of specific heats of the gas (mixed with strong radiation) is less than 4/3 and, according to a well-known theorem on the structure of gaseous spheres, the universe becomes intrinsically unstable and cannot cease collapsing while in a gaseous state. Only when the perfect-gas laws no longer are valid, i.e., when the stage of nuclear fluid is reached, will there develop enough resistance to stop the collapse and invert the trend of events.
In the case where open space appears to be required by the physical characteristics of our neighborhood, we never will be sure of its validity for the universe as a whole. The possibility should not be overlooked that what we observe now is merely the metagalaxy—only a step in the hierarchy of physical systems. The observed expansion may refer only to this limited, although large, material system; in other parts of the world conditions may be different. The finite intensity of the sky background has often been advocated to prove the finiteness of the world. However, as shown by C. V. L. Charlier on purely classical lines, an infinite world is compatible with a finite intensity of the sky background if the universe is built on a hierarchical principle, systems of each order (atomic nuclei, atoms, planets, stars, clusters, galaxies, metagalaxies, etc.) being separated by distances considerably greater than their diameters. Such a "hierarchically diluted" infinite universe has a finite and small surface brightness even in the absence of absorption or Doppler shifts.

In the case of closed space the universe (the whole, or the observable metagalaxy), with all its energy content, including radiation, is bound to return to the initial state of nuclear fluid. This course of events is likely to repeat itself, the universe oscillating without external loss, implying an unlimited age in the past and in the future (time here meaning simply a succession of events, irrespective of its numerical value). All the structural phases will return time and again without, however, an "eternal recurrence of all things" in Nietzsche's sense—the individual celestial bodies in successive oscillations would not be identical, nor would their inhabitants. On the contrary, an unlimited variety of combinations and of prospects of evolution would be possible during each phase of the oscillation.

Some have expressed disgust at the idea of an oscillating universe, periodically repeating its general features (2). The present writer cannot see why this great repetition should claim a lesser esthetic value than, e. g., the annual succession of seasons so praised by poet and layman. Besides, not only is the repetition never literally exact, but, alas, we have no say in the matter—the Plan was laid down without our being consulted beforehand.

**LITERATURE CITED**


Reprints of the various articles in this Report may be obtained, as long as the supply lasts, on request addressed to the Editorial and Publications Division, Smithsonian Institution, Washington 25, D. C.