Date	Julian Date	Comparisons	Ded. Mag.	Date	Julian Date	Comparisons	Ded. Mag.
1933				1933			
Feb. 17	7121	=22	9·9	Mar. 30	7162	4(3)v(1)5	8·2
23	7127	15(2)v(1)22	9·8	Apr. 3	7166	4 − 4, 5 + 2	8·1
28	7132	15(1)v(2)22	9·6	13	7176	4(1)v(1)5	8·0
Mar. 7	7139	=14, 15 + 1	9·4	19	7182	4 − 3, 5 + 3	8·0
13	7145	7(2)v(1)11	8·9	30	7193	4 − 3, 5 + 3	8·0
20	7152	7(1)v(3)11	8·7	May 5	7198	5(3)v(1)7	8·5
25	7157	5(1)v(1)7	8·5	14	7207	7(3)v(1)11	9·0
27	7159	5(1)v(2)7	8·4	22	7215	15(2)v(1)17	9·6

The Comparison Stars Used

Hagen's Number	H.P. Magnitude	Hagen's Number	H.P. Magnitude	Hagen's Number	H.P. Magnitude
4	7·7	17	9·6	43	10·9
5	8·3	19	9·7	51	11·2
6	8·6	21	9·8	55	11·4
7	8·6	22	9·9	64	11·8
9	8·9	24	9·9	69	12·0
11	9·1	32	10·3	73	12·3
14	9·4	34	10·4	75	12·4
15	9·5	40	10·7		

The values given above are obtained from the H.P. column in Hagen's *Atlas Stellarum Variabilium*, Series VI.

Brackenhurst,
Bucklebury Common, Berks:
1933 *June* 1.

ON THE EXPANDING UNIVERSE AND THE TIME-SCALE.

W. de Sitter.

1. I may be allowed to begin by recalling briefly the principal facts relating to the theory of the expanding universe.

The line-element of time-space of a universe containing matter and radiation, homogeneously and isotropically distributed over the three-dimensional space, is

$$ds^2 = R_1^2[-y^2 d\sigma^2 + d\tau^2], \qquad (1)$$

where $d\sigma$ is the line-element of a three-dimensional space of constant unit curvature k ($k = -1$, 0 or $+1$) and y is a function of τ alone. R_1 is a constant, which arises in the integration of the equation of energy.* In the simple

* Empty universes, having $R_1 = 0$, are not considered in this paper. The actual universe is not empty.

case of no pressure, which is a good approximation to the truth, the integral of this equation is

$$\kappa\rho R_1^2 y^3 = 3, \tag{2}$$

where ρ is the proper density. The factor 3 arises naturally: it is the number of dimensions of the three-dimensional space. In the actual universe, containing material pressure and radiation, and in which the distribution of matter and radiation differs from complete homogeneity and isotropy, the integral is more complicated, and cannot be explicitly stated, but it still exists, and contains a constant of integration. Since the integral must in the idealized case degenerate into (2), this constant is the same R_1. It depends on the material contents of the universe, and is the natural unit in which lengths and times are measured. In the actual universe it is probably of the order of 10^{27} cm. or 10^9 years. The instantaneous curvature of three-dimensional space is k/R^2, where $R = R_1 y$.

The differential equation determining y as a function of τ, *i.e.* Lemaître's equation, is, in the simple case in which the pressure is neglected,

$$\left(\frac{dy}{d\tau}\right)^2 = \frac{1}{y} - k + \gamma y^2, \tag{3}$$

where $k = -1$, 0 or $+1$ as before, and $\gamma = \frac{1}{3}R_1^2\lambda$, λ being the "cosmical constant" introduced by Einstein. The value of λ, or γ, is entirely unknown.

The acceleration, we may remark in parenthesis, is

$$\frac{d^2y}{d\tau^2} = \gamma y - \frac{1}{2y^2}. \tag{3'}$$

It consequently is independent of the curvature k, and can only be positive for positive values of γ.

The possible solutions of (3) are of three types, which I have called the expanding universes of the first and of the second kind and the oscillating universes, depending on the values of k and γ. For negative values of γ the solution is of the oscillating type, the value of y oscillating between zero and a maximum in a finite period (of the order of π). For $\gamma > +4/27$ the solution is an expanding universe of the first kind, in which y increases from zero to infinity, the velocity of increase $dy/d\tau$ being infinite both for $y = 0$ and for $y = \infty$. The time needed to increase from zero to a moderate value y is of the order of y itself. For very large y this time becomes of the order of log y, the solution approaching asymptotically to $y = e^{(\tau - \tau_0)\sqrt{\gamma}}$.* For values of γ between zero and $+4/27$ the solution is an expanding one of the first kind if k is either -1 or 0. For $k = +1$ two solutions are possible

* For large y the solution is:

$$\tau - \tau_0 = \frac{1}{\sqrt{\gamma}} lg y + \frac{k}{4\gamma^{3/2}} \cdot \frac{1}{y^2} - \frac{1}{6\gamma^{3/2}} \cdot \frac{1}{y^3} + \frac{3k^2}{32\gamma^{5/2}} \cdot \frac{1}{y^4} + \frac{3k}{20\gamma^{5/2}} \cdot \frac{1}{y^5} + \dots$$

For small y it is:

$$\tau - \tau_0 = \tfrac{2}{3}y^{\frac{3}{2}}[1 + \tfrac{3}{10}ky + \tfrac{9}{56}k^2y^2 + (\tfrac{5}{48}k - \tfrac{1}{6}\gamma)y^3 + \dots].$$

for these values of γ, viz. either an oscillating one with a period exceeding π and a maximum value of y smaller than $1\cdot5$, or an expanding universe of the second kind, in which y has a minimum value exceeding $1\cdot5$, and increases to infinity on both sides of the minimum. The time from the minimum to any moderate value of y is again of the order of y itself, and for very large y of the order of $\log y$. In the limiting case $\gamma = 4/27$ the maximum value of y in the oscillating universe is $1\cdot5$ and the period becomes infinite, the maximum being reached asymptotically for $\tau = +\infty$, whilst in the expanding universe of the second kind the minimum is $1\cdot5$ and is reached asymptotically at the time $\tau = -\infty$. The particular solution $y = \text{constant} = 1\cdot5$ ($k = +1$, $\gamma = 4/27$) is "Einstein's universe," and is unstable. For $\gamma = 0$ the universe is expanding of the first kind for $k = -1$ or 0, and oscillating for $k = +1$.

Astronomical observations give us no means whatever to decide which of these possible solutions corresponds to the actual universe. The choice must, as Sir Arthur Eddington says, depend on æsthetical considerations. Reasons will be given below for preferring the minimum value $y = 0$ to a finite minimum, thus limiting the choice to the expanding universes of the first kind and the oscillating ones. Personally I have, like Eddington, a strong dislike to a periodic universe, but that is a purely personal idiosyncrasy, not based on any physical or astronomical data. The values of k and γ remain undetermined (except that negative values of γ are excluded if oscillating universes are not admitted), but the following considerations are independent of these values.

2. The short time-scale of these expanding universes has been repeatedly commented upon. With the exception of the very special case $k = +1$, $\gamma = 4/27$, the period in the oscillating case, and the time elapsed since y had its minimum value (either zero or a finite value) in the expanding case, is only a few thousand million years. The time-scale is short not only in the past, but also in the future; except in the special case $k = 0$, $\gamma = 0$, the value of y in the expanding universes of both kinds will be so large (say a thousand times its present value) that the universe will practically be disintegrated into independent galaxies, after a time of the order of 10^{10} years. The universe consists of galaxies, which are at present separated from each other by distances of the order of 30 or 50 times their diameters. The galaxies consist of stars, for which this ratio is of the order of 10^7 or 10^8, and the stars consist of electrons and protons, for which the same ratio is of the order of 10^5 (so far as the "diameter" of an electron or a proton has any meaning). The present structure of the universe, so different from that of the galaxies and of the stars in the ratio of the mutual distances and the diameters of their constituents, is, however, only an episode of a very ephemeral character, lasting not longer than a few times 10^{10} years. The constitution and the dimensions of the stars and the galaxies are not affected by the expansion of the universe.

Nevertheless the "age of the universe," *i.e.* the time since y passed through its minimum, has very generally been assumed to be also the age of the stars, and consequently the opinion has generally been held that the long time-scale of 10^{12} or 10^{13} years, demanded by modern theories for the evolution of the stars, would have to be given up, and the theories of evolution would have to be modified so as to give a shorter time-scale. I think this identification of the time of minimum of y with the "beginning of the world" is entirely gratuitous. In the case of the expanding universes of the second kind it is at once evident that the minimum is not a very remarkable point on the curve at all. There is no singularity at that point and no discontinuity in the motion, no more than at the perihelion of a planetary or cometary orbit.

In the case of the expanding universes of the first kind and the oscillating universes, where the minimum of y is zero, the case is different. The equation (1) is an idealization of the problem, taking account of inertia only and neglecting gravitation, or, in other words, replacing the mutual gravitational action between the separate galaxies by the averaged action of all galaxies in the universe combined. This approximation is, of course, entirely sufficient so long as the mutual distances are large, but ceases to be an approximation when these distances become very small. The co-ordinates of a galaxy, expressed in natural measure, are $x_i = R_1 y \xi_i$, if ξ_i are the co-ordinates in the three-dimensional unit space of line-element $d\sigma$. Described in these co-ordinates x_i, the orbit of a galaxy in the idealized case, corresponding to the line-element (1), is a curve which near the origin assumes a parabolical form, having its apex in the neighbourhood of (but not at) the origin. The body passes through the origin at the time $t = t_0$, which is the same for all galaxies, being the time at which $y = 0$, and the velocity in natural measure at that time is the velocity of light. The accelerations are, however, different for the individual bodies, and at a small distance from the origin the velocities are also very different. The mutual perturbations near the origin will be large and different for different bodies, and the exact simultaneity and exact equality of all velocities will consequently be destroyed. An attempt has been made to estimate the amount of these perturbations, which will be referred to later on.

The conception of a universe shrinking to a mathematical point at one particular moment of time $t = t_0$ must thus be replaced by that of a near approach of all galaxies during a short interval of time near $t = t_0$. The minimum distance and the time when this is reached is different for each pair of galaxies, and the velocities, though very large, are not rigorously equal to the velocity of light, and are also individually different. The dimensions of the galaxies themselves are not directly affected. It should be remembered that the distribution of the stars in the galaxies is so very rare that they can easily penetrate each other. If we put a million galaxies in the space now occupied by one, the ratio of the mutual distances and the diameters of the stars will still be of the order of 10^5 or 10^6, and, though collisions or near approaches will be more frequent than at present, they are still exceptions, and there is no reason to suppose that the epoch of minimum y is of any special importance in the evolution of the stars.

I wish to emphasize that the above is not a new theory of the expanding universe, like Dr. Milne's, but simply an elaboration, or consequence, of Lemaître's general theory. The large velocities of the galaxies near $t = t_0$, just like their present systematic positive radial motions, are not constants of integration, different for each individual body, but are imparted to them by the gravitational action of the whole of the universe. The individual peculiar velocities are assumed to be small and accidental.

3. The adoption of the long time-scale for the evolution of the stars dates from the publication of Eddington's mass-luminosity curve in 1924 March. Theories of the generation of energy in the stars were then developed to account for this long time-scale. Lately other theories, ascribing the generation of energy to atomic processes instead of to the annihilation of matter, have been proposed, leading to a time-scale of the order of 10^{10} years. Primarily, however, what led to the adoption of the long time-scale was not the belief in annihilation of matter or the corresponding theory of evolution, but the observational fact of the existence of the mass-luminosity relation. If the stars of different luminosity (and consequently, according to the spectrum-luminosity relation, of different spectrum) represent different stages in one and the same process of evolution, then the rate of this evolution is determined by the rate at which the star can lose mass, since it must necessarily slide down the mass-luminosity curve. As there seems to be no other method for a star of losing mass than by radiation, this rate is necessarily slow, and the time-scale is long. With a short time-scale the stars must retain practically the same luminosity during their whole life, and the "main series" does not represent a sequence of evolutionary stages, but a variety of individually different creations. Apart from any judgment regarding the relative merits of different theories of the generation of energy in a star, I think that the empirically ascertained mass-luminosity relation is a very strong argument in favour of the long time-scale. We must therefore accept the paradox that the stars are older than the universe, if by the "age of the universe" we mean the time elapsed since y passed through its minimum. It has been shown, however, that this minimum must not be conceived as the "beginning of the world," but as a transitory episode in the history of the universe, so that there is nothing paradoxical left in the paradox.

4. The time $t = t_0$ of the minimum of y is between 10^9 and 10^{10} years ago. Now it is worthy of remark that there are at least two other entirely independent lines of reasoning, which point to this same time, say 5.10^9 years ago, as a critical epoch.

The age of the Earth's crust is of the order of a few thousand million years. This determination, resting on the chemical analysis of minerals and the laws of radioactive processes, leaves only a small margin of uncertainty. The age of the Earth itself is probably not more, and possibly much less, than twice the age of the crust. All modern theories of the origin of the solar system agree in ascribing this origin to a near approach or a collision of the Sun with another star, consequently the age of the Earth is also the age of the solar system. Near approaches in our galactic system as it is now are extremely rare. Unless we are prepared to consider our own system as a freak, practically unique amongst stars, we must believe that about 5.10^9 years ago the chance of encounters must have been much greater than it is now. The near approach and mutual penetration of many galaxies at that epoch, as explained above, does provide this increased chance of encounters.

The structure of our own galactic system, and of those external galaxies which are near enough to enable us to study their structure, is extremely complicated. The distribution of density is very irregular, and very far from homogeneous. The systems are all rotating in periods, which are of the order of a few hundred million years. As a consequence of this rotation the unhomogeneity of structure cannot exist for ever, but must be gradually smoothed out. The age of the galactic systems in their present form can therefore not be more than a small number of periods of revolution. Probably ten can still be considered as a small number in this respect, but almost certainly a hundred would be a large number. The galactic system and the spiral nebulæ must thus have acquired their present structure at a definite epoch in the past, their ages being of the order of $10 \times$ (a few times 10^8) years, *i.e.* of the same order of 5.10^9 years as the age of the planetary system and of the universe. It need not be pointed out expressly that again the near approach of all galaxies at that epoch affords sufficient explanation for the setting up of the rotation, the formation of the spiral arms and the unhomogeneous distribution of matter.

The great increase of density and consequent increased chance of encounters needed for the explanation of the origin of the spiral galaxies and the planetary systems is, however, only provided at the epoch of minimum if the universe is either an oscillating one or an expanding one of the first kind. In the expanding universes of the second kind the increase of density is hardly sufficient to produce these effects, since at the minimum the distances between the galaxies were probably still equal to several times the diameters.

5. The line-element (1) supposes a uniform density. If we wish to take account of the mutual gravitation of the individual galaxies, this must be replaced by a discontinuous distribution, the matter being concentrated in concrete separate points. The formation of a line-element corresponding to this discontinuous distribution is a formidable mathematical problem, which I have not attempted to tackle. In order to arrive at an estimate of the size of the perturbations in the motion of the galaxies by their mutual gravitational attraction, I have contented myself with considering only two galaxies, one of which is situated in the origin of co-ordinates. The problem thus reduces to that of finding the motion of a material point, expressed in natural measure, in the field produced by one galaxy at the origin of co-ordinates and the rest of the universe. The assumption is made that the line-element representing this field is of the form

$$ds^2 = -e^{-2\gamma}y^2 d\sigma^2 + e^{2\gamma} d\tau^2, \tag{4}$$

where R_1 has been taken as unit of length and time, and γ is supposed to be a function of the radius vector

$$r = y\chi$$

alone, χ being the radius vector in the unit space $d\sigma$. This line-element satisfies the field equations to the first order of γ, and it agrees, also to the first order of γ, with that recently derived by Dr. McVittie,* if we take for γ the ordinary Newtonian potential:

$$\gamma = -\frac{m}{r}. \tag{5}$$

For distances r smaller than the diameter of a galaxy, *i.e.* when the two galaxies considered interpenetrate each other, (5) may, in analogy with Newtonian mechanics, be replaced by

$$\gamma = ar^2 + b, \tag{5'}$$

the constants a and b being so chosen that at the boundary, $(r = r_1)$, γ and $\gamma' = d\gamma/dr$ are the same by both formulæ. For m I have taken $\frac{1}{4}\cdot 10^{-10}$, which in the units chosen corresponds to about $10^{11}\odot$ for the sum of the masses. The "unperturbed" motion is a geodesic in the space $d\sigma$ described with a variable velocity. Expressed in the co-ordinates $x_i = y\xi_i$ it becomes a parabola-like curve, passing through the origin, the velocity at the origin being the velocity of light.† The "perturbation" is the difference between the geodesic corresponding to the line-element (4) and that corresponding to (1). The detailed computation of the perturbations will be published in one of the next numbers of the *B.A.N.* The perturbations are of the order of one or two parsecs, or a few years, both in the co-ordinates x_i and in the time τ_0 of passing through the origin. This rather small amount is the perturbation of the orbit of a galaxy by *one* other. That it is so small is evidently due to the enormous relative velocity, making the time during which there is any appreciable action very short. If there are n other galaxies their effects will be cumulative, and as a very rough approximation we can take their sum to be of the order of $\sqrt{n}$ times that by one unit. For $n = 10^6$ this would already give a total perturbation comparable with the dimensions of a galaxy. The total number of galaxies in the universe is, of course, much larger than 10^6, probably more nearly 10^{12}, if not infinite, but for such large numbers the proportionality of the total perturbation with $\sqrt{n}$ can no longer be supposed to be even a rough approximation.

Although this computation does not have the pretension to be more than a rough guess at the order of magnitude of the deviation of the motions in the actual universe from the idealized mathematical abstraction represented by (1), it seems to show that the hypothesis as to what will actually happen (or has happened) at the time $t = t_0$ developed in the paragraphs 2 and 4 is at least not improbable.

* *M.N.*, **93**, 325, 1933. McVittie's r is my χ, my r being his r_1. McVittie's $e^{\beta(t)}$ is my y.

† See *B.A.N.*, Nos. 193 and 200 (1930), and my Hitchcock Lectures, *University of California Publications in Mathematics*, **2**, No. 8, p. 186 (1933).