World Structure and the Expansion of the Universe

By Prof. E. A. MILNE, F.R.S.

THE most distant nebulæ appear to be receding from us, and the velocity of recession is proportional to the distance. The most commonly accepted explanation of this phenomenon is that due to Friedman and Lemaître. The principle of their explanation is that it is possible to describe the observed facts by assigning fixed co-ordinates to a distant nebula in a curved space-time in which the metric involves the time t. The spatial interval-distance between the nebula and the observer is then a function of t. The difficulties of this explanation are (1) that it involves the existence of ' cosmic time ' and restores the distinction between time and space abolished by Minkowski ; (2) that it has been impossible to explain why ' space ' is expanding and not contracting. This theory is a development of the remarkable pioneer theories of Einstein and de Sitter, which contemplated static metrics for space-time. De Sitter's world, it is true, placed time on the same footing as space, but Einstein's cylindrical world introduced " cosmic time ". More recently (*Proc. Nat. Acad. Sci.*, **18**, 213 ; 1932), Einstein and de Sitter have concluded that at the present moment it is impossible to determine the algebraic sign of the curvature of ' space ', and that the facts of observation can be described by assigning fixed co-ordinates to a distant nebula in a quasi-Euclidean space expanding with the time.

A much simpler explanation of the facts may be obtained as follows. The explanation abandons the curvature of space and the notion of expanding space, and regards the observed motions of distant nebulæ as their actual motions in Euclidean space.

Consider a spherical region of Euclidean space, occupied at time $t=0$ by a uniform spatial distribution of moving particles, moving in random directions with velocities u, v, w, distributed according to a law $f(u, v, w) \, du \, dv \, dw$. The density is supposed to be so small that collisions do not occur and forces of interaction are supposed negligible. Outside the sphere (say of radius r_0) let space be empty. Then the *outward* moving particles will move into the empty space outside and the faster particles will gain on the slower. At any time t the fastest moving particles will form an expanding spherical frontier, followed by the next fastest, and so on. The *inward* moving particles will traverse the sphere of radius r_0, emerge at the other side, and then move outwards. Thus at any sufficiently large time t all the particles moving with a given speed V will be found between the spheres of radii $Vt - r_0$ and $Vt + r_0$. We see at once that after the lapse of sufficient time, all the distant particles will have velocities of recession ; and the mean velocity of recession at any distant point will be

ultimately proportional to the distance, the constant of proportionality being simply $1/t$. The interior of the original sphere, and indeed all space inside the distant moving portion, remain occupied throughout. For example, near the centre of the original sphere, after a long time t only slow-moving particles will be found, namely, those which started inwards from the original frontier with sufficiently small velocities. The density everywhere decreases with the time, and the particles sort themselves out in velocity, the sorting becoming more perfect the larger the velocity.

Clearly the restriction to an initially uniformly occupied sphere is unnecessary. Any initial density distribution which decreases sufficiently rapidly with distance will give rise to the same effect. This is true even if the initial distribution fills infinite space. The only difference is that *every* small element of volume always contains a few slow-moving particles, even after a long time ; but the sorting goes on. The essential aspect of the situation is that we are dealing with an unenclosed system.

The above explanation is applicable at once to the system of nebulæ. The fastest moving ones will have velocities exceeding the velocity of escape against gravity, and will ultimately pursue curves indistinguishable from their linear asymptotes.

This common-sense explanation has many advantages in addition to that of rendering unnecessary the introduction of a curved ' space ' and a non-static metric. It shows at once that the system is necessarily an expanding system after a sufficiently long time. If at any instant all the velocities are imagined to be reversed in direction, the system appears to contract to its configuration at time $t=0$ and then expands again. Moreover, if at time $t=0$ all the velocities are reversed, the system still expands. The instant $t=0$ affords a natural origin of time for the observer recording the distribution-laws. From this instant evolution proceeds in an inevitable direction, namely, in the direction of expansion. We may say that creation and uni-directional evolution are brought into a single mathematical scheme.

It is quite unnecessary, however, to introduce a ' cosmic time '. Let an observer count the particles with velocities inside the range u to $u + du$, v to $v + dv$, w to $w + dw$, and arrive at the law $f(u, v, w) \, du \, dv \, dw$. Let a second observer, moving with uniform velocity V_0 with respect to the first, make a similar enumeration, and suppose he arrives at the law $f_1(u, v, w) \, du \, dv \, dw$. We may inquire what must be the actual distribution-law in any frame so that $f \equiv f_1$, that is, so that the velocity-pictures of the universe recorded by the two observers are identical. Neglecting the curvatures

of the paths, we find that the distribution-law must be

$$\frac{B\,c\,du\,dv\,dw}{(c^2 - u^2 - v^2 - w^2)^2}$$

where B is a pure number and c is the velocity of light. This permits a continuous distribution of velocities up to c. Since actual recession speeds have been recorded up to one-fifteenth of that of light (19,700 km./sec., see Humason, *Mt. Wilson Contributions*, No. 426, 1931), and since still higher velocities may be expected, there seems nothing fantastic about this distribution-law.

Suppose now two observers start at the centre $x = 0$, $y = 0$, $z = 0$ of the initial spherically symmetrical distribution at time $t = 0$, with relative velocity V_0 with respect to one another. We may inquire what must be the spatial distribution of particles such that the universe will for ever appear the same to each observer. This requires a more complicated analysis, for the two observers will disagree as to what is meant by simultaneity. To solve this problem it is necessary to introduce the concept of the intensity I of world-lines at a point in time-space in a given direction. This is defined as the number of world-lines per unit solid angle in 4-space per unit 3-space cross section normal to the direction of the world lines. The conditions that any proposed function $I = F(x, y, z, u, v, w, t)$ shall represent a concourse of permanent objects are that (1) F shall be invariant for a Lorentz transformation, (2) F shall be constant along a world-line ; the second condition is readily shown to be equivalent to the satisfaction of Boltzmann's equation in gas-kinetic theory. When F is determined and we return to the co-ordinates x, y, z, t of any particular observer measured from the natural origin of time-space, the distribution law is found to be

$$\frac{c^2\,A\,du\,dv\,dw\,dx\,dy\,dz}{(c^2 - \Sigma u^2)[(c^2 t - \Sigma u x)^2 - (c^2 t^2 - \Sigma x^2)(c^2 - \Sigma u^2)]^{\frac{3}{2}}}$$

which may also be written in the form

$$\frac{c^2\,A\,du\,dv\,dw\,dx\,dy\,dz}{(c^2 - \Sigma u^2)[c^2\Sigma(x - ut)^2 - \Sigma\{v(z - wt) - w(y - vt)\}^2]^{\frac{3}{2}}}$$

This shows at once that for given $x^2 + y^2 + z^2$ and given $u^2 + v^2 + w^2$, the density is a maximum (for sufficiently large t) when $u : v : w = x : y : z$, that is, it gives the recession predominance. It also gives the velocity-distance correlation for t large. Integrated over the complete spatial solid-angle for 3-space it gives for time $t = 0$ the distribution-law

$$\frac{4\pi c\,A\,du\,dv\,dw}{(c^2 - \Sigma u^2)^2} \cdot \frac{dr}{r}$$

showing that the density-distribution at any time t may be derived as the natural expansion of an initial distribution * with a density law $\rho \propto 1/r^3$. This of course gives a congestion of matter at the origin at time $t = 0$, and expanding spheres of

* For $t = 0$, the last formula gives the distribution *between* the radii r and $r + dr$. But at $t = 0$, at any given *point* r, the velocities are predominantly tangential.

singularities at any other time. This in turn shows that it is necessary to introduce gravitation and so curvature of world-lines as a second approximation. But the density law $\rho \propto 1/r^3$ gives a gravitational potential tending to a constant at large distances, and so preserves the analysis as representing the uniform rectilinear motions at great distances. Presumably a Gaussian metric is required for the vicinity of the space-time origin, but it tends to the Galileian form at great distances.

According to these very elementary considerations, which only involve the principles of the special theory of relativity, the continuum of time-space is occupied at large distances from the natural space-time origin by a hyper-complex of world-lines having spherical symmetry about the space-time origin. There is no such thing as cosmic time. But at each point of space-time there is a unique direction in which the space-time origin lies. Every observer sees the same velocity distribution at great distances. Every observer can regard himself as the centre of the universe by choosing his time axis so as to point away from the time-space origin—in other words, by choosing the motion of his frame of reference so that it is at rest with respect to the vector average of the motions in his immediate neighbourhood. The world is then perfectly ego-centric at all points, and the moving picture of the world as made by any one observer is identical with that made by any other observer.

The principle of relativity in its original form asserted that all frames of reference are equivalent for the description of the *laws* of Nature. The foregoing ideas rest on an extension of the principle to the assertion that the world itself, when local irregularities are disregarded, must appear to have the same structure to all observers ; in other words, the principle of relativity is extended from the *laws* of Nature to the *phenomena* occurring in Nature.

We cannot observe space. We observe point-events. But we can recognise the continued existence of material objects, and hence we can arrange these observed point-events in world-lines. It seems best to avoid the phrase ' the structure of space ' or of ' space-time ', and consider simply the structure of the hyper-complex of world-lines which can be reconstructed from our observations. The preceding analysis discusses the simplest ideal system of world-lines that is compatible with the observed permanence of material objects and satisfies the extended principle of relativity. It is not difficult to generalise the analysis so as to describe a hyper-complex of curved world-lines and to connect the distribution of intensity with the distribution of curvature, and thus to make some progress towards the introduction of gravitation even inside the limits of Galileian space-time. But these developments are deferred to another occasion. I conclude by emphasising the very simple explanation of the expansion of the universe of material objects obtained by examining the kinematics of an unenclosed system with central condensation.