

RED SHIFTS AND THE DISTRIBUTION OF NEBULÆ.

Edwin Hubble.

1. In *M.N.*, 97, 1937 January, appear two rediscussions, by Sir Arthur Eddington and by G. C. McVittie, of results of an analysis of nebular distribution recently published by the writer.* Five surveys to apparent magnitudes $m = 18.47$ to 21.03 furnished a relation $N_m = f(m, \Delta m)$, where N_m is the number of nebulæ per square degree brighter than m , and Δm is the average correction for the dimming effects of red shifts on nebulæ of magnitude m . The true, or corrected, distribution was found to be so nearly uniform, on any interpretation of red shifts, that the investigation could be reversed. Assuming true uniformity and referring apparent departures to effects of red shifts, an attempt was made to distinguish between the alternative, possible interpretations of red shifts as Doppler effects or not Doppler effects.

The problem may be stated in another way. A rapidly receding object will appear fainter than a stationary object at the same momentary distance. Consequently, apparent luminosities of nebulæ offer different scales of distance depending upon whether or not the recession factor is involved. It seemed possible that, when the laws of red shifts and of nebular distribution were formulated on the alternative scales, the wrong scale might introduce anomalies which could be detected or at least suspected. The possibility was partially realized, for apparent anomalies were found in the scale involving recession factors, which could be eliminated only by a forced interpretation of the data.

2. The numerical results of the investigation were as follows: The law of nebular distribution, regardless of interpretation of red shifts, is well represented by the relation

$$\log_{10} N_m = 0.6(m - \Delta m) - 9.052 \pm 0.013, \quad (1)$$

where †

$$\log_{10} \Delta m = 0.2(m - \Delta m) - 4.239 \pm 0.024. \quad (2)$$

* "Effects of Red Shifts on the Distribution of Nebulæ," *Mt. Wilson Contr.*, No. 557; *Ap. J.*, 84, 517, 1937.

† The probable error in (2) was incorrectly published as ± 0.008 , which refers to a single regression curve, assuming a particular relation.

If we represent Δm by $B \cdot d\lambda/\lambda$, then B is about 3.0 if red shifts are not velocity-shifts, or about 4.0 if red shifts are velocity-shifts.

Case I. Assume that red shifts are not velocity-shifts. The law of red shifts, for isolated nebulae, is then *

$$\log_{10} d\lambda/\lambda = 0.2(m - \Delta m) - 4.707 \pm 0.016. \quad (3)$$

Combining (2) and (3),

$$\log_{10} \Delta m = \log_{10} d\lambda/\lambda + 0.468 \pm 0.029 \quad (4)$$

and

$$\Delta m = (2.94 \pm 0.22)d\lambda/\lambda. \quad (5)$$

Thus the assumption leads to a consistent picture of the observable region. The law of red shifts is linear, and the distribution is uniform over the complete range of the observations. There is no evidence of spatial curvature. The sample is homogeneous, and the universe presumably extends indefinitely beyond the limits of our telescopes.

Case II. Assume that red shifts are velocity-shifts ($B = 4$). The expected N, m relation, derived from the theory of expanding universes which obey the relativistic laws of gravitation, is †

$$\log_{10} N_m = 0.6(m - \Delta m) + F + \text{const.}, \quad (6)$$

where F is a term involving spatial curvature. The empirical law of red shifts may be written

$$d\lambda/\lambda = 5.37 \times 10^{-10}r + 2.54 \times 10^{-19}r^2, \quad (7)$$

where r is expressed in light years.

Since B is already given, and r is known in terms of $(m - \Delta m)$, these equations furnish a value of R_0 , the present radius of curvature. Moreover, since the coefficients in (7) represent the first- and second-time derivatives of R_0 , it is possible to determine the parameters of a relativistic expanding universe that is consistent with the observations. The model is closed, small and dense—"an ever-expanding universe of the first kind" in which

$$k = +1, \quad R_0 = 4.7 \times 10^8 \text{ lt. yr.}, \quad \rho = \text{about } 6 \times 10^{-27} \text{ gm./cm.}^3$$

and

$$\Lambda = +(4.4 \text{ to } 6.9) \times 10^{-18} \text{ yr.}^{-2}.$$

3. Eddington makes the criticism that, in deriving equations (1) to (5), the possible effects of dispersion in the absolute magnitudes of nebulae

* *Mt. Wilson Contr.*, No. 549; *Ap. J.*, **84**, 270, 1936. The probable error in equation (3) was not given, but it is readily calculated from the data listed in Table II.

† "Two Methods of Investigating the Nature of the Nebular Red Shifts," Hubble and Tolman, *Mt. Wilson Contr.*, No. 527; *Ap. J.*, **82**, 302, 1935. This paper adapted the general formulæ previously developed by Tolman for application to the observational data. A preliminary statement of the present status of the observational work was included, which was developed in detail in the later contribution now under discussion.

have been neglected. He then investigates the effects of dispersion by a very neat, original method, and finds that the necessary correction, $B = 3.06$ instead of 2.94 , "proves to be unimportant." This independent confirmation of the analysis of the data is most welcome.

However, it may be pointed out that effects of dispersion had already been examined. A rigorous expression for N_m , taking full account of dispersion in M , was developed in the above-mentioned discussion of methods and principles by R. C. Tolman and the writer.* Actual solutions of the equation by numerical integration indicated "that on account of the fairly sharp maximum exhibited by the exponential factor $e^{-\frac{(M-\bar{M})^2}{2\sigma^2}}$, the integral . . . can often be satisfactorily approximated by assigning to the rest of the integrand the value that it has at the maximum point where $M = \bar{M}$." In other words, the rather complicated, rigorous method could be replaced by the very simple method actually used, without introducing significant errors. It is this statement that has been checked by Eddington's independent investigation.

4. His second and more serious criticism is that, although the value of B indicated by the data is approximately correct, it is not sufficiently reliable to constitute a discordance with the expected value 4.0 . He states that the discordance (3.06 against 4.0) is equivalent to a discordance † of 0.066 mag. in the range of the measured m . For, he adds, let δm be this range (2.56 mag.). From δm and the assumption of uniform distribution we derive $\delta(m - \Delta m)$ and, consequently, $\delta(\Delta m)$, from which B is determined. "An error of 0.01 mag. in δm will produce an equal error in $\delta(\Delta m)$ and an error of 0.13 in k " (where k , in his notation, is B in the present discussion).

But, by differentiating (2),

$$d(\Delta m) = \frac{\Delta m}{2.171 + \Delta m} \cdot dm, \quad (8)$$

and an error in m evidently produces a much smaller error in Δm . Errors in δm , of course, produce comparable errors in $\delta(m - \Delta m)$, which are reflected in the constants of equations (1) and (2), and, ultimately, in B . The calculation is delicate, and the probable error in (5), of the order of $1/5$ the discordance, does not inspire complete confidence. To this extent the criticism is justified.

The uncertainty arises from the fact that, over the short range δm , the constants can be adjusted with considerable latitude. A material increase in δm improves the situation, and, as it happens, the only other published survey, the Harvard catalogue, complete to $m = 12.8$, and with $\log \bar{N} = -1.38$, meets this requirement. When these data are included in the least-squares solution, the constants in (1) and (2) are essentially unchanged but the weights are considerably increased. The coefficient is

* *Mt. Wilson Contr.*, No. 527; *Ap. J.*, **82**, 302, 1935.

† By the writer's method of analysis, least-squares solution of the data in all the surveys, an error in δm of the order of 0.12 mag., evenly divided between the two extreme limits, is required to furnish $B = 4.0$.

then $B = 2.92 \pm 0.16$ which, in a formal sense, is significant. Because of the small numbers of nebulæ involved, the propriety of including the Harvard list on the same basis as the deeper surveys may be questioned. They are mentioned merely to emphasize the point that the only additional information now available tends to improve the reliability of the original results.

However, the measure of confidence to be placed in equation (5) does not depend primarily on the probable errors. The major uncertainties must be sought in the possibilities of systematic errors, and there is found the real significance of the conclusions. When it became evident that B would fall below the expected value, 4.0, the observations were reduced by methods which, at each step, favoured a large value. The methods were described, and it was specifically mentioned that the "examination of possible sources of systematic errors strongly suggest that the data . . . exhibit a maximum range in m and a minimum range in $\log \bar{N}$. The conclusion has an important bearing on the interpretation of the data."

As an example, consider the determination of δm . A zero-point was established at 19.11 ± 0.02 on the scale of the polar sequence, for standard exposures of 20 min. with the 100-inch, from measures of *schraffierkassette* images. The range, δm , is not very sensitive to errors in the zero-point, and precision was sought primarily for the purpose of comparing (2) and (3). The brightest limit, 18.47 ± 0.03 , for 20 min. exposures with the 60-inch, was determined from the same plates that furnished the zero-point, and, consequently, the uncertainty in that section of the range, aside from the small accidental errors, is the presumably negligible scale-error of the polar sequence over the short interval of 0.64 mag. The faintest limit, 21.03 ± 0.03 , for 120 min. exposures with the 100-inch, corresponds with an extrapolation from the standard 20 min. exposures, using a reciprocity law with an exponent of 0.99.

This section of the range, 1.92 mag., is clearly more unreliable than the short, brighter section, and the uncertainties operate in one direction only, namely, to reduce the interval. In order to appreciably increase δm , it would be necessary to postulate an exponent greater than unity in the reciprocity law; in other words, to assume that, when the 20 min. exposures are lengthened six times, nebulæ are recorded that are *more than* six times fainter. Under the particular conditions, faint objects and Eastman 40 plates, the assumption would be wholly unreasonable. The adopted range is clearly a maximum, and any significant revision would be expected to reduce the interval. The other observable quantity, N_m , and the calculation of the expected values of B , were treated from the same point of view.

The interpretation of red shifts by the theory of expanding universes is so plausible and so widely current that, in making a delicate test of the theory, it is desirable to push uncertainties in the favourable direction before admitting a discordance. Nevertheless—and this is perhaps the significant result of the investigation—when the observational data are weighted in favour of the theory as heavily as can reasonably be allowed, they still fall short of expectations.

With the theory in its present form, and with the observations now available, the discordance exists when spatial curvature is neglected.*

5. McVittie is concerned only with the determination of parameters in Case II. He attempts to show that the observational data are consistent with an infinite universe with negative curvature rather than a finite universe with positive curvature. The point of departure is the well-known pair of equations,

$$\kappa\rho = -\Lambda + 3(\dot{R}/R)^2 + 3c^2k/R^2, \quad (9)$$

$$\kappa p = \Lambda - (\dot{R}/R)^2 - 2\ddot{R}/R - c^2k/R^2. \quad (10)$$

Making the usual assumption that the pressure p is negligible, and eliminating Λ , the result is

$$2c^2k/R_0^2 = \kappa\rho - 2(\dot{R}_0/R_0)^2 + 2\ddot{R}_0/R_0, \quad (11)$$

where the subscript zero is used to indicate the present epoch.

Now the derivatives, \dot{R}_0/R_0 and \ddot{R}_0/R_0 , are determined, at least as far as their signs and orders of magnitude are concerned, from the observational data. The writer, following Tolman,† calculates them (or their equivalents) directly from the coefficients in the law of red shifts represented by (7). McVittie adopts an approximate value of \dot{R}_0/R_0 corresponding with the first-order term in (7), but derives \ddot{R}_0/R_0 by another, rather arbitrary method that will be discussed presently. However, the two sets of values are of the same general order. Thus

	\dot{R}_0/R_0	\ddot{R}_0/R_0
Hubble	1.70×10^{-17}	-2.84×10^{-34} .
McVittie	2×10^{-17}	-7×10^{-34} .

Using the first set, and remembering that $\kappa = 1.674 \times 10^{-6}$, equation (11) becomes

$$1.075 \times 10^{27}k/R_0^2 = \rho - 6.85 \times 10^{-28}. \quad (12)$$

Thus k , which can be $+1$, 0 or -1 , depends upon whether or not the density ρ is greater than, equal to or less than the critical value 6.85×10^{-28} . If ρ is sensibly less than the critical value, k will be negative and R_0 will be of the order of 10^9 light years, as McVittie concludes. On the other hand, if the curvature is positive, ρ must be greater than the critical value, and,

* The statement is believed to be properly worded. The observations, of course, may involve errors sufficient to vitiate the conclusions, but, if so, the trouble will probably be found in hidden systematic errors. Further investigations of two kinds are desirable. Extensive surveys to bright limits, where the dimming effects are negligible, will tie down the constant in equation (1) and, consequently, restrict the possible adjustments of the constant in (2). Surveys to still deeper limits (which will be possible with the 200-inch reflector), where the dimming effects reach a magnitude and more, should then establish the value of B with a significance that is beyond controversy.

† The writer's estimate of parameters are essentially a revision, suggested by new data, of the earlier determinations made by Tolman, *Relativity, Thermodynamics and Cosmology*, p. 474, 1934.

as will appear later, greater than can be accounted for without introducing large quantities of internebular material. This latter situation is one of the anomalies which results from analysing the observational data on the basis of a distance scale involving recession factors.

McVittie approaches the problem by calculating ρ from the luminous nebulae alone, determining $\bar{\alpha}$, the number of nebulae per unit volume, and using Smith's value of the mean mass, $2 \times 10^{11} \odot$. By an oversight he has compared coefficients in two equations, one referring to nebulae per square degree, and the other to nebulae over the entire sky. Consequently, his result must be increased by a factor of the order of 4.1×10^4 . Moreover, when the Sun's mass is taken as 2×10^{33} gm. rather than the round number 10^{33} , McVittie's value of the density (8.2×10^{-33}) must be revised to

$$\begin{aligned} \rho &= 8.2 \times 10^{-33} \times 8.2 \times 10^4 \\ &= 6.7 \times 10^{-28}, \end{aligned} \quad (13)$$

which is so near the critical density that the particular methods of approximation employed in the calculation become all-important.

In the writer's opinion, the most reliable value of α , the number of nebulae per unit volume, is derived from equation (1), together with the expression for distance in terms of $(m - \Delta m)$ and \bar{M} , the mean absolute magnitude of nebulae of a given m . Using the most recent value, $\bar{M} = -15.15$, and Smith's value for the mean mass, the density is

$$\begin{aligned} \rho &= 2.42 \times 10^{-73} \times 4 \times 10^{44} \\ &= 10^{-28}. \end{aligned} \quad (14)$$

Since Smith's result is considered as an upper limit, the density appears to be well below the critical value, and the anomaly persists—the derived curvature is negative and the radius is of the order of 10^9 light years.

The result is considered as an anomaly because, if the curvature were negative, then, after the corrections Δm are applied, the apparent distribution-density should fade out with distance, whereas, actually, it increases with distance. It was for this reason that the writer adopted the positive curvature necessary to compensate the apparently increasing numbers of nebulae (in other words, to compensate the recession factor, without which the distribution appears to be uniform), and then derived the corresponding density (about 60 times the critical density).

6. The investigator must necessarily follow this course or else reject the apparently increasing α as errors in the observations or their interpretation. It seems highly improbable that the mean mass of nebulae can be increased, or the mean absolute magnitude reduced, by amounts sufficient to remove the discrepancy. Consequently it is necessary, in order to preserve the theory intact, either to introduce large quantities of internebular material, or to question the validity of the corrections Δm as calculated from the relation, $\Delta m = 4d\lambda/\lambda$.

McVittie follows the latter course, and analyses the data by a method which seems to evade the difficulty at the cost of a rather arbitrary procedure.

He derived an N, δ relation (δ in his notation is the red shift $d\lambda/\lambda$) as follows:—

$$\log_{10} N = 0.501m - 7.371, \quad (15)$$

$$\log_{10} \delta = 0.2m - 4.967, \quad (16)$$

$$N = 1.113 \times 10^{5\delta^{5/2}} \quad (17)$$

and

$$dN/d\delta = 2.78 \times 10^{5\delta^{3/2}}. \quad (18)$$

Equation (15) is an interpolation formula, using the observed, uncorrected magnitudes, which satisfies the surveys well enough over the observed range in m , but which breaks down for brighter limits (for instance, the Harvard survey). Equation (16) represents an early formulation of the law of red shifts, in which the dispersion in M was not taken into account (the most recent value of the constant is about -4.707). Corrections Δm were applied but, since the distances were small, they may be neglected.

Equations (15) and (16) may be considered as linear interpolation formulæ, representing certain limited ranges of general non-linear relations. Near the observer the general relations are believed to be sensibly linear, the coefficients of m being 0.6 and 0.2 respectively. With increasing distance (increasing m) the slopes steadily diminish (due to the omission of the dimming effects, Δm). Consequently, the exponent of δ in (17), which is the ratio of the coefficients in (15) and (16), is arbitrary because it depends upon the particular ranges of the general functions which are selected for comparison.

The same ranges would presumably lead to an exponent, very nearly 3. But McVittie has compared a very remote region in (15) with a nearby region in (16), and necessarily derives a smaller exponent in (17) which happens to be almost precisely $5/2$.

He wishes to compare coefficients in (18) with those in McCrea's theoretical expression for $dN/d\delta$. But the latter is a power series in δ , the first term of which is of the second order. Therefore, McVittie takes the further arbitrary step of expanding $\delta^{3/2}$ as a power series, also beginning with the second-order term. The step, in a sense, amounts to a rough approximation toward comparing the original functions in the same range. In that case the exponent, 2, would presumably appear as a matter of course, just as it appears when the Δm are included in the general relations. It is partly for this reason that the values of α and \ddot{R}_0/R_0 , which are derived from the coefficients of δ^2 and δ^3 , are of the same general order as those derived with the use of Δm , except for the oversight * in the case of α .

7. The advantages of an empirical N, δ relation have long been recognized, provided it extends out to large values of δ . Direct formulations, of course, are unpractical, and the comparisons of the N, m with the

* In equation (15) N refers to nebulæ per square degree, while in McCrea's formula N refers to nebulæ over the entire sky. Consequently, equation (17) should read $N = 4.59 \times 10^9 \delta^{5/2}$, and (18), $dN/d\delta = 1.148 \times 10^{10} \delta^{3/2}$, in order to make the comparison with McCrea's formula.