# Remarks on the Redshift from Nebulae

F. ZWICKY, *California Institute of Technology*
(Received May 20, 1935)

Some critical considerations are advanced, concerning the theory of the expansion of the universe. The main purpose of this paper is to indicate how a *statistical theory* can be developed which makes it possible to discuss in a very general way a number of features of the redshift of light through intergalactic space. In particular the problem of the width of spectral lines from distant nebulae is treated. Finally some observational tests are proposed which promise to throw new light on the problem of the redshift.

## 1. INTRODUCTION

IN recent years numerous attempts have been made to treat theoretically the problem of the redshift of light from extragalactic nebulae. Best known are the various theories which have sought to interpret this redshift on the basis of expanding models of the universe treated with the help of relativistic mechanics. Many other proposals, however, have been advanced. At the present time the relativistic theories are perhaps the only ones which derive their justification more or less directly from a set of equations whose validity is established to a high degree in other fields of physics and astronomy. In spite of this alluring feature the theory of the expanding universe can hardly be considered as a completely satisfactory solution of the problem of the redshift. Some of the following objections might be raised.

(a) The fundamental equations of the general theory of relativity are admittedly incomplete, inasmuch as they do not organically include the electromagnetic field and the phenomena related to the quantum of action. A *unified* field theory may well lead to conceptions about time and space which are fundamentally different from our present notions. For instance, the now abandoned theory of distant parallelism did not even allow of isotropy of space. The assumption of isotropy, however, constitutes an essential part of the current theory of the expanding universe.

(b) Even if we admit the validity of the equations of the ordinary theory of general relativity for limited regions of space and time, it is by no means certain whether these equations can be applied to the universe as a whole without taking into account retardation effects caused by the finiteness of the speed of light. Gravitation, according to the general theory of relativity, is a *cooperative phenomenon*,[1] inasmuch as all physical actions related to gravitation and inertia in a given point are not wholly determined by any limited regions which surround this point. The influence of even the most distant masses is indeed essential. In a universe, populated by moving masses, it is therefore not permissible to assume homogeneity without proving that the forces which are available in the universe are capable of maintaining homogeneity.[2] Stationary solutions of the gravitational equations of general relativity can give reliable results only in regions whose linear dimensions $L$ are small, compared with a certain critical distance $d$, which is

$$d=cT, \tag{1}$$

where $c$ denotes the velocity of light and $T$ is equal to the average time in which are completed radical local rearrangements of the biggest individual stable agglomerations (systems) of matter. If we may identify these systems with the individual nebulae, then

$$T=\Lambda/\bar{v}, \tag{2}$$

where $\Lambda$ is the "mean free path" of a nebula and $\bar{v}$ its average velocity. From observations on extragalactic nebulae[3] we know that approximately $\Lambda=10^6$ L.Y. (light years) and $\bar{v}=1000$ km/sec. The regions of space to which stationary solutions can safely be applied therefore have

[1] On the classification of cooperative phenomena see F. Zwicky, Phys. Rev. **43**, 270 (1933).
[2] See also R. C. Tolman, *Relativity, Thermodynamics & Cosmology* (Oxford, 1934), p. 483.
[3] F. Zwicky, Helv. Phys. Acta **6**, 110 (1933).

linear dimensions $L$ which must satisfy the inequality

$$L \ll cT = 3\times 10^8 \text{ L.Y.} \qquad (3)$$

On all of the present cosmological theories the diameter $D$ of the universe is

$$D > 2\times 10^9 \text{ L.Y.} \qquad (4)$$

The inequality (4) is also borne out by direct observations. The faintest nebulae which can be reached with the 100″-telescope are probably at distances of the order of $10^9$ L.Y. All of the present relativistic stability considerations are based on stationary solutions of Einstein's equations and cannot be considered as conclusive.

Another serious difficulty is related to the fact, that in all of the recently advanced relativistic "models" of the universe use is made of notions, such as homogeneity, transfer of rigid meter sticks, etc., which have never been rigorously defined by actual physical operations.

(c) E. Hubble[4] has shown from his counts of nebulae, that a uniform distribution of these objects in a flat space is obtained if one corrects their apparent brightness for the loss of energy caused by the observed redshift. If, on the other hand, one assumes that the universe is expanding, the uniformity of the distribution of nebulae follows only if some additional assumption is introduced to compensate the effects of the general expansion, e.g., curvature of space, etc. As long as no further observations are available, it is therefore scientifically more economical *not* to link the redshift from nebulae with any *purely hypothetical* curvature and expansion of space.

(d) The dispersion of apparent velocities in clusters of nebulae constitutes a considerable part of the average apparent velocity of recession of these clusters. The coma cluster, for example, has a redshift of approximately 7000 km/sec. with a scattering of about 3000 km/sec. between its individual member nebulae. On the theory of the expanding universe no explanation has been advanced for this very important fact.

For the above reasons cautiousness requires not to interpret too dogmatically the observed redshifts as caused by an actual expansion. More experimental information is badly needed before we can hope to arrive at a satisfactory theory. The following lines contain an attempt to guide the theoretical and experimental researches on the subject in question in new directions.

[4] E. Hubble, *The Halley Lecture* (Oxford, May, 1934).

## 2. Probability Considerations Related to the Redshift from Nebulae

We consider the following problem. A light quantum $h\nu_0$ is emitted from a nebula $N$ located at a distance $D$ from the observer $O$. Suppose that on the path $NO$ there are $n$ equidistant points or *stations*, which act as obstacles in such manner that the quantum $h\nu_0$, with a probability $p$, may lose an increment $h\delta$ at any station. For simplicity we assume that $p$ and $\delta$ are independent of the frequency $\nu$, inasmuch as we consider only small changes of $\nu_0$. The probability for the quantum $h\nu_0$ to arrive at $O$ with the reduced energy $h(\nu_0 - i\delta)$ is $I_i^{(n)}$, where

$$I_i^{(n)} = \binom{n}{i} p^i (1-p)^{n-i}. \qquad (5)$$

We shall call $I_i^{(n)}$ the intensity factor of the frequency $\nu_0 - i\delta$. The relation (5) is obtained from the following obvious recurrency relation for the intensity factors

$$\mathrm{I}_i^{(n)} = pI_{i-1}^{(n-1)} + (1-p)I_i^{(n-1)} \qquad (6)$$

in conjunction with the theorem (7) for the binomial coefficients

$$\binom{n}{i} = \binom{n-1}{i-1} + \binom{n-1}{i}. \qquad (7)$$

Monochromatic light of the frequency $\nu_0$ from the nebula $N$ therefore arrives at $O$ as a spectrum of frequencies $\nu_0,\ \nu_0-\delta,\ \cdots\ \nu_0 - i\delta,\ \cdots\ \nu_0 - n\delta$, with the intensity factors $I_0^{(n)},\ I_1^{(n)},\ \cdots\ I_i^{(n)},\ \cdots\ I_n^{(n)}$. As we have assumed that our quanta are not deflected from the straight path, all of them arrive at $O$, with more or less reduced energies, and the sum of the intensity factors must be

$$\sum_{i=1}^{n} \mathrm{I}_i^{(n)} = 1. \qquad (8)$$

This condition is satisfied, because

$$\sum_i I_i^{(n)} = \sum_i \binom{n}{i} p^i (1-p)^{n-i}$$
$$= [(1-p)+p]^n = 1. \quad (9)$$

## 3. Values of the Intensity Factors for a Large Number of Stations

In discussing the spectral distribution of the intensity factors, two cases A and B must be distinguished.

### Case A

Monochromatic light of frequency $\nu_0$, emitted at $N$, arrives at $O$ as a spectrum, whose maximum intensity is still at $\nu_0$, or, $I_{i+1}^{(n)} < I_i^{(n)}$ and in particular $I_1^{(n)} < I_0^{(n)}$. The condition for this to be true is

$$np/(1-p) < 1. \quad (10)$$

I shall not discuss this case any further, as it does not correspond to a shift of the maximum of the spectral lines and therefore has little significance for the theory of the redshift from nebulae.

### Case B

The maximum intensity of a spectral line is shifted to a lower frequency when light travels from $N$ to $O$. This implies that at least $I_1^{(n)} > I_0^{(n)}$ or

$$np/(1-p) > 1. \quad (11)$$

Three simple subcases $B_1$, $B_2$ and $B_3$ may be distinguished, which we now proceed to analyze.

*Subcase* $B_1$. We assume that $p = 1$, independent of how many stations we choose on the path $NO$. We also put $\delta = \Delta/n$ where $\Delta$ is a function of $D$ and $\nu$ only, but is independent of $n$.

The interpretation of redshift as caused by an expanding universe falls into this category. We evidently have

$$I_0 = I_1 = \cdots = I_{n-1} = 0 \qquad I_n = 1. \quad (12)$$

A spectral line on its way from $N$ to $O$ is shifted "bodily" from $\nu_0$ to $\nu_0 - \Delta$. There is no widening of the line.

The remaining two cases deal with probabilities $p < 1$. For reasons which will become apparent later on, we are particularly interested to find the *displacement* and the *widening* of spectral lines for large values of $n$.

*Subcase* $B_2$. We assume, $p = \text{constant} < 1$, and simultaneously $\lim n = \infty$, $\lim_{n=\infty} \delta = 0$ in such a manner that always $n\delta = \Delta\nu(D)$. This means that at *every* station the incoming frequency can lose an increment $\delta$. As we increase the number of stations $n$ on a given path $NO = D$ the probability $p$ for a change of frequency to take place at any individual station stays constant, but the magnitude of the change decreases. Every station has a real physical significance. In real physical problems $n$ will of course never be infinitely large, so that our calculations for $n = \infty$ will give asymptotic representations of the actual phenomena.

The proposal which I made a few years ago,[5] that a "gravitational drag" on light is responsible for the redshift from nebulae, falls into this category $B_2$. A gravitational drag arises in this manner. We consider a stationary gravitational system, in our case the universe, which contains matter and radiation. Suppose that a given body $C$ of this system suddenly divides into two parts $C_1$ and $C_2$. In our case $C_1$ is, for instance, a star and $C_2$ is a light quantum $h\nu$ which is emitted from it and which sooner or later is absorbed as a quantum $h\nu'$ by some other body $C'$. If now all of the bodies $C$, $C'$, $C''$, $\cdots$, were immovable relative to one another, then according to the theory of relativity

$$(\nu - \nu')/\nu = (\Phi' - \Phi)/c^2, \quad (13)$$

where $\Phi$ and $\Phi'$ are the gravitational potentials on $C$ and $C'$, respectively. If, however, all the bodies are freely movable under their mutual gravitational actions, the rearrangement of matter caused by the transfer of the mass $h\nu/c^2$ from $C$ to $C'$ involves in general an additional loss of energy for the quantum $h\nu$. The magnitude of this loss is essentially determined by retardation effects which are not included in formula (13).

Mathematically, subcase $B_2$ is closely related to a problem of probabilities first treated by de Moivre. It is seen at once that the terminal intensity factors $I_0^{(n)} = (1-p)^n$ and $I_n^{(n)} = p^n$, for $n = \infty$ tend towards zero faster than $1/n$. The intensity factors therefore possess a maxi-

[5] F. Zwicky, Proc. Nat. Acad. Sci. **15**, 773 (1929); Phys. Rev. **34**, 1623 (1929).

mum $I_i$ between $I_0$ and $I_n$ at some index $i=\alpha n$. We proceed to determine $\alpha$, assuming that $n$, $\alpha n$ and $n(1-\alpha)$ are simultaneously large numbers. We make use of Stirling's asymptotic formula for the factorials

$$\lg x! = (x+\tfrac{1}{2}) \lg x - x + \tfrac{1}{2} \lg 2\pi + R(x), \quad (14)$$

where $R(x) < 1/12x$. We shall neglect the term $R(x)$. In this approximation we obtain as characteristic for the maximum

$$\alpha = p - (1-2p)/2n. \quad (15)$$

We may neglect $(1-2p)/2n$ as compared with $p$ and expand $I$, approximating it by a Gaussian error curve. Putting $i = pn + \xi$ we finally have

$$I_{pn+\xi} = e^{-\xi^2/2np(1-p)}/[2\pi np(1-p)]^{\frac{1}{2}}. \quad (16)$$

The sum, taken over all of the intensity factors is unity, as it should be.

$$\sum_i I_i^{(n)} = \int_{-\infty}^{+\infty} I d\xi = 1. \quad (17)$$

From (16) we see that the maximum intensity factor has a value which is proportional to $n^{-\frac{1}{2}}$. At the same time the characteristic width of the "spectrum" of the intensity factors comprises the range of indices from $pn-\xi_c$ to $pn+\xi_c$, where $\xi_c \propto n^{\frac{1}{2}}$. The intensity factor is indeed reduced to the value $I_{\max}/e$ for

$$\xi_c = [2np(1-p)]^{\frac{1}{2}}. \quad (18)$$

Our initially monochromatic beam of frequency $\nu_0$ arrives at $O$ as a spectrum whose intensity maximum lies at the frequency $\nu_m = \nu_0 - p\Delta\nu$ and whose characteristic width $W$ is

$$W = 2\xi_c\delta = [8p(1-p)/n]^{\frac{1}{2}}\Delta\nu, \quad (19)$$

whereas the shift $S$ of the spectral line is

$$S = \nu_0 - \nu_m = p\Delta\nu, \quad (20)$$

the ratio of width to shift is

$$W/S = [8(1-p)/np]^{\frac{1}{2}} \propto n^{-\frac{1}{2}}. \quad (21)$$

As the number of stations $n$ on the path $D = NO$ increases, the ratio $W/S$ decreases proportional to $n^{-\frac{1}{2}}$. Hence, for great distances $D$ the sharpness of the initial spectral line is preserved.

If the redshift has a cause whose statistical interpretation is covered by our subcase $B_2$ no appreciable widening of the spectral line can be expected.

*Subcase $B_3$.* We have simultaneously $\lim n = \infty$, $\lim_{n=\infty} p = 0$, in such a manner that always $np = P(D)$, where $P > 1$. $P$ depends only on $D$, but is independent of $n$. At the same time $\delta$ = constant, independent of $n$. This case corresponds to a problem of probabilities first treated by Poisson. It embraces such interpretations of the redshift which postulate the presence in space of a definite average number of obstacles which are distributed at random over the $n$ stations on the path $D$. Changes of the frequency, caused by "direct contact" of light with electrons and atoms, such as the Compton effect, Raman effect, etc., may be classified under subcase $B_3$.

The distribution of the intensity factors is immediately obtained from our previous result (16) by putting $np = P$. The maximum intensity factor $I_i$ has the index $i = P$ which is independent of the number $n$ of stations chosen. We have

$$I_{p+\xi} = e^{-\xi^2/2P(1-P/n)}/[2\pi P(1-P/n)]^{\frac{1}{2}} \quad (22)$$

or, for $n = \infty$ $\quad I_{p+\xi} = e^{-\xi^2/2P}/(2\pi P)^{\frac{1}{2}}. \quad (23)$

This result is also obtained immediately by a direct derivation of the intensity factor $I_i$ for $n = \infty$, namely,

$$\lim_{n=\infty} I_i = \lim_{n=\infty} \frac{n!}{i!(n-i)!}(P/n)^i(1-P/n)^{n-i}$$
$$= P^i e^{-P}/i!. \quad (24)$$

Putting $i = P + \xi$ and assuming that $P$ is a fairly large number we have asymptotically the approximation (23) for the intensity factors. The shift of the spectral line in this case becomes

$$S = P\delta \quad (25)$$

and the width of the resulting line

$$W = 2\xi_c\delta = 2(2P)^{\frac{1}{2}}\delta. \quad (26)$$

The ratio of width to shift is

$$W/S = (8/P)^{\frac{1}{2}} \quad (27)$$

tends toward zero as $D$ and $P$ increase. For great distances again there results no widening of the spectral lines.

All of those interpretations of the redshift, which with sufficient accuracy can be described by our simple statistical schemes will therefore not be associated with any appreciable widening of the spectrum of light which travels over long distances. In actuality no widening of spectral lines from more and more distant nebulae has so far been observed.

## 4. Generalizations of the Statistical Theory of the Redshift

The theory which I have presented in this paper of course does not give more than the most elementary ground work for a statistical interpretation of the redshift. The following generalizations might be considered:

(1) The probability $p$ depends on the frequency $\nu$ and on the index $i$ of the station.

(2) The change $\delta$ of the frequency $\nu$ depends on $\nu$ and $i$.

(3) The change $\delta$ at a given station is not unambiguous. The possible changes $\delta_{i1}$, $\delta_{i2}$, $\cdots$ occur with the probabilities $p_{i1}$, $p_{i2}$, $\cdots$.

(4) Finally a three-dimensional system of stations, which are distributed over the entire space, must be considered in order to incorporate such interpretations of the redshift which involve scattering of light.

Professor H. Bateman at our Institute has been kind enough to take up some of these problems. An account of his researches will appear in this journal.

## 5. Concluding Remarks

From the above it becomes apparent that considerations concerning the *spectral spreading* of light traveling over long distances are not likely to provide any definitely distinguishing tests between various interpretations of the redshift, even if they are as radically different from one another as was assumed in the cases $B_1$, $B_2$ and $B_3$.

If one looks for other effects which will furnish more decisive clues about the correct theory of the redshift, a consideration of the *spatial spreading* or scattering of light immediately suggests itself. Qualitatively the following conclusions are obtained.

All of the proposals which seek to interpret the redshift as being caused by obscuring matter, such as electrons, atoms, etc., must be discarded. Redshifts which result from the Compton effect, the Raman effect, and so on, are associated with a scattering of light which would produce impossibly large washing out of nebular images which actually are not observed.

The expansion of the universe, on the other hand, calls for no changes of nebular images except such as are caused by purely geometrical conditions which determine the path of light. The images therefore are not washed out.

The theory of the gravitational drag occupies a middle position inasmuch as different light quanta from the same nebula go through almost but not exactly identical changes on their path from the nebula to the observer. An initially parallel beam of light, on this theory, will gradually open itself because of small angle scattering. Observational tests on this point will be important.

In conclusion I mention in a preliminary fashion one more possibility to gain new information about the redshift from nebulae. From Hubble's counts of nebulae it follows, that in our "immediate" neighborhood of about 50 million light years about one percent of all nebulae are members of giant clusters. The question suggests itself whether this ratio between cluster nebulae and field nebulae is independent of distance. The mode of formation and duration of clusters of nebulae must be closely associated to the physical history of the universe as a whole. Observations of the relative number of nebulae in clusters promise to provide important material for further theories. It will also be essential to find out if the dispersion of velocities among the member nebulae of cluster is a function of distance. The proposal to test a possible relation between distance and dispersion of velocities has been advanced by Mt. Wilson astronomers in their research program.