# THE LAW OF RED-SHIFTS

*George Darwin Lecture*, delivered by Dr Edwin Hubble on* 1953 *May* 8

I propose to discuss the law of red-shifts—the correlation between distances of nebulae and displacements in their spectra. It is one of the two known characteristics of the sample of the universe that can be explored and it seems to concern the behaviour of the universe as a whole. For this reason it is important that the law be formulated as an empirical relation between observed data out to the limits of the greatest telescope. Then, as precision increases, the array of possible interpretations permitted by uncertainties in the observations will be correspondingly reduced. Ultimately, when a definite formulation has been achieved, free from systematic errors and with reasonably small probable errors, the number of competing interpretations will be reduced to a minimum.

The path towards such a definitive formulation is now clear and the investigations are under way at Mount Wilson and Palomar. I shall discuss this programme in some detail after sketching the history of the law of red-shifts as a convenient background.

There are three phases in the history, namely, the discovery phase which ended with a crude formulation in 1928–29, the rapid extension and improved formulation out to the limit of the 100-inch in 1929–36, and the current attempts to reach the limit of the 200-inch with the definite formulation, which began about two years ago.

## I. The First Phase (Discovery)

### A. *Radial Velocities*

The discovery emerged from a combination of radial velocities measured by Slipher at Flagstaff with distances derived at Mount Wilson. You will recall that the first velocity of a spiral nebula was measured by Slipher in 1912 ($-300$ km/sec for M 31). I mention the incident because the first step in a new field is the great step. Once it is taken, the way is clear and all may follow. In this case, however, Slipher worked almost alone, and ten years later, when he turned to other problems, he had contributed 42 out of the 46 nebular velocities then available. The list was completely dominated by large positive velocities, ranging up to $+1800$ km/sec for NGC 584.

Attempts were made, from 1916 onward, to derive the solar motion with respect to the nebulae; but the positive signs, indicating general recession, proved intractable. Then, in 1918, Wirtz attempted to " save the phenomena " by introducing a constant $K$ term in the equations. The device improved the results somewhat but not enough to render them acceptable. Nevertheless,

* *Editorial Note.*—This paper comprises the text of the George Darwin Lecture, which the late author had intended to revise before publication. His notes, together with the manuscript from which he spoke, made it clear what form he wished the published material to take. A reorganization of the original manuscript according to his marginal notes, with the addition of a few connecting sentences, was the extent of the editing required.—A. R. Sandage.

it did suggest that the next logical step would be to replace the constant $K$ with $rK$, a term which varied with the distances of the individual nebulae. This procedure was suggested by the run of residuals in the earlier solutions but it was also encouraged by the rapidly developing studies of relativistic cosmology.

Various solutions were made in the early and middle twenties, using apparent luminosities and apparent diameters as measures of relative distance. They failed but, as later appeared, only because the distance criteria were not sufficiently accurate to give reliable individual distances over the relatively short range covered by the radial velocities. The uncertainties are evident in the triple system comprising M 31 and its two companions, M 32 and NGC 205, where luminosities ranged from 100 to 1, and diameters from 60 to 1.

### B. *Distances from Brightest Stars*

The problem was finally solved when the brightest stars in late-type spirals and irregular nebulae were established as suitable distance indicators, and were calibrated by Cepheids in a few of the nearest nebulae.

This new development started in 1923–24 when Cepheids were found in M 31 and M 33. The variables indicated the distances of the spirals, confirmed their suspected resolution and identified them as extragalactic stellar systems. Following up the clue, stars were soon recognized in many of the more conspicuous spirals, and the upper limits of brightness of these stars appeared to be substantially constant. It was this upper limit for the three or four brightest stars in great stellar systems which was used as a measure of distance. When, in 1928, it was applied to the nebulae with measured velocities, the law of red-shifts, or, as it was then called, the velocity-distance relation, emerged at once in approximately its present form. The law was established out as far as the Virgo Cluster, or about 7 million light-years, as an approximately linear relation, according to which velocities of recession increased at a rate of the order of 500 km/sec per million parsecs of distance.

## II. The Second Phase (Reconnaissance)

### A. *Humason's Adventures among the Clusters*

The second phase was entirely a Mount Wilson project, and lasted from 1928 to 1936. Humason assembled spectra of nebulae and I attempted to estimate their distances. Attention was concentrated on clusters because their brightest members could be regarded as distance indicators, and, at the same time, offered maximum distances for a given apparent magnitude. Furthermore, the populations of great clusters are dominated by elliptical systems, which are the most highly concentrated of all nebulae, and hence the most easily observed with spectrographs. Faint elliptical nebulae give semi-stellar images and, therefore, they respond to the increased light-gathering power of great telescopes.

Humason's adventures were spectacular. He first observed some of Slipher's nebulae, and then, when he was sure of his techniques, and confident of his results, he set forth. From cluster to cluster he marched with giant strides right out to the limit of the 100-inch. Spectrographs were improved from time to time—the names of Rayton and Schmidt and others are associated with the cameras—and each improvement was followed by a new surge forward. The limit was bracketed when apparent velocities of 40 000 km/sec were recorded in the Boötes Cluster and in Ursa Major No. II, and readable spectra could not be obtained in the fainter Hydra Cluster (now known to have shifts half again as large). And always, throughout the march, the data fell in line with the relation as originally formulated.

When clusters could not be observed to advantage, spectra of field nebulae, isolated or members of small groups, were assembled in large numbers. These data furnished an account of red-shifts in the nearer, more conspicuous nebulae of all types.

### B. *Results of the Second Phase*

When the second phase ended, the following results could be listed.

(1) The law of red-shifts appeared to be linear (within the uncertainties of the distance estimates) out to approximately 250 million light-years, and the red-shifts, when expressed as velocities ($c \, . \, d\lambda/\lambda$), increased at the rate of about 530 km/sec per million parsecs.

(2) The solar motion is largely accounted for by galactic rotation, the residual representing the peculiar motion of the stellar system being small and uncertain (order of 100–150 km/sec towards galactic latitude about $+30^\circ$ and longitude in the quadrant $50^\circ$ to $140^\circ$).

(3) The law of red-shifts does not operate within the Local Group.

(4) Magnitude-velocity relations, fitted to data for clusters, resolved nebulae and field nebulae, and, of course, displaced from one another by magnitude intervals indicating relative absolute luminosities, could be used to calibrate all the relations in terms of brightest stars. Assuming $M_{pg} = -6{\cdot}35$ for Brightest Stars, the corresponding values for Field Nebulae and Fifth Nebulae in Clusters were $-15{\cdot}20$ and $-16{\cdot}45$, respectively.

## III. Interim between Second and Third Phases

### A. *Lists of Velocities Assembled by Humason and Mayall*

Towards the end of the second phase, Mayall entered the field and began assembling spectra with a very fine ultra-violet spectrograph on the 36-inch Crossley at Mount Hamilton. He paid special attention to the later-type spirals and other systems of low surface brightness for which his reflector was fully as efficient as the larger telescopes on Mount Wilson. From 1935 onward, the data increased in volume until now he has a list of the order of 280 nebular velocities as compared with Humason's list of about 570. There are some 100 nebulae in common which ensure a reliable comparison of the two systems, and the two lists when published will include about 750 individual nebulae.

These data furnish a wealth of information concerning nebulae of all kinds. Although full discussions must await publication of the lists, it is permitted to say that the data confirm and refine the results previously mentioned and, in particular, establish the order of precision in the measures (probable errors of the order of 30 km/sec or less) and of the dispersions among nebulae of different categories. Thus, by making reasonable allowance for the spread of absolute luminosities, the dispersion among isolated nebulae and small groups appears to be less than 150 km/sec. The dispersion among 11 members of the Local Group is about 60 km/sec, and, in general, the dispersions within groups and clusters increase steadily with compactness of the clusters. For instance,

$\sigma$ is 140 km/sec for the small loose M 81 Group, 275 km/sec for the Large Ursa Major Cloud, 700 km/sec for the Virgo Cluster, 1050 km/sec for the compact Coma Cluster, and perhaps 1200 km/sec for the Corona Borealis Cluster (from velocities of eight nebulae by Humason).

These results are important for the present discussion because they indicate the reliability of the measured red-shifts for clusters used in the formulation of the law of red-shifts. The percentage errors are trivial except possibly for the nearby Virgo Cluster where it is about 5 per cent (in addition to the unknown peculiar motion of the cluster). With increasing distance the percentage errors diminish, and, beyond the Coma Cluster, they can safely be ignored. Evidently the uncertainties in the formulation of the law of red-shifts are entirely those in the distances. Furthermore, because distances are derived from the apparent faintness of objects of supposedly known luminosity, the difficulties seem to be largely those of photometry, except for uncertainties in the fundamental unit of distance.

## IV. The Third Phase (Attempts at a Definitive Formulation with the 200-inch Reflector)

### A. *Extension of the Observed Range*

This was the situation when the completion of the 200-inch, and its accessories, initiated the third and current phase of the development. Because this phase represents an attempt at the definitive formulation of the law, I will be a little more precise in the use of terms. Red-shifts, for instance, will be used for the fractions, $d\lambda/\lambda$, by which details in the spectra are shifted. The shifts cannot be distinguished from Doppler shifts; they are constant throughout any given spectrum within the errors of measurement, the most reliable tests being those made on emission lines in large-scale spectra from H$\alpha$ to $\lambda$ 3727. The term " apparent velocity " will be discarded, and replaced by " velocity " signifying $c \,.\, d\lambda/\lambda$, or red-shifts expressed on a scale of velocities. The procedure is not formally correct but it is convenient.

The current programme has two parts: one is the extension of the observed range and the other is the revision of the distance scale. The first part is straightforward, once suitable clusters are found. During the first season that a nebular spectrograph could be used on the 200-inch (1950–51), Humason obtained spectra in three clusters beyond the limit reached with the 100-inch. They gave velocities of the order of 50 000, 54 000 and 61 000 km/sec, respectively, the last being the Hydra Cluster which had been tried unsuccessfully with the 100-inch. Out to these limits, clusters can be selected from plates with the 48-inch Schmidt reflector, and the programme calls for groups of clusters at successive intervals to furnish mean points for the correlation curve. Beyond 60 000 km/sec, clusters cannot be identified with certainty on the survey plates. Many possible cases suggested by the 48-inch are being checked with the larger reflectors, but otherwise we must depend upon purely chance finds with the 100-inch and 200-inch. Humason has now observed groups of three clusters each at 25 000 km/sec, at 40 000, at 50 000 (representing both galactic hemispheres), and two clusters at about 58 000 to group with Hydra. He is confident that larger velocities, at least up to one-quarter of the velocity of light, can be recorded when such clusters have been located.

## B. *Distances of Clusters*

The second part of the programme, namely, the revision of the distance scale, lags behind the measurement of red-shifts. There are several steps in the programme, namely, the revision of the unit of distance, the setting up of faint magnitude sequences in Selected Areas, the actual measurement of nebular magnitudes, corrections for effects of red-shift on apparent magnitudes, and, finally, the examination of the possibility of errors arising from inter-nebular obscuration and systematic evolutional changes in luminosities during the travel-time of the light from the nebulae. Let us consider these steps in their logical order, beginning with the revision of the unit of distance.

1. *The unit of distance.*—A review of the unit of distance has been pending for more than twenty years—ever since it was realized that the many novae and globular clusters in M 31 were fainter with respect to the Cepheids than was the case in our own system. The discrepancy was emphasized when Baade resolved the body of M 31 and recognized it as a Type II stellar population whose brightest stars were known to be comparable with those in globular clusters. Then, recently, Baade and his younger colleagues calibrated the cluster-type variables (and hence the brightest stars) in globular clusters by extending the colour-magnitude arrays in some clusters down to very faint limits on the main sequence, and comparing them with the corresponding arrays for stars in the vicinity of the Sun (stars so near that their distances were known from direct triangulation). Shapley's value for the cluster-type variables was confirmed within small uncertainties.*

The results clearly indicated that the classical Cepheids in M 31 are nearly $1^{m}\cdot5$ brighter than previously supposed, and, consequently, all our current estimates of absolute distances must be nearly doubled. The revision does not affect the estimates of relative distances, of course, but it nearly doubles the so-called "age of the universe". Further study will be required to reach a definitive value for the revision but meanwhile the factor 2 can be used in discussions. At the moment it seems likely that the final value may be less rather than greater than the even number.

The revision was unexpected because Shapley made his primary calibration from parallactic motions of classical Cepheids in the galactic system (in 1918), and used it to derive his secondary calibration of cluster-type variables. It now appears that the latter is about correct while the former was in error by nearly $1^{m}\cdot5$.

2. *Magnitude standards.*—The faint stellar magnitude standards and colours required for photometry near the limit of the 200-inch are being established in nine Selected Areas forming a convenient pattern around the sky. The first measures, reaching to about $m_{pg}=18\cdot5$, were made by Stebbins and Whitford in three of the Areas, using their photo-cells on the Mount Wilson reflectors. The programme has been continued and expanded by Baum, who has reached the 19th magnitude in all nine Areas and had pushed experimental measures in one or two of them down beyond the 21st magnitude. He is now putting in operation a new instrument designed to reach the limit of the 200-inch. The standards already available are sufficient for the faintest

* Since this was written, preliminary results by Baum in the globular cluster M 13 suggest that grave uncertainties exist in this method of calibration. The colour-magnitude diagrams of the globular clusters M 3 and M 13 appear to differ by two magnitudes in the region of the main sequence. However, Blaauw's direct calibration of the classical Cepheids and Savedoff's fit of pulsation theory to the observations leave little doubt that the order of the correction to the classical Cepheids is $\Delta M=-1^{m}\cdot5$.—A.R.S.

clusters in which red-shifts have been measured, and Baum's new programme should soon reach the faintest clusters in which red-shifts can be measured.

3. *Measurement of nebular magnitudes.*—The measurement of total magnitudes of nebulae has been revolutionized since the war, both by the photo-cell in the hands of Stebbins and Whitford, Pettit, and others, and by the precise photographic techniques of Bigay and Holmberg. All these investigations tend to reach general agreement on magnitudes (and colours, where measured) for given diameters. However, they all agree on the necessity of increasing apertures until no further change in luminosity can be detected. The formulation of a practical definition of diameter, however arbitrary, remains a problem for the future. The subject is under investigation but, meanwhile, the application of uniform principles to apparently similar nebulae in different clusters, using apertures adjusted to red-shifts, is expected to control the relative errors within reasonable limits (possibly of the order of $0^{m}\cdot 1$ either way).

The character of the energy-distribution curves for E nebulae shows that the corrections to observed magnitudes due to red-shifts are smallest in the wave-length region centred about $\lambda$ 6300 A. With this circumstance, the definitive formulation of the law of red-shifts can be most accurately made in the red. Unfortunately, the most reliable method of measuring magnitudes—the photoelectric technique—does not yet reach the necessary limits in the red, although of course it does so easily in the blue and yellow. The compromise adopted at Mount Wilson and Palomar is to use a jiggle-camera in all three colours, blue, yellow and red, and to use the photo-cell for the blue and yellow. This procedure furnishes valuable information on colours as a function of distance, and also permits the photo-cell calibrations in blue and yellow to be applied to the red as well. One additional safeguard has been added to the jiggle-camera programme. Because these images must be large compared with the in-focus images, and the size of throw must be adjusted accordingly, it is found convenient to make images in all three colours with a uniform, medium size, and to repeat in one colour only (generally yellow) with double the throw. Corrections for size of in-focus images can then be transferred to the other colours on the plausible assumption that colour does not vary much throughout the images of elliptical nebulae.

A jiggle-camera for the prime focus of the 200-inch, giving smooth, square images with throws of 0·5, 1, 2 and 4 mm on a side, was put into operation last autumn by my colleague, A. R. Sandage. Plates have now been assembled on half a dozen of the fainter clusters, although the delay in getting transfers of stellar magnitudes from Selected Areas, due to exceptionally poor winter weather, has seriously hampered the progress in final reductions. Baum proposes to join in the programme, now that his new photo-cell is ready for operation, and he has already contributed some data to the calibration problem. Meanwhile, Pettit has measured many of the nebulae in the nearer clusters with a photo-cell on the Mount Wilson reflectors in the course of his programme of getting magnitudes and colours of all nebulae for which radial velocities have been measured (except the very faint cluster nebulae observed with the 200-inch).

These remarks summarize the state of the programme. They indicate why the present discussion must necessarily be a progress report, with perhaps some indications of the direction in which any revisions of the earlier results are likely to fall.

4. *Effects of red-shifts on apparent magnitudes.*—The total effects of red-shifts on apparent luminosities depend upon their interpretation. The mere fact that we observe the shifts clearly indicates that each light-quantum from the nebulae reaches us with reduced energy. Regardless of the interpretation of red-shifts, we must accept the loss of energy by the individual quanta (or reject the fundamental relation $E \, . \, \lambda = \text{constant}$) and we must correct the apparent magnitudes accordingly. These corrections are routine and have the same kind of validity as those for atmospheric extinction or instrumental characteristics. The bolometric luminosities are evidently reduced by the factors $(1 + d\lambda/\lambda)$, but since the red-shifts change the wave-length scales of the energy-distribution curves, the effects are selective. If the true, unshifted energy curves are known, they can be distorted by the observed shifts, and the effects then traced through the atmosphere, the telescopes, and the measuring devices to give the desired corrections. This procedure has been possible only since Stebbins and Whitford in 1947 derived the energy curve for the elliptical nebula, M 32, reduced to no atmosphere. The curve is derived from 6-colour photo-cell measures, ranging from $\lambda$ 10300 to $\lambda$ 3530.

Because elliptical nebulae, which dominate cluster populations, are remarkably homogeneous as to spectral types, colours and distribution of colour throughout their images, the energy curve for M 32 can be applied directly to the estimation of corrections for red-shifts to apparent magnitudes of all elliptical nebulae in any region of the observed range in the spectrum. Minor uncertainties may arise from the facts that M 32 is a dwarfish nebula, that it is superposed on an outer arm of M 31 and, finally, that it is in the low galactic latitude of $-22°$. The curve will be redetermined eventually from giant elliptical nebulae in the Virgo Cluster, but the present formulation seems to be fairly reliable and it is all we have.

Luminosities of nebulae are measured through " windows " consisting of the atmosphere, telescope and plate-filter or other receiving devices. By means of the energy-curve for M 32, the energy getting through a window from a nebula with a known red-shift can be compared with the energy that would have got through in the absence of a red-shift. These ratios furnish the desired corrections required by the mere observed presence of the red-shifts. These corrections I once called the " energy effects " and, I repeat, they must be applied to the measured luminosities, regardless of the interpretation of red-shifts.

5. *Progress report.*—In eleven clusters blue and yellow magnitudes are available for a number of the brightest nebulae, and in two of the faint clusters red magnitudes also have been measured. Among these data, corrected for galactic obscuration and for red-shifts, the 1st, 3rd, 5th and 10th nebula in each cluster have been used as distance indicators, and the magnitudes plotted against logarithms of red-shifts. The curves suggest that the brightest nebulae in the different clusters are about equally luminous, but that the spread between 1st and 10th varies with the richness of the cluster or with the size of the sample. On this interpretation of the data, the most accurate results should be furnished by the 10th nebula in the great clusters only. This procedure represents a compromise, because a higher ordinal number, say the 25th nebula, would be expected to reduce accidental errors but at the price of introducing systematic errors due to variations in size of samples.

The yellow magnitudes offer the most reliable results, because corrections for red-shifts in the blue are large and uncertain while the red measures are not yet assembled in sufficient numbers. The favoured correlation, at the moment, is

$$\log_{10} v = 0{\cdot}2\, m_{\rm pv} + 1{\cdot}16 \text{ (10th nebula)},$$

where

$$v = c\,.\,d\lambda/\lambda.$$

This correlation is shown in Fig. 1. Evidently the correlation is linear within the uncertainties of the data, and the residuals are surprisingly small. There are no indications of absorption in space in these particular directions. As the number of observed clusters increases, the residuals should offer a critical test of the presence of inter-nebular absorption. Meanwhile the colours offer an approximate criterion where spectra have not yet been recorded.

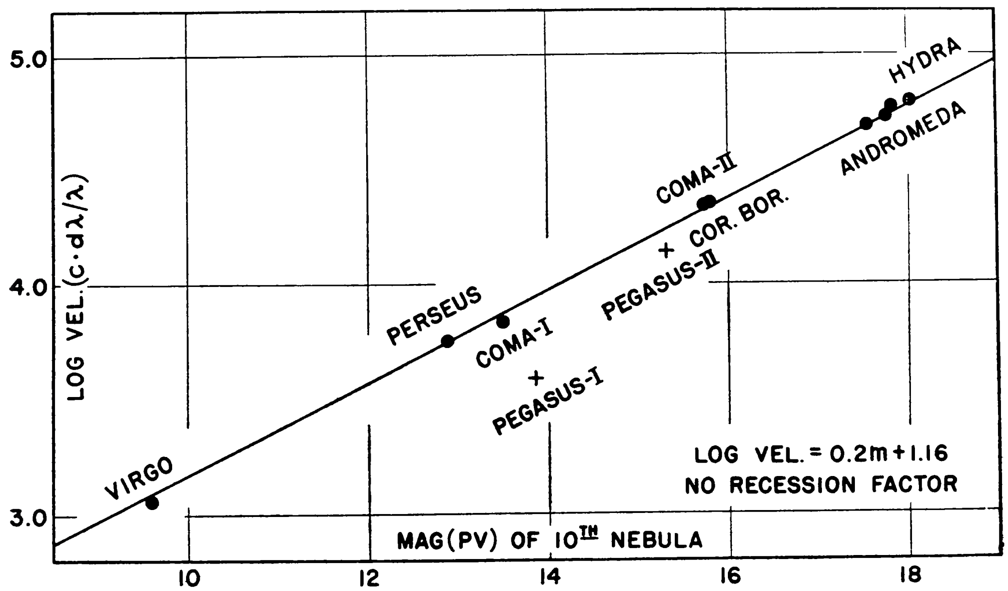


FIG. 1.—*The relation between velocity and apparent magnitude. The new data obtained with the 200-inch are represented by the last four points on the regression line. Humason's red-shifts are expressed on a scale of velocities as c . dλ/λ in km/sec. The photovisual magnitudes have been corrected for the energy effect only. They do not include the recession factor.*

Similar correlations are readily derived for other distance indicators in clusters, for isolated field nebulae, for isolated pairs, and for brightest stars (blue supergiants) in late-type spirals and irregulars with Type I populations. The curves are all parallel, and are displaced by 0·2 times the difference in the absolute magnitudes of the distance indicators. In this way it is possible to calibrate all the correlations in terms of any one, the most accessible, of course, being the brightest stars in nearby nebulae. Thus, the $M_{\rm pv}$ of the 10th nebula seems to be of the order of $-17{\cdot}4$ on the old distance scale, and $-18{\cdot}9$ or somewhat less on the new.

Another method is offered by novae and globular clusters in Type II populations found in the Virgo Cluster. The rather meagre results now in hand appear to be consistent with those from brightest stars in Type I populations. When precise photometric data are available, the two methods together should furnish a reliable calibration for all the correlations.

Meanwhile, the relations between red-shift and apparent magnitude represent relative distances that should be reliable, and the slope of the curve,

$\log_{10} v = 0{\cdot}2m + C$, indicates a linear relation to a very close approximation. It should be emphasized that the magnitudes have not been corrected for recession of the nebulae.

Such corrections are closely represented by the term, $\Delta m = d\lambda/\lambda$, which must be subtracted from the measured magnitudes if the red-shifts are interpreted as Doppler shifts. The corrections range up to a maximum of $0^{\mathrm{m}}{\cdot}2$ for the Hydra Cluster, and, if the uncorrected data define a linear relation, the corrected data would lead to a non-linear relation in the sense of an accelerated expansion. However, the data now available are not sufficient to furnish a critical test of the two interpretations; either one is permitted by the uncertainties, although it seems unlikely that a decelerated expansion can be represented.

In conclusion it should be repeated that the discussion necessarily ends with a progress report, and not with a definitive solution of the problems considered. Nevertheless, it can be stated with some confidence that a definitive formulation of the law of red-shifts will be available before long, in the form of a relation between red-shifts and apparent magnitudes, out to red-shifts of the order of 0·25.

When no recession factors are included, the law will represent approximately a linear relation between red-shifts and distance. When recession factors are included, the distance relation is expected to be linear or non-linear in the sense of accelerated expansion. The "age of the universe" is likely to be between 3000 and 4000 million years, and thus comparable with the age of rock in the crust of the Earth.

The implications of the law established over this distance may be traced out to nearly double that distance by its effect on the apparent distribution of nebulae in depth, $N_m = f(m)$.

Thus, if red-shifts do measure the expansion of the universe, we may be able to gather reliable information over a quarter of its history since expansion began, and some information over nearly a half of the history.

As for the future, it is possible to penetrate still deeper into space—to follow the red-shifts still farther back in time—but we are already in the region of diminishing returns; instruments will be increasingly expensive, and progress increasingly slow. The most promising programmes for the immediate future accept the observable region as presently defined, hope for only modest extensions in space, but concentrate on increased precision and reliability in the recorded description. The reconnaissance is being followed by an accurate survey; the explorations are pushed towards the next decimal place instead of the next cipher. This procedure promises to reduce the array of possible worlds as surely as did the early rapid inspections of the new territory. And later perhaps, in a happier generation, when the cost of a battleship can safely be diverted from insurance of survival to the consolations of philosophy, the march outward may be resumed.

For I can end as I began. From our home on the Earth, we look out into the distances and strive to imagine the sort of world into which we are born. Today we have reached far out into space. Our immediate neighbourhood we know rather intimately. But with increasing distance our knowledge fades, and fades rapidly, until at the last dim horizon we search among ghostly errors of observations for landmarks that are scarcely more substantial. The search will continue. The urge is older than history. It is not satisfied and it will not be suppressed.