YALE UNIVERSITY

MRS. HEPSA ELY SILLIMAN

MEMORIAL LECTURES

---

# THE REALM OF THE NEBULÆ

# FRONTISPIECE

## *Outer Region of the Great Spiral in Andromeda, Messier 31. (The Open Cluster Is about 48′ South-Preceding the Nucleus.)*

THE observer looks out through the swarm of stars which surrounds him, past the borders and across empty space, to find another stellar system —the nebula Messier 31. The brightest objects in the nebula can be seen individually, and among them the observer recognizes various types that are well known in his own stellar system. The apparent faintness of these familiar objects, as they lie in the nebula, indicates the distance of the nebula—a distance so great that light requires seven hundred thousand years to make the journey.

A star-cloud (catalogued as NGC 206), an open star-cluster, a globular star-cluster, and a Cepheid variable star at maximum luminosity are marked on the photograph. The Cepheid has a period of 18.28 days and a magnitude at maximum of 18.75.

Plate by Duncan with the 100-inch reflector, August 24, 1925; north is at the top; 1 mm. = 11″.8.

# THE

# REALM OF THE NEBULÆ

BY EDWIN HUBBLE

OF THE

MOUNT WILSON OBSERVATORY

CARNEGIE INSTITUTION OF WASHINGTON

OXFORD UNIVERSITY PRESS

LONDON: HUMPHREY MILFORD

1936

Printed in the United States of America

# THE SILLIMAN FOUNDATION

In the year 1883 a legacy of eighty thousand dollars was left to the President and Fellows of Yale College in the city of New Haven, to be held in trust, as a gift from her children, in memory of their beloved and honored mother, Mrs. Hepsa Ely Silliman.

On this foundation Yale College was requested and directed to establish an annual course of lectures designed to illustrate the presence and providence, the wisdom and goodness of God, as manifested in the natural and moral world. These were to be designated as the Mrs. Hepsa Ely Silliman Memorial Lectures. It was the belief of the testator that any orderly presentation of the facts of nature or history contributed to the end of this foundation more effectively than any attempt to emphasize the elements of doctrine or of creed; and he therefore provided that lectures on dogmatic or polemical theology should be excluded from the scope of this foundation, and that the subjects should be selected rather from the domains of natural science and history, giving special prominence to astronomy, chemistry, geology, and anatomy.

It was further directed that each annual course should be made the basis of a volume to form part of a series constituting a memorial to Mrs. Silliman. The memorial fund came into the possession of the Corporation of Yale University in the year 1901; and the present work constitutes the twenty-fifth volume published on this foundation.

# PREFACE

THIS book consists of the Silliman Lectures, delivered at Yale University in the autumn of 1935, with the addition of an introductory chapter. The subject is The Realm of the Nebulæ, which is the portion of the universe, thinly populated with nebulæ, that can be explored with existing telescopes. The discussion, since it was designed for a general audience, is necessarily incomplete, but it includes most of the important questions concerning which definite statements can be made at the moment. The subject is presented from the observer's point of view; the rich field of theoretical literature has scarcely been touched. With these limitations, the book is believed to furnish an authentic picture of a typical case of scientific research in the process of development.

The writer is responsible for whatever is not specifically attributed to others. He gratefully acknowledges an indebtedness to his colleagues at the Mount Wilson Observatory; in particular, to Milton Humason, who has contributed almost the whole of the recent data on nebular spectra and red-shifts, to Walter Baade, whose interests and exact knowledge range over the entire field of nebular research, and to Sinclair Smith, who is investigating the physical constitution of nebulæ and of clusters. In the field of cosmology, the writer has had the privilege of consulting Richard Tolman and Fritz Zwicky of the California Institute of Technology. Daily contact with these men has engendered a common atmosphere in which ideas develop that cannot always be assigned to particular sources. The individual, in a sense, speaks for the group.

The conquest of the Realm of the Nebulæ is an achievement of great telescopes. It began with the identification of nebulæ as independent stellar systems, comparable with our own system of the Milky Way. Once the nature of the nebulæ was known, methods of estimating distances were readily developed, and the new field was open to investigation.

The instrument which definitely established the identification—and enlarged the domain of positive knowledge a thousand-million-fold—is the Hooker telescope—the 100-inch reflector of the Mount Wilson Observatory of the Carnegie Institution of Washington. It is the largest telescope in operation, it has the greatest light-gathering power, and it penetrates to the greatest distance. For these reasons, it defines the present extent of the observable region of space, and it has contributed the most significant data to the study of the region as a sample of the universe.

In recognition of the unique rôle which the 100-inch has played in the progress of nebular research, the illustrations in this book are nearly all reproductions of photographs made with the great telescope. The only exceptions, Plate XII and parts of Plates I and II, are reproductions of photographs made with the companion of the 100-inch, namely, the 60-inch reflector of the Mount Wilson Observatory. The reproductions were prepared at the Observatory by Ferdinand Ellerman and Edison Hoge.

Finally, the writer wishes to express his gratitude to Godfrey Davies, of the Huntington Library, who has criticized the form of presentation. His kindly intervention has very materially reduced the number of obstacles which the reader would otherwise have encountered.

E. H.

Mount Wilson Observatory,
February, 1936.

# CONTENTS

# ILLUSTRATIONS

## PLATES



## FIGURES

# TABLES

# THE REALM OF THE NEBULÆ

## INTRODUCTION

## SCIENTIFIC RESEARCH

SCIENCE is the one human activity that is truly progressive. The body of positive knowledge is transmitted from generation to generation, and each contributes to the growing structure. Newton said, "If I have seen farther it was by standing on the shoulders of giants."[1] Today, the least of the men of science commands a wider prospect. Even the giants are dwarfed by the great edifice in which their achievements are incorporated. What a Newton might see today, we do not know. And tomorrow, or a thousand years hence, even our dreams may be forgotten.

This remarkable attribute of science is bought at a price—the strict limitation of the subject matter. "Science," as Campbell remarks, "deals with judgments concerning which it is possible to obtain universal agreement."[2] These data are not individual events, but the invariable associations of events or properties which are known as laws of science. Agreement is secured by means of observation and experiment. The tests represent external authorities which all men must acknowledge, by their actions if not by their words, in order to survive.

Science, since it deals only with such judgments, is necessarily barred from the world of values. There no external authority is known. Each man appeals to his private god and recognizes no superior court of appeal. Wisdom is a personal achievement and is difficult to transmit. Sarton writes:

1 Sarton has traced this quotation to Bernard of Chartres who died in 1126. *Isis*, No. 67, 107, 1935.

2 Norman Campbell, *What Is Science?* (1921), p. 27; quoted with permission of the author and the publisher (Methuen, London).

The saints of today are not necessarily more saintly than those of a thousand years ago; our artists are not necessarily greater than those of early Greece; they are likely to be inferior; and, of course, our men of science are not necessarily more intelligent than those of old; yet one thing is certain, their knowledge is at once more extensive and more accurate. The acquisition and systemization of positive knowledge is the only human activity that is truly cumulative and progressive.[3]

The special methods of science are efficient and powerful when they are used within their proper field—the region closed and bounded by the necessity for agreement on the subject matter. The methods are so successful, indeed, that attempts are constantly made to apply them in other fields—to the study of things as they should be rather than things as they are. The results are seldom convincing. The calculus of values, if it is ever formulated, will probably have little in common with the calculus of science. Nevertheless the atmosphere of scientific research—the disinterested curiosity, controlled imagination, and passion for impersonal tests—is by no means unique. It may, and often does, produce a definite influence in those fields from which scientific technique is barred. That influence, the man of science likes to believe, is generally beneficial. The special methods of science will be discussed hereafter, but they are characterized by the remark that research attempts to discover laws and to explain the laws by theories, the ultimate goal being the understanding of the physical structure and operation of the world in which we live.[4]

Actually, investigations are carried out in various ways, two of which may be mentioned as typical examples. One emphasizes the observational approach, the other, the theoretical point of view. The observer commonly starts by accumulating an isolated group of data,

[3] George Sarton, *Introduction to the History of Science* (1927), I, 3.

[4] The reader is referred to Campbell's *What Is Science?* for further discussion and in particular for the significance of the word "explain" as used above.

together with their estimated uncertainties. The material is studied, usually by graphical methods, and relations are found between various features.

The data, for instance, may consist of apparent luminosities of nebulæ and red-shifts in their spectra. These terms will be explained later; for the moment, they may be considered merely as two measurable features, *A* and *B*. When one feature is plotted against the other, it appears that, on the average, red-shifts increase as the luminosities diminish—the fainter the nebula, the larger the red-shift. The conclusion is important even in this qualitative form, but the significance would be enormously enhanced if the precise form of the relation could be established in numerical terms.

The points that have been plotted make a scatter-diagram through which many different correlation curves might be drawn, all of which would represent the particular data in a fairly satisfactory way. The observer selects from the possible relations, the simplest one that is consistent with the body of general knowledge. In the particular case under discussion, the adopted relation, as will appear later, corresponds to a linear relation between red-shifts and distances as indicated by the faintness of the nebulæ.

The relation is plausible but not unique. The true relation might be a curve which was nearly linear within the range covered by the observations, but which departed widely from a straight line in the regions beyond the faintest nebulæ in the group. This possibility was investigated by *extrapolating* the adopted relation—extending it far out into the hitherto unobserved regions—and testing it by new observations. Such a procedure often leads to minor, or even to major, revisions in the relation first selected: it has been said that research proceeds by successive approximations. However, in the investigation of red-shifts, no revision was definitely indicated. The linear relation has survived repeated tests of this nature and is known to hold, at least approximately, as far out into space as the observations can be carried with existing instruments.

The investigation thus leads to a new law—out to a certain limiting distance, red-shifts are an approximately linear function of distances. Beyond the limit of the observations, the form of the function, or relation, is speculative. For this reason, the law is empirical and it must remain empirical until it is explained by an accepted theory.[5] Some students believe that suitable theories have already been formulated, and they may be right in their opinion. The question will be decided by further investigations.

The particular case of the red-shifts, or rather a simplified version of that investigation, has been discussed at some length because it offers a fair example of the observational approach. An isolated group of data is studied and the results are interpreted against the background of general knowledge. Then follows the process of extrapolation, and tests, and appropriate revision. The observations and the laws which express their relations are permanent contributions to the body of knowledge; the interpretations and the theories change with the spreading background. The research sweeps outward and develops an observable region around a given center—a realm of positive knowledge. Beyond the horizon, is the realm of speculation. The observer, if he ventures therein, can only throw his empirical relation into the blue, and search for inconsistencies with extrapolations from other centers.

The theoretical investigator, employing another method of approach, studies the isolated, empirical laws established by observers. He searches among them for some common element, some generalization by which he may

[5] Red-shifts are directly measured out to distances of the order of 250 million light-years. The linear relation is consistent with the body of general (observational) knowledge out to about 400 million light-years. Beyond this limit, where observational data are not available, the extrapolation must be tested by its consistency with accepted theories, and these are not yet established in the detailed form required for the tests.

collect various observed relations into a single statement. In short, he strives to invent a theory to explain the laws. The approach to the theory may be by logic or by intuition—it is immaterial which method is used. The important feature is the ability to explain the relations already observed, and to predict new relations.

A fundamental theory and the various relations that can be readily deduced from it, form a consistent pattern which may apply to certain aspects of the universe, or even to the universe as a whole. The author projects his pattern on the world around him in order to see how closely the two coincide. The known empirical laws naturally fall into place, if the author is competent, and new laws may be predicted. The success of the venture depends largely on the verification of predicted relations. If no such test is possible, the value of the pattern must be measured by the order and kinship which it introduces among known, but hitherto unrelated, phenomena. Unless the systemization is of a high order, the theory will be regarded as speculation.

Many theories are formulated but relatively few endure the tests. The survivors, in general, must be occasionally revised to conform with the growing body of knowledge. The ability to theorize is highly personal; it involves art, imagination, logic, and something more. An outstanding genius may invent a successful new type of theory; first-rate men may follow the lead and develop other theories on the same pattern; less competent minds are embarrassed by the custom of testing predictions.

The theoretical investigator frequently works from the circumference toward the center, while the observer works from the center toward the circumference. The latter extrapolates outward, while the former, in a sense, extrapolates inward. When the two agree, they inspire some confidence in the significance of the general pattern.

The distinction is seldom so clear as the foregoing discussion would imply. Almost every investigation combines both methods of approach, although they may be present in varying proportions. Research men attempt to satisfy their curiosity, and are accustomed to use any reasonable means that may assist them toward the receding goal. One of the few universal characteristics is a healthy skepticism toward unverified speculations. These are regarded as topics for conversation until tests can be devised. Only then do they attain the dignity of subjects for investigation.

The present writer is primarily an observer. The chapters which follow describe the development of a new phase of astronomical research—the exploration of the realm of the nebulæ. Emphasis is placed upon observational data—upon the positive knowledge that has been assembled—rather than upon interpretations, whether theoretical or speculative. The latter aspect has been exploited in popular literature, and a number of books have been written, some of which stir the imagination. The observational data, for the most part, rest quietly in the technical journals. The references in this book are largely restricted to original sources rather than to rediscussions of the data.

Much of the source material may be of interest to the general reader. It presents a case history of scientific research in a rather simple form. Some insight may be gained concerning that little known activity, without the preliminary acquisition of an extensive technical vocabulary. The subject was new, the data were rough, and the dangers of overdiscussion in a formative period were fully realized; therefore, the treatment was generally direct and free from the intricate devices used in the case of refined, quantitative investigations. Analysis of the data, for instance, was almost always graphical rather than mathematical. Naturally, certain special terms are constantly met in the source material. Several of them are so convenient that they will be used throughout the present book. Their explanation forms a necessary introduction to the chapters which follow.

*The Language of Astronomy.*

Astronomy, as other disciplines, has its own technical vocabulary of words and phrases with precise definitions. The terms always carry the same significance and substitutes are not employed. Variety is sacrificed for precision. Some of the terms have long histories. The words themselves are familiar, but the technical definitions are far removed from common usage. Other terms are recent additions, deliberately coined in efforts to avoid the confusion of associated ideas. The result is a vocabulary so strange to the general reader that scientific reports, many of which are relatively simple, appear to be wrapt in a mantle of obscurity. Translation into the language of general discourse is a difficult art and frequently blurs the meaning for the dubious advantages of specious familiarity. It is for this reason that a few of the more common terms will be used in their technical purity. They are restricted to units of distance and of luminosity, and to certain types of variable stars which can be recognized by their behavior, wherever they are found. The glossary follows, and concludes with a brief discussion of the word *nebula.*

*Units of Distance.*

Miles and kilometers (1 mile = 1.6 km.) are occasionally used, but great distances are expressed in light-years (l.y.) or in parsecs (par.). The light-year is simply the distance which light travels in a year. Since the speed of light is about 186,000 miles per second, the number of miles in a light-year is roughly six followed by twelve ciphers ($5.88 \times 10^{12}$ miles = $9.46 \times 10^{12}$ km.).

Light reaches the earth from the moon in about 1⅓ seconds; from the sun, in about 8⅓ minutes; from Pluto, the outermost planet, in roughly 6 hours. The nearest star (Alpha Centauri) is at a distance of 4.3 light-years; the nearest nebula (Large Magellanic Cloud), about 85 thousand light-years; the faintest nebulæ that have been photographed (limit of the 100-inch reflector) are at an average distance of the order of 500 million light-years.

Beyond the nearer stars, distances cannot be determined with precision. Errors of 10 per cent are considered to be very small and uncertainties of 25 per cent represent reasonable accuracy. Under these conditions, distances are generally expressed in round numbers, using only one or two significant figures.

The term *parsec* was coined for the distance corresponding to a *parallax* of *one second* of arc. The unit is very convenient in many calculations and for this reason it is generally employed in technical papers. It is used only rarely in the following chapters and in these cases the distances are also expressed in light-years (1 par. = 3.258 l.y.).

For those who may be interested, the derivation is as follows. The *astronomical unit* (not used in this book) is the mean distance from the earth to the sun, $9.29 \times 10^7$ miles = $1.49 \times 10^8$ km. The *parallax* of an object is the angle subtended by the astronomical unit as seen from the distance of the object. Now an angle of one second of arc is subtended by an object whose distance is about 206,000 times its diameter. Therefore, the parsec is about $1.92 \times 10^{13}$ miles, or, as mentioned above, about 3.258 light-years.

Parallaxes of the nearer stars are measured by direct triangulation from opposite points on the earth's orbit around the sun. The largest known stellar parallax, that of Alpha Centauri, is about three quarters of a second of arc[6] (distance = 1⅓ par. = 4⅓ l.y.), and parallaxes of one hundredth of a second (distance = 100 par. = 326 l.y.) can be measured with reasonable accuracy. These directly determined distances are used to calibrate the many indirect methods of estimating greater distances.

[6] This angle is about that subtended by a dime at a distance of three miles.

*Apparent Magnitudes.*

LUMINOSITIES are expressed as *magnitudes*. The usage is traditional, although the precise calibration of the scale is modern. The ancient astronomers recorded apparent luminosities of stars primarily as aids in their identifications. The earliest classification was presumably into the natural groups: bright, intermediate, and faint. Later, each group was probably subdivided into two sections. At any rate the oldest surviving catalogue of stars—that given in Ptolemy's *Almagest*, composed in the first half of the second century A.D.—employed a 6-group classification. The scheme persisted into modern times and furnished the basis of the present system.

The groups came to be known as *magnitudes*. About fifteen of the brightest stars were included in the first magnitude, and the faintest stars that could be seen with the naked eye formed the sixth magnitude. The five intervening steps represented roughly equal ratios in the luminosities. Each magnitude was brighter or fainter than the next magnitude by an approximately constant but undetermined factor, which is now known to have been of the order of 2.5. Thus the first magnitude was about 2.5 times brighter than the second, $(2.5)^2 = 6.25$ times brighter than the third, $(2.5)^3 = 16$ times brighter than the fourth, $(2.5)^4 = 40$ times brighter than the fifth, and $(2.5)^5 = 100$ times brighter than the sixth. The scheme was reached instinctively since, according to the relation now known as Fechner's law of stimuli, the eye distinguishes equal ratios rather than equal increments of luminosity.

Ptolemy's magnitudes were accepted almost uncritically for many centuries. Even in modern times, when independent estimates of apparent luminosities began to accumulate, and whole magnitudes were split into halves, thirds, and tenths, the same system was generally employed. Telescopic stars were assigned to magnitudes numerically larger than six. Eventually, when the importance of a precise, uniform scale was realized, the value of the constant factor, or ratio of luminosities, was carefully investigated. The results varied considerably, but in most cases they were of the order of 2.5. Finally, in 1856, Pogson (1829–91) at the Radcliffe Observatory in Oxford, made a suggestion that met with general approval. He said in effect, let us adopt, as an arbitrary but very convenient value of the ratio, the number 2.512···, whose common logarithm[7] is exactly 0.4.[8] The assumption is made that the first magnitude is precisely 100 times brighter than the sixth, and the range is divided into five steps whose ratios are equal. The logarithm of 100—namely 2.0—divided by 5, is 0.4, which is the logarithm of the luminosity-ratio between successive magnitudes.

This scale, which is not very different from those employed in the older catalogues, is in current use today. The zero-point has been adopted by international agreement to conform with certain published magnitudes for a standard sequence of stars in the vicinity of the north pole of the heavens. Magnitudes of other stars are determined by comparisons, direct or indirect, with the standard sequence.

Magnitudes are proportional, not to luminosities, but to logarithms of luminosities. If $L_0$ is the luminosity of a standard star whose magnitude (defined by agreement)

[7] No mathematics is included in this book except simple linear equations (statements of equality) and the use of common logarithms instead of exponentials. The common logarithm of a number is merely the power of ten which represents the number. Thus, for any (positive) number, *a*, the logarithm is defined as follows:

$$a = 10^{\log a}.$$

In other words, if $a = 10^b$, then $b = \log a$, and $a$ is sometimes called the antilog of $b$. The quantity mentioned above, 0.4, is the logarithm of 2.512···, since $10^{0.4} = 2.512\cdots$.

Because magnitudes are logarithmic functions of luminosities, and luminosities vary as a power of the distance, it will often be convenient to use logarithms of distances, red-shifts, or other characteristics in order to express relations by simple linear equations rather than by more complicated expressions.

[8] *Monthly Notices of the Royal Astronomical Society*, 17, 12, 1856.

is $m_0$, then the magnitude, $m$, of any other star whose luminosity is $L$, is given by the relation

$$0.4\ (m - m_0) = \log L_0/L$$
$$m = m_0 + 2.5 \log L_0/L.$$

The scheme is convenient because the ratio of the luminosities, $L_0/L$, can be measured easily and accurately, although the absolute values of the individual luminosities, $L_0$ and $L$, are very difficult to determine.

Two points should be noticed. First, the magnitudes increase slowly while the corresponding luminosity-ratios increase rapidly. Thus a difference of 0.1 mag., represents a luminosity-ratio of 1.1 to 1.0, while a difference of 10 mag., represents a ratio of 10,000 to 1. A short table of corresponding values emphasizes the relation.

TABLE I.

*Magnitude-Differences and Luminosity-Ratios.*

| $m - m_0$ | $L_0/L$ | $m - m_0$ | $L_0/L$ |
|---|---|---|---|
| 0.1 | 1.1 | 5. | 100. |
| 0.5 | 1.6 | 7.5 | 1,000. |
| 1.0 | 2.5 | 10. | 10,000. |
| 2.0 | 6.3 | 15. | 1,000,000. |
| 2.5 | 10.0 | 20. | 100,000,000. |

The second point is that the numerical values of the magnitudes *increase* as the luminosities *diminish*. Magnitudes measure faintness. A large magnitude, say +20, refers to an exceedingly faint star, while a small magnitude, say +0.1, indicates a bright star (Vega). Still brighter luminosities are expressed by negative magnitudes. The photographic magnitude of the brightest object in the sky, namely the sun, is about −26; the full moon, about −11; Venus, about −3. Two stars have negative magnitudes—Sirius is −1.6, and Canopus is about −0.5. Otherwise, the magnitudes of stars are all positive (except those of occasional novæ near their maxima). The faintest stars that have been photographed with the largest telescope are of the twenty-second magnitude, or about 3,000 million times fainter than Sirius.

There are many systems of magnitudes, but all are on the same scale—namely,

$$m = m_0 + 2.5 \log L_0/L,$$

where $m_0$ is arbitrarily defined. The systems are distinguished by appropriate subscripts to the general symbol $m$. Thus $m_{pg}$ represents photographic magnitudes. However, since this one system will be used almost exclusively in the following chapters, the subscript will be dropped and the symbol, $m$, will henceforth be used for photographic magnitudes alone.

These quantities represent blue-violet luminosities. Visual, or the nearly equivalent photo-visual, magnitudes indicate yellow luminosities. A red star is brighter visually than photographically, and for a blue star the relation is reversed. Thus the difference between the visual and the photographic magnitude, called the color-index (C.I.), measures the color of an object. The two systems of magnitudes are adjusted so that the color-index is zero for a white star (spectral type A0). Therefore, the color-index is negative for a blue star and positive for a yellow or a red star. The range among normal stars is from about −0.4 to +2.0 magnitudes, although exceptional cases (for instance, the very red N-type stars) may be found beyond these limits. The sun, a yellow star, has a color index of about +0.6 magnitudes.

The magnitudes hitherto discussed are called *apparent* magnitudes and are expressed by the symbol $m$. They indicate the luminosities of objects as they appear in the sky, and represent a combination of distance and intrinsic luminosity (or candle power). For instance, a star of the eleventh apparent magnitude ($m = 11$), might be a dwarf star at a small distance, or a giant star at a great distance, or any intermediate combination.

### *Absolute Magnitudes.*

INTRINSIC luminosities are measured by *absolute* magnitudes, expressed by the symbol $M$. They are on the same scale as apparent magnitudes and, as before, the zero-point is arbitrarily defined. Actually, the *absolute* magnitude, $M$, is merely the *apparent* magnitude, $m$, which an object would exhibit if that object were at a certain standard distance from the observer. The standard distance, by definition, is 10 parsecs or 32.6 light-years. At this distance, the fainter dwarf stars could not be seen with the naked eye, the sun would be just comfortably visible, the brightest giants would surpass the planet Venus and would be seen in the daytime. The average nebula would appear several times brighter than the full moon.

At the standard distance of 32.6 light-years, $m = M$. At any other distance, the difference, $m - M$, is a known function of the distance (in fact it is sometimes called the modulus of the distance). The relation is

$$\log d \text{ (in parsecs)} = 0.2\ (m - M) + 1$$

or

$$\log d \text{ (in light-years)} = 0.2\ (m - M) + 1.513.$$

Now $m$ can always be observed. Therefore, if either of the two quantities, $d$, or $M$, is known, the other can be readily calculated. The methods of estimating great distances are based almost entirely upon this simple relation. The absolute magnitudes of various types of stars have been determined from objects whose distances were already known. Thus, wherever the types can be recognized, the apparent magnitudes can be measured and the distances can be derived from the differences, $m - M$.

### *Period-Luminosity Relation for Cepheids.*

ONE application of the method is of special interest in the study of nebulæ. The stars concerned are called Cepheid variables, after the type example, Delta Cephei. They are pulsating stars which brighten rapidly and fade slowly, repeating the cycle continuously and faithfully. The period (length of the cycle) is constant for an individual star, but varies from one star to another, ranging from about one day to a hundred days. The range in luminosity is also constant for a given star, but varies within the group from about 0.8 to 2.0 magnitudes. From these features the Cepheids are easily recognized wherever they may be found.

Several dozen are known among the stars in the galactic system, but they are scattered at wide intervals and even the nearest is very remote from the earth. For this reason the determination of distances, and hence absolute magnitudes, has been a difficult problem. Before it was completely solved, a new feature, of extraordinary significance, was discovered among Cepheids in the small Magellanic Cloud.

The Cloud is an independent stellar system and a close neighbor—actually a satellite—of the galactic system. It offers a unique opportunity to study a sample collection of stars which are all at about the same distance from the observer. It is so remote that only the brighter stars (giants and supergiants) can be observed, but this drawback is more than compensated for by the fact that relative apparent luminosities, within the Cloud, are also relative absolute luminosities.[9]

Investigations of the Cloud, made at the Harvard College Observatory, led to the discovery of several hundred variable stars. Some were carefully followed and most of these were identified as Cepheids. As early as 1908, Miss Leavitt, who made the investigations, remarked that the brightest Cepheids had longer periods (pulsated more

[9] Since the diameter of the Cloud is small compared to its distance, $m - M =$ constant, for all objects in the Cloud. Thus $M = m -$ constant, and differences in $m$ represent differences in $M$ ($\Delta M = \Delta m$). Later, when the value of the constant was determined for a few stars (Cepheids), the distance of the Cloud and the absolute magnitudes of its many members were immediately known.

slowly) than the fainter Cepheids. In 1912,[10] she announced a definite period-luminosity relation. The logarithms of the periods increased directly with the median magnitudes (the mid-points between maxima and minima). Thus, if the period of any Cepheid in the Cloud was known, the apparent magnitude was determined. The relation evidently reflected certain inherent characteristics of Cepheids which would presumably be found in all such stars wherever they might be located—in the Cloud, in the galactic system, or elsewhere. If the relation could be numerically calibrated—if the *absolute* magnitude for any one period could be established—the Cepheids, since they are so readily identified, would furnish a powerful method of estimating great distances.

Hertzsprung,[11] who immediately recognized the full significance of the period-luminosity relation, made the first calibration in 1913. He determined the mean distance of thirteen galactic Cepheids from their parallactic motions (reflections of the motion of the sun among the stars). The individual distances were very uncertain, but the mean for the group was fairly reliable and furnished a mean absolute magnitude corresponding to a particular mean period. These data permitted him to calibrate the period-luminosity relation, make a provisional estimate of the distance of the Cloud, and examine the distribution of Cepheids in the galactic system.

Five years later (1918), Shapley[12] repeated the calcula-

[10] *Harvard College Observatory Circular,* No. 173 (1912), where references to earlier statements will be found.

[11] *Astronomische Nachrichten,* 196, 201, 1913. Russell had previously derived the mean absolute magnitude of presumably the same thirteen Cepheids, but no details were given and the period-luminosity relation was not discussed (*Science,* 37, 651, 1913).

[12] "Contribution of the Mt. Wilson Observatory," No. 151; *Astrophysical Journal,* 48, 89, 1918. Shapley rejected two of the thirteen galactic Cepheids as abnormal, but derived a mean *M* for a given period only 0.2 mag. brighter than Hertzsprung had found. The material revision in the calibration arose largely from new information concerning colors of Cepheids. The magnitudes of the bright galactic Cepheids were visual, while those of the

tions and materially revised the calibration. Later alterations, introduced by Shapley, led to the current form of the period-luminosity relation. Further revision is expected to be of minor importance. Thus, wherever a Cepheid may be found, the period will indicate the absolute luminosity, and the apparent faintness then measures the distance. It was by this method that the first reliable distances of nebulæ were determined.

## *Nebulæ and External Galaxies.*

The astronomical term *nebulæ* has come down through the centuries as the name for permanent, cloudy patches in the sky that are beyond the limits of the solar system. The interpretation of these objects has frequently changed, but the name has persisted. It was once believed that all nebulæ were clusters or systems of stars; later, it was found that some were composed of gas or dust. As new theories were developed, various new names were introduced, but in general the names did not survive. Only one revision has been permanent. Certain star-clusters, easily resolved with telescopes of moderate power, and obviously subordinate members of the galactic system, have been withdrawn from the list of nebulæ to form a separate and distinct class of objects.

Today, the term *nebula* is used for two quite different kinds of astronomical bodies. On the one hand are the clouds of dust and gas, numbering a few score in all, which are scattered among the stars of the galactic system. These have been called *galactic nebulæ.* On the other hand are the remaining objects, numbering many millions, which are now recognized as independent stellar systems scattered through space beyond the limits of the galactic system. These have been called *extragalactic*[18]

faint variables in the Cloud were photographic. Shapley had found a relation between period and color and was able to reduce one system of magnitudes to the other with confidence.

[18] The adjective *anagalactic,* proposed by Lundmark, is frequently used by Swedish writers, although it is seldom employed in this country.

*nebulæ.* The nomenclature is followed in this book with the exception that, since extragalactic nebulæ are mentioned so frequently, the adjective will be dropped. Therefore the term *nebulæ* will refer to extragalactic nebulæ alone, unless otherwise specified.

Some astronomers consider that since nebulæ are now known to be stellar systems they should be designated by some other name, which does not carry the connotation of clouds or mist. Such a revision might be useful, but, as yet, no entirely suitable alternative name has been suggested. The proposal most frequently discussed is a revival of the term *external galaxies.*[14]

The authoritative definition of *galaxy* is *the Milky Way,* and the word, especially in the adjective form, *galactic,* is commonly used in this sense. But a transferred and figurative use has also crept into the literature. The galactic system has sometimes been identified with its most conspicuous feature, and the term *galaxy* employed for the stellar system as a whole. Those who follow the practice commonly refer to other stellar systems as *external galaxies.*

The term is open to certain objections. A purist would say that our own stellar system is the galactic system, but is not the galaxy; that an independent stellar system is neither the one nor the other. Moreover, while the application of a new meaning to an old term is sometimes convenient, the continued use of both meanings is not advisable. Usage, however, is not always determined by logic. The established definition may be dropped and the variant, revived, may flourish. No prediction is ventured.

[14] The term occurs sporadically in the literature of the nineteenth century and had a certain vogue in the more popular astronomies. One example of the latter is *The Architecture of the Heavens,* by J. P. Nichol, which went through many editions after its first appearance in 1838. The ninth edition, 1851, is one of the most interesting. It is dedicated to the Countess of Rosse and presents the early observations with Lord Rosse's six-foot reflector, in glowing terms. A letter is included in which Lord Rosse expresses his conviction that the Orion nebula had been resolved. Nichol states that nebulæ are external galaxies and presents certain globular clusters as conspicuous examples.

The term *nebulæ* offers the values of tradition; the term *galaxies*, the glamour[15] of romance.

## *Designation of Individual Nebulæ.*

INDIVIDUAL nebulæ are generally designated by their numbers in certain catalogues which were made by Messier (1730–1817) and by Dreyer (1852–1926). During the latter half of the eighteenth century, Messier[16] made a list of 103 bright clusters and nebulæ (both galactic and extragalactic), and these conspicuous objects are still known by their Messier numbers. Some thirty-two extragalactic nebulæ are included. The great spiral in Triangulum, for instance, is Messier No. 33, or M33.

Dreyer's New General Catalogue—usually referred to as the NGC—is a compilation of all clusters and nebulæ (galactic as well as extragalactic) which were known at the end of 1887. Two supplements—the Index Catalogues (IC)—bring the lists up to the end of 1907.[17] Since the numbering in the second supplement is a continuation of that in the first, there is no necessity for distinguishing between the two Index Catalogues. The great majority of the 7840 NGC objects and the 5386 IC objects, are extragalactic nebulæ. The NGC nebulæ, in general, are brighter than the IC nebulæ, and, of course, include the Messier objects. Thus M33 is also known as NGC 598.

Since 1907, the numbers of nebulæ recorded on photographs have increased so rapidly that the compilation of general catalogues is neither practical nor important.

[15] The *Oxford English Dictionary* gives the established definition of *galaxy* ("a luminous band or track, encircling the heavens . . . ; the Milky Way"), and states that the transferred and figurative use is "now chiefly applied to a brilliant assemblage or crowd of beautiful women or distinguished persons."

[16] Messier's final catalogue was published in the *Connaissance des temps* for 1784. A finding list of the objects, together with appropriate references, is given by Shapley and Davis, *Publications of the Astronomical Society of the Pacific*, 29, 178, 1917.

[17] The NGC is found in the *Memoirs of the Royal Astronomical Society*, 49, 1, 1890. The IC are found in the same publication, 51, 185, 1895, and 59, 105, 1910.

Many lists have been made for special purposes, but only one that covers the entire sky in a homogeneous manner. This latter is the Harvard survey of nebulæ brighter than the thirteenth magnitude,[18] which contains 1,249 objects (1,188 NGC nebulæ, 48 IC nebulæ, and 13 others). Individual, uncatalogued nebulæ are designated by their positions in the sky, or with reference to some object whose coördinates are generally known.

[18] A Survey of the External Galaxies Brighter than the Thirteenth Magnitude, Harvard College Observatory, *Annals*, 88, No. 2, 1932.

CHAPTER I

# THE EXPLORATION OF SPACE[1]

THE exploration of space has penetrated only recently into the realm of the nebulæ. The advance into regions hitherto unknown has been made during the last dozen years with the aid of great telescopes. The observable region of the universe is now defined and a preliminary reconnaissance has been completed. The chapters which follow are reports on various phases of the reconnaissance.

The earth we inhabit is a member of the solar system—a minor satellite of the sun. The sun is a star among the many millions which form the stellar system. The stellar system is a swarm of stars isolated in space. It drifts through the universe as a swarm of bees drifts through the summer air. From our position somewhere within the system, we look out through the swarm of stars, past the borders, into the universe beyond.

The universe is empty, for the most part, but here and there, separated by immense intervals, we find other stellar systems, comparable with our own. They are so remote that, except in the nearest systems, we do not see the individual stars of which they are composed. These huge stellar systems appear as dim patches of light. Long ago they were named "nebulæ" or "clouds"—mysterious bodies whose nature was a favorite subject for speculation.

But now, thanks to great telescopes, we know something of their nature, something of their real size and brightness, and their mere appearance indicates the gen-

[1] This summary of nebular research may be compared with a "progress report" of results obtained up to the end of 1928, published under the same title in *Harper's Magazine* for May, 1929. Some of the material in the earlier report is included in the present summary, with the permission of Harper & Brothers.

eral order of their distances. They are scattered through space as far as telescopes can penetrate. We see a few that appear large and bright. These are the nearer nebulæ. Then we find them smaller and fainter, in constantly increasing numbers, and we know that we are reaching out into space, farther and ever farther, until, with the faintest nebulæ that can be detected with the greatest telescope, we arrive at the frontiers of the known universe.

This last horizon defines the observable region of space. It is a vast sphere, perhaps a thousand million light-years in diameter. Throughout the sphere are scattered a hundred million nebulæ—stellar systems—in various stages of their evolutionary history. The nebulæ are distributed singly, in groups, and occasionally in great clusters, but when large volumes of space are compared, the tendency to cluster averages out. To the very limits of the telescope, the large-scale distribution of nebulæ is approximately uniform.

One other general characteristic of the observable region has been found. Light which reaches us from the nebulæ is reddened in proportion to the distance it has traveled. This phenomenon is known as the velocity-distance relation, for it is often interpreted, in theory, as evidence that the nebulæ are all rushing away from our stellar system, with velocities that increase directly with distances.

### *Receding Horizons.*

THIS sketch roughly indicates the current conception of the realm of the nebulæ. It is the culmination of a line of research that began long ago. The history of astronomy is a history of receding horizons. Knowledge has spread in successive waves, each wave representing the exploitation of some new clew to the interpretation of observational data.

The exploration of space presents three such phases. At first the explorations were confined to the realm of the planets, then they spread through the realm of the stars, and finally they penetrated into the realm of the nebulæ.

The successive phases were separated by long intervals of time. Although the distance of the moon was well known to the Greeks, the order of the distance of the sun and the scale of planetary distances was not established until the latter part of the seventeenth century. Distances of stars were first determined almost exactly a century ago, and distances of nebulæ, in our own generation. The distances were the essential data. Until they were found, no progress was possible.

The early explorations halted at the edge of the solar system, facing a great void that stretched away to the nearer stars. The stars were unknown quantities. They might be little bodies, relatively near, or they might be gigantic bodies, vastly remote. Only when the gap was bridged, only when the distances of a small, sample collection of stars had been actually measured, was the nature determined of the inhabitants of the realm beyond the solar system. Then the explorations, operating from an established base among the now familiar stars, swept rapidly through the whole of the stellar system.

Again there was a halt, in the face of an even greater void, but again, when instruments and technique had sufficiently developed, the gap was bridged by the determination of the distances of a few of the nearer nebulæ. Once more, with the nature of the inhabitants known, the explorations swept even more rapidly through the realm of the nebulæ and halted only at the limits of the greatest telescope.

### *The Theory of Island Universes.*

This is the story of the explorations. They were made with measuring rods, and they enlarged the body of factual knowledge. They were always preceded by speculations. Speculations once ranged through the entire field, but they have been pushed steadily back by the explorations until now they lay undisputed claim only to the territory beyond the telescopes, to the dark unexplored regions of the universe at large.

The speculations took many forms and most of them have long since been forgotten. The few that survived the test of the measuring rod were based on the principle of the uniformity of nature—the assumption that any large sample of the universe is much like any other. The principle was applied to stars long before distances were determined. Since the stars were too far away for the measuring instruments, they must necessarily be very bright. The brightest object known was the sun. Therefore, the stars were assumed to be like the sun, and distances could be estimated from their apparent faintness. In this way, the conception of a stellar system, isolated in space, was formulated as early as 1750. The author was Thomas Wright (1711–86) an English instrument maker and private tutor.[2]

But Wright's speculations went beyond the Milky Way. A single stellar system, isolated in the universe, did not satisfy his philosophical mind. He imagined other, similar systems and, as visible evidence of their existence, referred to the mysterious clouds called "nebulæ."

Five years later, Immanuel Kant (1724–1804) developed Wright's conception in a form that endured, essentially unchanged, for the following century and a half. Some of Kant's remarks[3] concerning the theory furnish an excellent example of reasonable speculation based on the principle of uniformity. A rather free translation runs as follows:

I come now to another part of my system, and because it suggests a lofty idea of the plan of creation, it appears to me as the most seductive. The sequence of ideas that led us to it is very simple and natural. They are as follows: let us imagine a system of stars gathered together in a common plane, like those of the Milky Way, but situated so far away from us that even with the telescope we cannot distinguish the stars composing it; let us assume that its distance, compared to that separating us from the

[2] *An Original Theory or New Hypothesis of the Universe* (London, 1750).

[3] *Allgemeine Naturgeschichte und Theorie des Himmels,* published first in 1755. The passages are found in the First Part.

stars of the Milky Way, is in the same proportion as the distance of the Milky Way is to the distance from the earth to the sun; such a stellar world will appear to the observer, who contemplates it at so enormous a distance, only as a little spot feebly illumined and subtending a very small angle; its shape will be circular, if its plane is perpendicular to the line of sight, elliptical, if it is seen obliquely. The faintness of its light, its form, and its appreciable diameter will obviously distinguish such a phenomenon from the isolated stars around it.

We do not need to seek far in the observations of astronomers to meet with such phenomena. They have been seen by various observers, who have wondered at their strange appearance, have speculated about them, and have suggested sometimes the most amazing explanations, sometimes theories which were more rational, but which had no more foundation than the former. We refer to the nebulæ, or, more precisely, to a particular kind of celestial body which M. de Maupertius[4] describes as follows:

"These are small luminous patches, only slightly more brilliant than the dark background of the sky; they have this in common, that their shapes are more or less open ellipses; and their light is far more feeble than that of any other objects to be perceived in the heavens."

Kant then mentions and rejects the views of Derham that the patches are openings in the firmament, through which the fiery Empyrean is seen, and of Maupertius that the nebulæ are enormous single bodies, flattened by rapid rotation. Kant then continues:

It is much more natural and reasonable to assume that a nebula is not a unique and solitary sun, but a system of numerous suns, which appear crowded, because of their distance, into a space so limited that their light, which would be imperceptible were each of them isolated, suffices, owing to their enormous numbers, to give a pale and uniform luster. Their analogy with our own system of stars; their form, which is precisely what it should be according to our theory; the faintness of their light, which denotes an infinite distance; all are in admirable accord and lead us to consider these elliptical spots as systems of the same order as our own—in a word, to be Milky Ways similar to the one whose con-

[4] *Discours sur les différentes figures des astres* (Paris, 1742).

stitution we have explained. And if these hypotheses, in which analogy and observation consistently lend mutual support, have the same merit as formal demonstrations, we must consider the existence of such systems as demonstrated. . . .

We see that scattered through space out to infinite distances, there exist similar systems of stars [nebulous stars, nebulæ], and that creation, in the whole extent of its infinite grandeur, is everywhere organized into systems whose members are in relation with one another. . . . A vast field lies open to discoveries, and observation alone will give the key.

The theory, which came to be called the theory of island universes,[5] found a permanent place in the body of philosophical speculation. The astronomers themselves took little part in the discussions: they studied the nebulæ. Toward the end of the nineteenth century, however, the accumulation of observational data brought into prominence the problem of the status of the nebulæ and, with it, the theory of island universes as a possible solution.

## *The Nature of the Nebulæ.*

### (*a*) THE FORMULATION OF THE PROBLEM.

A FEW nebulæ had been known to the naked-eye observers and, with the development of telescopes, the numbers grew, slowly at first, then more and more rapidly. At the time Sir William Herschel (1738–1822), the first outstanding leader in nebular research, began his surveys, the most extensive published lists were those by Messier, the last of which (1784) contained 103 of the most conspicuous nebulæ and clusters. These objects are still known by the Messier numbers—for example, the great spiral in Andromeda is M31. Sir William Herschel cata-

[5] The realm of the stars was once known as the "universe of stars" and the term persisted after the isolation of the stellar system was recognized. The multiplication of stellar systems led to the term "Weltinseln"—Island Universes—used in von Humboldt's *Kosmos* (Vol. III [1850]), presumably for the first time. In the familiar English translation by Otté (1855), the word is translated literally as "world islands" (Vol. III, 149, 150). The transition to "island universes" is an obvious step, but the writer has not ascertained the first use of the term.

logued 2,500 objects, and his son, Sir John (1792–1871), transporting the telescopes to the southern hemisphere (near Capetown in South Africa) added many more.[6] Positions of about 20,000 nebulæ are now available, and perhaps ten times that number have been identified on photographic plates. The mere size of catalogues has long since ceased to be important. Now the desirable data are the numbers of nebulæ brighter than successive limits of apparent faintness, in sample areas widely distributed over the sky.

Galileo, with his first telescopes, resolved a typical "cloud"—Præsepe—into a cluster of stars. With larger telescopes and continued study, many of the more conspicuous nebulæ met the same fate. Sir William Herschel concluded that all nebulæ could be resolved into star-clusters, if only sufficient telescopic power were available. In his later days, however, he revised his position and admitted the existence, in certain cases, of a luminous "fluid" which was inherently unresolvable. Ingenious attempts were made to explain away these exceptional cases until Sir William Huggins (1824–1910), equipped with a spectrograph, fully demonstrated in 1864 that some of the nebulæ were masses of luminous gas.

Huggins' results clearly indicated that nebulæ were not all members of a single, homogeneous group and that some kind of classification would be necessary before they could be reduced to order. The nebulæ actually resolved into stars—the star-clusters—were weeded out of the lists to form a separate department of research. They were recognized as component parts of the galactic system, and thus had no bearing on the theory of island universes.

Among unresolved nebulæ, two entirely different types were eventually differentiated. One type consisted of the

[6] Sir John Herschel's general catalogue, representing the first systematic survey of the entire sky to a fairly uniform limit of apparent faintness, was published in 1864, and contained about 4,630 nebulæ and clusters observed by his father and himself, together with about 450 discovered by others. The catalogue was replaced by Dreyer's *New General Catalogue* in 1890.

relatively few nebulæ definitely known to be unresolvable —clouds of dust and gas mingled among, and intimately associated with, the stars in the galactic system. They were usually found within the belt of the Milky Way and were obviously, like the star-clusters, members of the galactic system. For this reason, they have since been called "galactic" nebulæ. They are further subdivided into two groups, "planetary" nebulæ and "diffuse" nebulæ, frequently shortened to "planetaries" and "nebulosities."

The other type consisted of the great numbers of small, symmetrical objects found everywhere in the sky except in the Milky Way. A spiral structure was found in most, although not in all, of the conspicuous objects. They had many features in common and appeared to form a single family. They were given various names but, to anticipate, they are now known as "extragalactic" nebulæ[7] and will be called simply "nebulæ."

The status of the nebulæ, as the group is now defined, was undetermined because the distances were wholly unknown. They were definitely beyond the limits of direct measurement, and the scanty, indirect evidence bearing on the problem could be interpreted in various ways. The nebulæ might be relatively nearby objects and hence members of the stellar system, or they might be very remote and hence inhabitants of outer space. At this point, the development of nebular research came into immediate contact with the philosophical theory of island universes. The theory represented, in principle, one of the alternative solutions of the problem of nebular distances. The question of distances was frequently put in the form: Are nebulæ island universes?

### (*b*) THE SOLUTION OF THE PROBLEM.

THE situation developed during the years between 1885 and 1914; from the appearance of the bright nova in the spiral M31, which stimulated a new interest in the ques-

[7] The term "external galaxies," revived by Shapley, is also widely used, as is a third term "anagalactic" nebulæ, introduced by Lundmark.

tion of distances, to the publication of Slipher's first extensive list of radial velocities of nebulæ, which furnished data of a new kind and encouraged serious attempts to find a solution of the problem.

The solution came ten years later, largely with the help of a great telescope, the 100-inch reflector, that had been completed in the interim. Several of the most conspicuous nebulæ were found to be far beyond the limits of the galactic system—they were independent, stellar systems in extragalactic space. Further investigations demonstrated that the other, fainter nebulæ were similar systems at greater distances, and the theory of island universes was confirmed.

The 100-inch reflector partially resolved a few of the nearest, neighboring nebulæ into swarms of stars. Among these stars various types were recognized which were well known among the brighter stars in the galactic system. The intrinsic luminosities (candle powers) were known, accurately in some cases, approximately in others. Therefore, the apparent faintness of the stars in the nebulæ indicated the distances of the nebulæ.

The most reliable results were furnished by Cepheid variables, but other types of stars furnished estimates of orders of distance, which were consistent with the Cepheids. Even the brightest stars, whose intrinsic luminosities appear to be nearly constant in certain types of nebulæ, have been used as statistical criteria to estimate mean distances for groups of systems.

### *The Inhabitants of Space.*

The nebulæ whose distances were known from the stars involved, furnished a sample collection from which new criteria, derived from the nebulæ and not from their contents, were formulated. It is now known that the nebulæ are all of the same order of intrinsic luminosity. Some are brighter than others, but at least half of them are within the narrow range from one half to twice the mean value, which is 85 million times the luminosity of the sun.

Thus, for statistical purposes, the apparent faintness of the nebulæ indicates their distances.

With the nature of the nebulæ known and the scale of nebular distances established, the investigations proceeded along two lines. In the first place the general features of the individual nebulæ were studied; in the second, the characteristics of the observable region as a whole were investigated.

The detailed classification of nebular forms has led to an ordered sequence ranging from globular nebulæ, through flattening, ellipsoidal figures, to a series of unwinding spirals. The fundamental pattern of rotational symmetry changes smoothly through the sequence in a manner that suggests increasing speed of rotation. Many features are found which vary systematically along the sequence, and the early impression that the nebulæ were members of a single family appears to be confirmed. The luminosities remain fairly constant through the sequence (mean value, $8.5 \times 10^7$ suns, as previously mentioned), but the diameters[8] steadily increase from about 1,800 light-years for the globular nebulæ to about 10,000 light-years for the most open spirals. The masses are uncertain, the estimates ranging from $2 \times 10^9$ to $2 \times 10^{11}$ times the mass of the sun.

## *The Realm of the Nebulæ.*

### (*a*) THE DISTRIBUTION OF NEBULÆ.

INVESTIGATIONS of the observable region as a whole have led to two results of major importance. One is the homogeneity of the region—the uniformity of the large-scale distribution of nebulæ. The other is the velocity-distance relation.

The small-scale distribution of nebulæ is very irregular. Nebulæ are found singly, in pairs, in groups of various sizes, and in clusters. The galactic system is the chief component of a triple nebula in which the Magellanic

[8] The numerical values refer to the main bodies, which, as will be explained later, represent the more conspicuous portions of the nebulæ.

Clouds are the other members. The triple system, together with a few additional nebulæ, forms a typical, small group that is isolated in the general field of nebulæ. The members of this local group furnished the first distances, and the Cepheid criterion of distance is still confined to the group.

When large regions of the sky, or large volumes of space, are compared, the irregularities average out and the large-scale distribution is sensibly uniform. The distribution over the sky is derived by comparing the numbers of nebulæ brighter than a specified limit of apparent faintness, in sample areas scattered at regular intervals.

The true distribution is confused by local obscuration. No nebulæ are seen within the Milky Way, and very few along the borders. Moreover, the apparent distribution thins out, slightly but systematically, from the poles to the borders of the Milky Way. The explanation is found in the great clouds of dust and gas which are scattered throughout the stellar system, largely in the galactic plane. These clouds hide the more distant stars and nebulæ. Moreover, the sun is embedded in a tenuous medium which behaves like a uniform layer extending more or less indefinitely along the galactic plane. Light from nebulæ near the galactic poles is reduced about one fourth by the obscuring layer, but in the lower latitudes, where the light-paths through the medium are longer, the absorption is correspondingly greater. It is only when these various effects of galactic obscuration are evaluated and removed, that the nebular distribution over the sky is revealed as uniform, or isotropic (the same in all directions).

The distribution in depth is found by comparing the numbers of nebulæ brighter than successive limits of apparent faintness, that is to say, the numbers within successive limits of distance. The comparison is effectively between numbers of nebulæ and the volumes of space which they occupy. Since the numbers increase directly with the volumes (certainly as far as the surveys have een carried, probably as far as telescopes will reach), he distribution of the nebulæ must be uniform. In this roblem, also, certain corrections must be applied to the pparent distribution in order to derive the true distribu-ion. These corrections are indicated by the velocity-dis-ance relation, and their observed values contribute to he interpretation of that strange phenomenon.

Thus the observable region is not only isotropic but omogeneous as well—it is much the same everywhere nd in all directions. The nebulæ are scattered at average ntervals of the order of two million light-years or per-aps two hundred times the mean diameters. The pattern night be represented by tennis balls fifty feet apart.

The order of the mean density of matter in space can lso be roughly estimated if the (unknown) material etween the nebulæ is ignored. If the nebular material vere spread evenly through the observable region, the moothed-out density would be of the general order of $0^{-29}$ or $10^{-28}$ grams per cubic centimeter—about one grain f sand per volume of space equal to the size of the earth.

The size of the observable region is a matter of defini-ion. The dwarf nebulæ can be detected only to moderate listances, while giants can be recorded far out in space. There is no way of distinguishing the two classes, and hus the limits of the telescope are most conveniently de-ined by average nebulæ. The faintest nebulæ that have een identified with the 100-inch reflector are at an aver-ge distance of the order of 500 million light-years, and o this limit about 100 million nebulæ would be observable xcept for the effects of galactic obscuration. Near the alactic pole, where the obscuration is least, the longest xposures record as many nebulæ as stars.

### (*b*) THE VELOCITY-DISTANCE RELATION.[9]

THE foregoing sketch of the observable region has been ased almost entirely upon results derived from direct

9 A more extensive, nontechnical discussion of the velocity-distance rela-on by the writer will be found in *Red-Shifts in the Spectra of Nebulæ,*

photographs. The region is homogeneous and the general order of the mean density is known. The next—and last—property to be discussed, the velocity-distance relation, emerged from the study of spectrograms.

When a ray of light passes through a glass prism (or other suitable device) the various colors of which the light is composed are spread out in an ordered sequence called a spectrum. The rainbow is, of course, a familiar example. The sequence never varies. The spectrum may be long or short, depending on the apparatus employed, but the order of the colors remains unchanged. Position in the spectrum is measured roughly by colors, and more precisely by wave-lengths, for each color represents light of a particular wave-length. From the short waves of the violet, they steadily lengthen to the long waves of the red.

The spectrum of a light source shows the particular colors or wave-lengths which are radiated, together with their relative abundance (or intensity), and thus gives information concerning the nature and the physical condition of the light source. An incandescent solid radiates all colors, and the spectrum is *continuous* from violet to red (and beyond in either direction). An incandescent gas radiates only a few isolated colors and the pattern, called an *emission* spectrum, is characteristic for any particular gas.

A third type, called an *absorption* spectrum and of special interest for astronomical research, is produced when an incandescent solid (or equivalent source), giving a continuous spectrum, is surrounded by a cooler gas. The gas absorbs from the continuous spectrum just those colors which the gas would radiate if it were itself incandescent. The result is a spectrum with a continuous background interrupted by dark spaces called absorption lines. The pattern of dark absorption lines indicates the

being the Halley Lecture delivered at Oxford University in 1934. Some of the material in the lecture is included in the present summary, with the permission of the Clarendon Press.

particular gas or gases that are responsible for the absorption.

The sun and the stars give absorption spectra and many of the known elements have been identified in their atmospheres. Hydrogen, iron, and calcium produce very strong lines in the solar spectrum, the most conspicuous being a pair of calcium lines in the violet, known as H and K.

The nebulæ in general show absorption spectra similar to the solar spectrum, as would be expected for systems of stars among which the solar type predominated. The spectra are necessarily short—the light is too faint to be spread over long spectra—but the H and K lines of calcium are readily identified and, in addition, the G-band of iron and a few hydrogen lines can generally be distinguished (Plates VII and VIII).

Nebular spectra are peculiar in that the lines are not in the usual positions found in nearby light sources. They are displaced toward the red of their normal position, as indicated by suitable comparison spectra. The displacements, called red-shifts, increase, on the average, with the apparent faintness of the nebula that is observed. Since apparent faintness measures distance, it follows that red-shifts increase with distance. Detailed investigation shows that the relation is linear.

Small microscopic shifts, either to the red or to the violet, have long been known in the spectra of astronomical bodies other than nebulæ. These displacements are confidently interpreted as the results of motion in the line of sight—radial velocities of recession (red-shifts) or of approach (violet-shifts). The same interpretation is frequently applied to the red-shifts in nebular spectra and has led to the term "velocity-distance" relation for the observed relation between red-shifts and apparent faintness. On this assumption, the nebulæ are supposed to be rushing away from our region of space, with velocities that increase directly with distance.

Although no other plausible explanation of red-shifts has been found, the interpretation as velocity-shifts may be considered as a theory still to be tested by actual observations. Critical tests can probably be made with existing instruments. Rapidly receding light sources should appear fainter than stationary sources at the same distances, and near the limits of telescopes the "apparent" velocities are so great that the effects should be appreciable.

*The Observable Region as a Sample of the Universe.*

A COMPLETELY satisfactory interpretation of red-shifts is a question of great importance, for the velocity-distance relation is a property of the observable region as a whole. The only other property that is known is the uniform distribution of nebulæ. Now the observable region is our sample of the universe. If the sample is fair, its observed characteristics will determine the physical nature of the universe as a whole.

And the sample may be fair. As long as explorations were confined to the stellar system, the possibility did not exist. The system was known to be isolated. Beyond lay a region, unknown, but necessarily different from the star-strewn space within the system. We now observe that region—a vast sphere, through which comparable stellar systems are uniformly distributed. There is no evidence of a thinning-out, no trace of a physical boundary. There is not the slightest suggestion of a supersystem of nebulæ isolated in a larger world. Thus, for purposes of speculation, we may apply the principle of uniformity, and suppose that any other equal portion of the universe, selected at random, is much the same as the observable region. We may assume that the realm of the nebulæ is the universe and that the observable region is a fair sample.

The conclusion, in a sense, summarizes the results of empirical investigations and offers a promising point of departure for the realm of speculation. That realm, dominated by cosmological theory, will not be entered in the present summary. The discussions will be largely restricted to the empirical data—to reports of the actual explorations—and their immediate interpretations.

Yet observation and theory are woven together, and it is futile to attempt their complete separation. Observations always involve theory. Pure theory may be found in mathematics but seldom in science. Mathematics, it has been said, deals with possible worlds—logically consistent systems. Science attempts to discover the actual world we inhabit. So in cosmology, theory presents an infinite array of possible universes, and observation is eliminating them, class by class, until now the different types among which our particular universe must be included have become increasingly comprehensible.

The reconnaissance of the observable region has contributed very materially to this process of elimination. It has described a large sample of the universe, and the sample may be fair. To this extent the study of the structure of the universe may be said to have entered the field of empirical investigation.

CHAPTER II

# FAMILY TRAITS OF NEBULÆ

## *The Classification of Nebulæ.*

THE first chapter presented a brief sketch of the current conception of the realm of the nebulæ. The sample available for inspection is a vast region of space through which comparable stellar systems are more or less uniformly distributed. The subject will now be developed systematically and in greater detail.

The first step is obviously a study of the apparent features of the systems under investigation. The nebulæ might be members of a single family or they might represent a mixture of utterly different kinds of objects. The question is very important for all investigations of a general nature. The nebulæ are so numerous that they cannot all be studied individually. Therefore, it is necessary to know whether a fair sample can be assembled from the more conspicuous objects and, if so, the size of the sample required. The answer to this question, and to many others, is sought in the classification of nebulæ.

The problem is essentially a photographic one, for the nebulæ are faint and structural details are extremely difficult to see. Visual observations, even with the largest telescopes, are less satisfactory than photographs made with moderate-size cameras. Photographs with large telescopes, of course, are correspondingly more informative.

The simplest procedure is to sort out the nebulæ, by inspection of photographs, into groups of objects showing similar features. The more conspicuous members of each group can then be studied in detail and the results used for the comparison of the groups themselves. The degree of success attained by the method depends largely upon the significance of the features selected as the basis of classification.

To a certain degree, the choice of these criteria represents a compromise. The features must be significant—they must indicate physical properties of the nebulæ themselves and not chance effects of orientation—and also they must be conspicuous enough to be seen in large numbers of nebulæ. Millions of nebulæ are within reach of existing telescopes, but relatively few are sufficiently large and bright for detailed investigation. Inconspicuous features, although they might be highly significant, would restrict the classification to a small number of nebulæ which might not be a fair sample.

Numbers of nebulæ increase rapidly with decreasing brightness, and the great majority are recorded on the photographic plates as mere formless specks, barely distinguishable from images of faint stars. These objects in general are beyond the limits of any useful classification. There are large numbers of nebulæ, somewhat brighter but still so small and faint that no details are discernible except elongation and concentration (luminosity-gradient or rate at which the luminosity fades from the center toward the edge of an image). Classifications have been based on these features, but depend largely on chance orientation. When the criteria are applied to conspicuous, well-known nebulæ, their significance appears to be slight.

### *The Common Pattern.*

THE current classifications were derived from several hundred of the bright nebulæ, on the assumption that the collection was sufficiently large to constitute a fair sample of nebulæ in general. These objects were sorted into groups, each representing an assemblage of characteristic features. The groups fell naturally into an ordered sequence with the criteria varying systematically from one end to the other. Many characteristic features could be described in the brightest nebulæ, but as fainter and still fainter nebulæ were examined, the features were gradually lost until finally the most conspicuous alone could be recognized. These last surviving criteria constitute the basis of the formal classification. Two such systems have been developed, but as they are rather similar, only one will be presented here in detail.[1] This classification reveals a common, fundamental pattern, whose continuous variation produces the observed sequence of nebular forms.

As a first step the nebulæ are divided into two very unequal groups. The great majority are called "regular nebulæ," since they exhibit, as a common pattern, conspicuous evidence of rotational symmetry about dominating, central nuclei. The remaining objects, about 2 or 3 per cent of the total number, are called "irregular," because they lack both rotational symmetry and, in general, dominating nuclei.

Regular nebulæ are either "elliptical nebulæ" or "spirals." Objects in each group fall naturally into ordered sequences of structural forms; and one end of the elliptical sequence is rather similar to one end of the spiral sequence. Accordingly, the two sequences are oriented, for purposes of description, as though they were two sections of a single, larger sequence containing all structural forms encountered among the regular nebulæ. The zero-point is arbitrarily selected at the free end of the elliptical section. The progression throughout the complete sequence thus runs from the most compact of the elliptical nebulæ to the most open of the spirals—a progression in dispersion or expansion. The terms "early" and "late" are used to denote relative position in the empirical sequence without regard to their temporal implications. These explanations emphasize the purely empirical na-

[1] Hubble, *Extra-Galactic Nebulæ*, "Mt. Wilson Contr.," No. 324; *Astrophysical Journal*, 64, 321, 1926. See also, Lundmark, *A Preliminary Classification of Nebulæ*, "Upsala Meddelanden (Arkiv för Mat., Astr. och Fysik)," 19b, No. 8, 1926.

ture of the sequence of classification. The consideration is important because the sequence closely resembles the line of development indicated by the current theory of nebular evolution as developed by Sir James Jeans.[2]

### *Elliptical Nebulæ.*[3]

ELLIPTICAL nebulæ are designated by the symbol *E*. They range from globular objects through ellipsoidal figures to a limiting, lenticular form with a ratio of axes about 3 to 1. It is probable that all regular nebulæ with main bodies flatter than this limiting form are spirals. Elliptical nebulæ are highly concentrated and show no indications of resolution into stars. The luminosity falls rapidly away from bright, semistellar nuclei to undefined boundaries. As far as exposures have been pushed, diameters, and hence total luminosities, increase steadily with increasing exposure-time. Small patches of obscuring material are occasionally silhouetted against the luminous background, but otherwise these nebulæ present no structural details.

The only general features by which the elliptical nebulæ can be further classified are (*a*) the shapes of the images, or, more precisely, the shapes of the isophotal contours (contours of equal luminosity); and (*b*) the luminosity-gradients. The gradients are difficult to estimate on a numerical scale and their measurement requires an elaborate technique. For these reasons they are not suitable as criteria for rapid classification.

The shapes of contours are readily estimated by simple inspection, but they refer, of course, to the projected images as seen on the photographic plates and not to the actual 3-dimensional nebulæ. Circular contours might

2 The most recent statement of the theory is in Jeans, *Astronomy and Cosmogony* (1928), XIII.

3 See Hubble, *Distribution of Luminosity in Elliptical Nebulæ,* "Mt. Wilson Contr.," No. 398; *Astrophysical Journal,* 71, 231, 1930; a rediscussion of these data by ten Bruggencate, *Zeitschrift für Astrophysik,* 1, 275, 1930; Smith, *Some Notes on the Structure of Elliptical Nebulæ,* "Mt. Wilson Contr.," No. 524; *Astrophysical Journal,* 82, 192, 1935.

# PLATE I

## *Types of Nebulæ (Elliptical and Irregular)*

*Elliptical Nebulæ.* The sequence of regular nebulæ is composed of two sections, one containing *Elliptical* nebulæ and the other, *Spirals.* Elliptical nebulæ range from globular bodies, E0, through flattening, ellipsodial figures to the limiting, lenticular form, E7. Nebulæ flatter than E7, are spirals.

The luminosity, in elliptical nebulæ, fades smoothly away from semi-stellar nuclei to undetermined boundaries, and the isophotal contours (contours of equal luminosity) are, in general, similar ellipses. Therefore, the size of an image grows with increasing exposure but the shape remains approximately constant.

E7 nebulæ are lenticular bodies seen on edge. Less flattened images, En, might represent true figures of any form between En and E7, which are properly oriented in space. The real existence of all forms from globular to lenticular is inferred from the frequency distribution of the ellipticities of the projected images.

*Irregulars.* Regular nebulæ are characterized by rotational symmetry around dominating central nuclei. About one nebula in forty is *irregular* in the sense that one or both characteristics are absent. The Magellanic Clouds are conspicuous examples of irregulars, and are similar to the object NGC 4449, which appears on the Plate.

represent globular nebulæ or any of the flatter forms whose polar axes happen to fall in the line of sight. Only when the flattest (lenticular) nebulæ are seen on edge, do the projected images indicate the true figures. The uncertainty is serious, but unavoidable. Except for the one case, no method is known by which the individual true figures can be determined.[4] Nevertheless, it is possible, by statistical analysis of the shapes of many projected images (assuming random orientation), to investigate the existence and the relative frequencies of the various true forms, without knowing those of individual objects. The analysis indicates that the true forms do range from globular to lenticular and that the latter are more frequent than the former.

Under these circumstances, a provisional classification has been based on the contours of projected images. The contours are ellipses[5] and in any single nebula, they are all similar. In other words, the shapes of images remain unchanged as the images grow with increasing exposure time.

Ellipticity is defined as $(a-b)/a$ where $a$ and $b$ are respectively the major and the minor diameters. Position in the sequence is very simply indicated by estimates of the ellipticity to one decimal, the decimal point being omitted. Thus circular contours (apparently globular nebulæ such as NGC 3379) are designated as E0; M32, the brighter satellite of M31, with a ratio of axes about 5 to 4, is E2; lenticular nebulæ such as NGC 3115, are E7. There the series stops—E8 or higher would doubtless refer to a spiral on edge mistaken for an unusually thin, lenticular object. The latter forms may possibly occur, but if so they must be very rare.

[4] Luminosity-gradients along the major axes may be more or less independent of orientation, but the possibility is a subject for investigation and not a present basis for classification.

[5] Contours in lenticular nebulæ deviate very slightly from ellipses—they are pointed at the extremities—but the deviations are so small that they can be disregarded in the discussion of elliptical nebulæ as a group.

# PLATE II

## *Types of Nebulæ (Normal and Barred Spirals)*

THE sequence of spirals is double, one branch containing normal spirals, the other containing barred spirals. Each branch is subdivided into three sections, designated by the subscripts, *a, b,* and *c,* to the symbols S and SB, and described as early, intermediate and late, normal or barred spirals. Positions in each branch are determined by the amount of material in the arms relative to that in the nuclear regions, by the openness of the arms and by the degree of resolution. The early spirals ($S_a$ and $SB_a$) appear to be closely related to the lenticular (E7) nebulae. The transition from E7 to $SB_a$ is smooth and continuous but that from E7 to $S_a$ may be cataclysmic—all known examples of $S_a$ have fully developed arms.

The complete sequence of regular nebulae, E0 to $S_c$, exhibits many features which vary systematically throughout the entire range. Total luminosities (absolute magnitudes) remain approximately constant but diameters increase and hence surface brightness diminishes; colors, spectral types, and degree of resolution, which characterize the nebular contents, also change continuously. The galactic system is presumably a late type normal spiral.

### *Spirals.*

SPIRAL nebulæ fall into two distinct but parallel sequences, containing *normal* and *barred* spirals, designated as S and SB, respectively. A thin scattering of mixed forms lies between the series. In the normal spiral, the two arms emerge smoothly from opposite segments of the periphery of a nuclear region resembling a lenticular nebula, and thence wind outward along spiral paths. In the barred spiral, the two arms spring abruptly from either end of a bar of nebulosity stretching diametrically across the nuclear region, and thence follow spiral paths similar to those found in the normal spirals. Normal spirals are more frequent than barred forms in the ratio of 2 or 3 to 1.

### *Normal Spirals.*

AT the beginning of the sequence, the normal spiral exhibits a bright, semistellar nucleus and a relatively large nuclear region of unresolved nebulosity which closely resembles a lenticular (E7) nebula. The arms which emerge from the periphery are also unresolved and are closely coiled. As the sequence progresses, the arms increase in bulk at the expense of the nuclear region, unwinding as they grow, until in the end they are widely open and the nucleus is inconspicuous. About the middle of the sequence, or slightly earlier, condensations begin to form. The resolution generally appears first in the outer arms and gradually spreads inward until, at the end of the sequence, it reaches the nucleus.

### *Barred Spirals.*[6]

THE barred spiral is first seen as a lenticular nebula in which the outer regions have condensed into a more or less conspicuous ring of nebulosity concentric with the nucleus, and a broad bar has condensed diametrically

**6 Attention was first directed to the barred spirals by Curtis, *Publications of the Lick Observatory,* 13, 12, 1918.**

across the nucleus from rim to rim. The appearance resembles that of the Greek letter *theta*, $\theta$. As the sequence progresses, the ring appears to break away from the bar at two opposite points, just above the bar at one end and just below the bar at the other—thus ϴ—and the spiral arms grow out of the free ends of the broken ring. Thereafter the development parallels that of the normal spiral; the arms build up at the expense of the nuclear region, unwinding as they grow; resolution appears first in the outer arms and works inward toward the nucleus. The last stage is the familiar *S*-shaped spiral with thin, well-resolved arms (NGC 7479).

## *Sequence of Spirals.*

PROGRESSION in the two series is fairly indicated by the relative luminosity of nuclear region and spiral arms, by openness of the arms and by the degree of resolution. The last of these criteria, of course, cannot be used in the case of faint, distant spirals, but the other two are generally applicable and, in principle, they are independent of orientation. Spirals seen on edge cannot always be placed with precision, but they can be assigned with some confidence to general regions of the sequence.

Provisionally, each sequence of spirals has been subdivided into three sections designated by the subscripts *a, b,* and *c.* Thus $S_a$, $S_b$, and $S_c$ represent, *early, intermediate,* and *late* types of normal spirals and $SB_a$, $SB_b$, and $SB_c$, represent the corresponding types of barred spirals. Sections are supposed to cover equal ranges in the sequences, but divisions have not been precisely specified and the classification of border-line objects is rather arbitrary. Such objects are sometimes indicated by the combined subscripts *ab* and *bc,* and nebulæ intermediate between E7 and $S_a$ are occasionally designated as S0.

The general trend of the progression appears to be well established, but refinements may be expected as the subject develops. Reynolds,[7] for instance, has introduced

[7] "A Classification of Spiral Nebulæ," *Observatory,* 50, 185, 1927.

the terms "massive" and "filamentary," denoting spirals with broad, bulky arms (M33), and with thin, stringy arms (M101). The distinction may depend upon total mass or some other inherent characteristic of the nebulæ and if such an interpretation can be established, the terms will be highly significant as well as descriptive.

*Sequence of Regular Nebulæ.*

SINCE early-type spirals, $S_a$ and $SB_a$, resemble lenticular nebulæ (E7) in many respects, the complete sequence of regular nebulæ may be represented by a diagram shaped like the letter *Y*, or, since the spiral series are roughly parallel, like a tuning fork (Fig. 1). Elliptical nebulæ

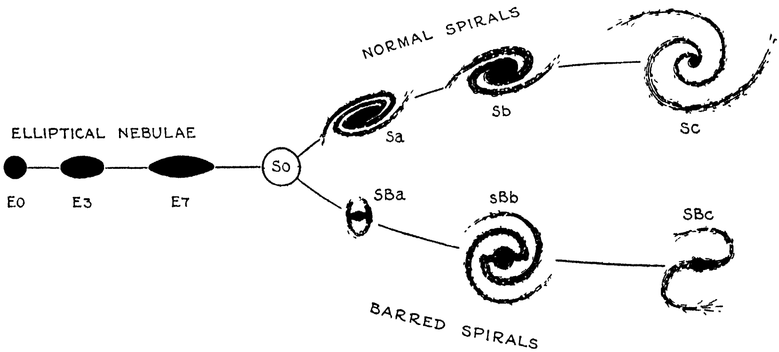



FIG. 1. *The Sequence of Nebular Types.*

The diagram is a schematic representation of the sequences of classification. A few nebulæ of mixed types are found between the two sequences of spirals. The transition stage, S0, is more or less hypothetical. The transition between E7 and $SB_a$ is smooth and continuous. Between E7 and $S_a$, no nebulæ are definitely recognized.

form the stem, with globular objects (E0) at the base and lenticular forms (E7) just below the junction. Normal and barred spirals lie along the two arms, and between the arms a few spirals of mixed types are found. The latter usually exhibit barred characteristics in small regions immediately surrounding the nuclei, but otherwise resemble normal spirals—M83 and M61 are examples.

The junction may be represented by the more or less hypothetical class S0—a very important stage in all theories of nebular evolution. Observations suggest a smooth transition between E7 and $SB_a$, but indicate a discontinuity between E7 and $S_a$ in the sense that $S_a$ spirals are always found with arms fully developed. Speculation concerning the discontinuity will be unprofitable until more detailed information has been assembled from large-scale photographs. At present, the suggestion of cataclysmic action at this critical point in the evolutional development of nebulæ is rather pronounced.

### *Additional Characteristics.*

Before leaving the description of regular nebulæ, a few additional features may be mentioned. Nuclei in general are semistellar, and too small for thorough investigation by photographic methods. In very late-type spirals (M33), where the nuclei stand out more or less conspicuously against the relatively faint nebulosity in the immediate region, the nuclei resemble globular clusters. In rare cases the nuclei appear stellar in all direct tests that have been applied. Such nuclei in general are relatively bright and exhibit emission spectra similar to the spectra of planetary nebulæ ($N_2$ brighter than Hβ). Thus, regardless of their appearance, they cannot be considered as single stars in the ordinary sense of the term.

Spiral structures appear to be embedded in faint, unresolved nebulosity which can often be traced well beyond the main bodies of the nebulæ. Obscuration plays a conspicuous rôle. Peripheral bands of obscuring material, presumably dust or gas, are frequently found in the earlier types of normal spirals, $S_a$ and $S_b$, and are seen in silhouette when the nebulæ are oriented nearly edge-on (NGC 4594). These bands have not been found among the early barred spirals. Obscuration in definite patches, ranging widely in dimensions, is especially conspicuous in the later types, but more so in the normal than in the barred spirals. The patches are presumably similar to obscuring clouds in the galactic system, and comparisons of angular diameters were once used as rough indications of the order of distances of spirals.[8]

Occasional nebulæ show unusual features and their precise positions in the sequence of classification are rather uncertain. Such objects are placed according to the judgment of the investigator, and the letter *p* (for peculiar) is added to the symbol of classification. The device is required for perhaps 2 per cent of the regular nebulæ, and is more frequently applied to elliptical nebulæ than to spirals. The fainter companion of M31, and the companion of M51, both of which are classed as Ep, may be mentioned as examples.

### *Irregular Nebulæ.*

OTHER nebulæ, between 2 and 3 per cent of the total number, show no evidence of rotational symmetry and hence do not find a place in the sequence of classification. These objects are called irregular nebulæ and are designated as Ir. About half of the irregulars form a homogeneous group, in which the Magellanic Clouds are typical examples, and their importance probably merits a separate division.[9] Since their stellar contents resemble those of very late-type spirals, they are sometimes considered as representing the last stage in the sequence of regular nebulæ. Their status, however, is speculative, and the absence of conspicuous nuclei may be of more fundamental significance than the absence of rotational symmetry, which is a possible consequence.

The remaining irregulars might be arbitrarily placed in the regular sequence as highly peculiar objects, rather

[8] The most extensive discussion of obscuring material in spirals is that by Curtis, *Publications of the Lick Observatory*, Vol. XIII (1918), to which the reader is referred for further information.

[9] Lundmark's classification does include a division—Magellanic Nebulæ—for these objects.

than in a separate class. Some, such as NGC 5363 and 1275, could be described as elliptical nebulæ which have disintegrated without developing spiral structures. Others, such as M82, are merely nondescript. Almost all of them require individual consideration but, in view of their very limited numbers, they can be neglected in preliminary surveys of nebular forms.

### *The Standard Nebula.*[10]

Nebulæ at any given stage in the regular sequence are constructed on a fairly uniform pattern. They present not only similar structures but also a constant, mean surface-brightness. Some are large and bright, others are small and faint; in appearance they resemble a single, standard nebula placed at various appropriate distances. This conclusion is derived from the observed fact that total luminosity, on the average, varies directly with the square of the major diameter. Now the square of the diameter measures the area of a nebular image if the nebula is oriented face-on (polar axis in the line of sight). For such nebulæ, the relation

$$\text{Luminosity} = \text{Constant} \times (\text{diameter})^2$$

indicates constant, mean surface-brightness. Moreover, since the nebulæ are moderately transparent, total luminosities are, in the same degree, independent of orientation. Hence, if major diameters of projected images are used consistently, the above relation holds approximately for all nebulæ, regardless of foreshortening.

The relation is expressed in astronomical units as

$$m + 5 \log d = C$$

where $m$ is total apparent magnitude, $d$ is apparent angular diameter, say in minutes of arc, and $C$, the sum, is constant for nebulæ at a particular stage in the sequence. By means of this relation all nebulæ at a given stage can be reduced to a standard apparent magnitude and the

10 Hubble, *Extra-Galactic Nebulæ.*

dispersion in diameters can then be examined, or vice versa.

When standard nebulæ, all reduced to a given apparent magnitude, are established at various stages along the sequence, it is found that the diameters increase steadily from globular nebulæ to open spirals (Fig. 2). This statement is another way of saying that $C$ increases systematically throughout the sequence. Once the law of variation is established, it is possible to reduce all the regular nebulæ (and, in a statistical sense, the irregular nebulæ as well) to a given stage in the sequence—for instance, to the junction point, S0—and to discuss them as a single, homogeneous group. This procedure emphasizes the convenience, as well as the significance, of the classification

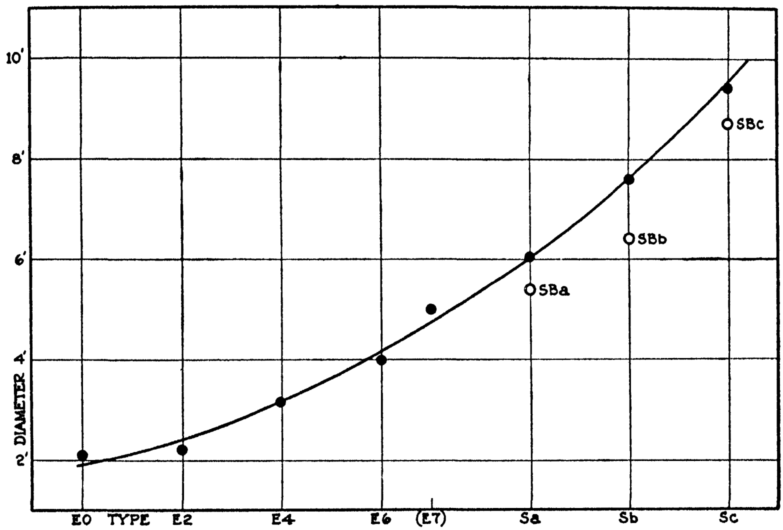



FIG. 2. *Nebular Diameters along the Sequence of Types* (m = *10*).

For nebulæ of the same apparent luminosity, the diameters increase steadily from the globular nebulæ to the open spirals. The diagram shows the mean diameters (of the main bodies) in minutes of arc, for nebulæ of the tenth apparent magnitude. The horizontal scale is arbitrary.

and permits the application of quantitative methods so desirable in all investigations.

In actual practice, certain difficulties are encountered. Diameters and luminosities are very arbitrary quantities, depending upon exposures and methods of measurement. It is for this reason that the term "main body" has been employed, referring to the portion of a nebula which is readily visible on simple inspection of well-exposed photographs. The main body must be distinguished from the much larger area over which the nebula can be traced with laborious photometric methods.

The diameter-luminosity relation and the variation of $C$ throughout the sequence, may be expected to emerge from any homogeneous set of data. The numerical results, however, depend upon the particular set of data employed. The first investigation, for instance, which included about 400 nebulæ, was based upon Holetschek's visual magnitudes, and diameters of main bodies estimated by simple inspection of hour-exposures on fast plates with large reflectors. Improved results, derived by combining these diameters with estimated photographic magnitudes listed in the Harvard Survey[11] of bright nebulæ, are given in Table II. The data represent the general pattern of the variation, but the numerical values refer to a particular set of conditions and require appropriate revisions when other conditions are specified.

## *Features which Vary Systematically through the Sequence.*

### (*a*) SPECTRAL TYPES.[12]

OTHER characteristics which vary systematically through the sequence are spectral types, colors, and among the

[11] "A Survey of the External Galaxies Brighter than the Thirteenth Magnitude," Harvard College Observatory, *Annals*, 88, No. 2, 1932.

[12] The classification of nebular spectra is largely the work of Humason. Most of the data are found in "Mt. Wilson Contrs.," No. 426 and No. 531; *Astrophysical Journal*, 74, 35, 1931, and 83, 10, 1936.

spirals, the upper limits of stellar luminosity (intrinsic luminosities of brightest stars). Spectral types are known for the nuclear regions of about 150 nebulæ. The solar type, early G, predominates, although an occasional K or F is found. Dwarf characteristics are conspicuous in the few spectra which have been recorded on considerable scales, and hence they are provisionally assumed to be

TABLE II.

*Diameter-Luminosity Relation.*

| Type | $C$* | Type | $C$ |
|---|---|---|---|
| E 0 | 11.4 | $S_a$ | 13.9 |
| 1 | 11.6 | $S_b$ | 14.4 |
| 2 | 11.9 | $S_c$ | 14.9 |
| 3 | 12.2 | $SB_a$ | 13.7 |
| 4 | 12.5 | $SB_b$ | 14.0 |
| 5 | 12.8 | $SB_c$ | 14.7 |
| 6 | 13.1 | Irr | 14.0 |
| 7 | 13.4 | | |

* $C = m + 5 \log d$, where $m$ is the total apparent magnitude of a nebula and $d$ is the angular diameter in minutes of arc.

normal features of nebular spectra. In the best-known spectra, those of the nuclear regions of M31 and M32, the relative intensities of the absorption lines correspond to those found in spectra of stars (dG3) whose absolute magnitudes are about +4.3. Such stars are closely comparable with the sun.

The relatively rare "stellar" nuclei with emission spectra resembling spectra of planetary nebulæ, have already been mentioned. The significance of these spectra is unknown, but in view of their rarity they may be neglected in a preliminary survey of the field. Another type of emission spectrum is fairly common in irregular nebulæ and in the outer regions of open spirals. These spectra are localized in isolated patches within the nebulæ and resemble the spectra produced by certain diffuse nebulosities (clouds of gas in the neighborhood of very hot stars, for example, the Orion nebulosity) in the galactic system. The phenomena represent one of the many analogies between nebulæ and our own stellar system.

When emission spectra are disregarded, the mean type of the nuclear spectra varies systematically through the sequence of classification, ranging from about G4 for the early, elliptical nebulæ to about F9, or slightly earlier, for the open spirals. The range is small but is quite definitely established, since the dispersion in the correlation is also small. The mean type of all available absorption spectra is about dG3.

### (*b*) COLORS.

SPECTRAL types are derived from absorption lines without regard to the distribution of the continuous spectra. Colors, on the other hand, represent the distribution of the continuous spectra without regard to absorption lines. Among stars, there is a definite relation between colors and spectral types, and when one is known the other can be inferred with some confidence. Departures from the normal relation are termed color-deficiencies and color-excesses. The latter are by far the more common, especially in the region of the Milky Way, and are generally interpreted as due to selective absorption by diffuse, interstellar material.

Among the nebulæ, however, the normal relation between color and spectral type appears only in the open spirals. Globular nebulæ exhibit a conspicuous color excess of the order of 0.3 mag., and the excess diminishes along the sequence until it vanishes at $S_c$.

The cause of the phenomena is unknown, although the variation is well established. The colors (for about 80 nebulæ) represent very precise measures through blue and yellow filters made with photoelectric cells by Stebbins and his associates directly at the foci of the large reflectors on Mount Wilson.[13] Other methods give com-

[13] These data are not yet published, but provisional values were available by the courtesy of Professor Stebbins.

parable, although less precise, results for a wide range of additional nebulæ.

The color-excesses exhibit no appreciable dependence upon galactic latitude nor upon apparent magnitude. The entire range is observed in the Virgo cluster alone (at latitude +75°). The source of the color-excess must, therefore, be assigned to the nebulæ themselves and not to intervening diffuse material either within or without the galactic system. Diffuse material within the nebulæ might be responsible, but this suggestion raises certain difficulties which have not yet been satisfactorily explained.

The observational data are summarized in Table III, where mean spectral types by Humason, and mean color classes (on the scale of giant stars) by Stebbins, are given for various stages in the sequence of classification.

TABLE III.

*Spectral Types and Colors of Nebulæ.*

| Nebular Type | Spectral Type | Color Class |
|---|---|---|
| E0–E9 | G4 | g6 |
| $S_a$, $SB_a$ | G3 | g5 |
| $S_b$, $SB_b$ | G2 | g4 |
| $S_c$, $SB_c$ | F9 | f7 |

### (*c*) RESOLUTION.

PHOTOGRAPHS of many of the conspicuous spirals and irregular nebulæ show numerous knots and condensations which are now known to represent individual stars and groups of stars. The identification of stars in nebulæ, which was of fundamental importance since it led directly to the determination of distances, will be described at some length in Chapter IV. The subject is mentioned at this point because it is simpler and more significant to discuss "resolution into stars" rather than "resolution into condensations."

Resolution first appears among the earlier, intermediate spirals, probably about $S_{ab}$. Thereafter the stars become increasingly conspicuous. Lack of resolution among earlier types does not necessarily mean that stars are absent; it merely indicates that, if stars are present, the brightest of them are fainter than the brightest stars in the later spirals. Thus it is not impossible that all nebulæ are composed of stars, but that the upper limit of stellar luminosity increases systematically through the sequence of classification, passing above the observable threshold in the vicinity of the stage $S_{ab}$.

The hypothesis is admittedly speculative, but it finds some support in the observed luminosities of nebulæ and stars in the neighboring Virgo cluster. This cluster is a compact group of several hundred nebulæ among which all types are represented (except the irregulars). The mean luminosities of the different types are of the same general order, but the stars are systematically brighter in the $S_c$ than in the $S_b$ spirals, and no stars at all are found in the $S_a$ spirals. These data further suggest that increasing stellar luminosities may compensate for the fading of unresolved nebulosity, leaving the total luminosities of the nebulæ relatively constant.

The fact that the combined photographic luminosity of stars and unresolved nebulosity is fairly constant, is intimately associated with diminishing color-excess through the sequence. The brightest stars in open spirals are blue, and are presumably O-type supergiants such as those observed in the galactic system and in the Magellanic Clouds. Seares,[14] in 1922, found that the outer arms of open spirals, the regions where resolution is most conspicuous, are bluer (have smaller color-indices) than the nuclear regions. No interpretation of the phenomenon was offered at the time, but later, when the condensations were identified as stars, it seemed probable that the color-

14 "Preliminary Results on the Color of Nebulæ," *Proceedings of the National Academy of Sciences*, 2, 553, 1916. See also another discussion by Seares in *Publications of the Astronomical Society of the Pacific*, 28, 123, 1916.

effect was due to blue (early-type) stars. Finally, when the colors of individual stars were measured and found to be blue, the interpretation seemed established. Since no differential distribution of color is found in elliptical nebulæ, the systematic decrease in color along the sequence appears to be definitely associated with the progressive development of blue giants.

### (*d*) RELATIVE FREQUENCY OF TYPES.

FINALLY, there appears to be a systematic increase in the relative frequency, or numbers of nebulæ, along the sequence of classification. The great clusters, which are dominated by the early types, do not conform to the rule. In the general field of isolated nebulæ, however, the increasing frequency is found in all large collections of nebulæ which are reasonably complete or representative, to a definite limit of apparent magnitude. The only essential requirement is that the classification must be made from photographs on a scale sufficient to avoid effects of selection. Such effects generally operate in favor of early types at the expense of later types.

The first tabulation according to the present classification was made from Holetschek's list of about 400 nebulæ observed from a northern latitude. A more comprehensive summary may be derived from the Harvard Survey of bright nebulæ covering the entire sky. The relative frequencies for 600 nebulæ with types estimated from large-reflector plates, is given in Table IV. The increasing

TABLE IV.

*Relative Frequencies of Nebular Types.*

| Type | Frequency (Per Cent) |
|---|---|
| E0–E7 | 17 |
| $S_a$, $SB_a$ | 19 |
| $S_b$, $SB_b$ | 25 |
| $S_c$, $SB_c$ | 36 |
| Irr | 2.5 |

frequency along the sequence is conspicuous among the spirals, where effects of orientation are not serious.

Elliptical nebulæ, except the lenticular forms, E7, cannot be treated individually since the actual forms cannot be differentiated from effects of orientation in the projected images. An E0 image, for instance, might represent any figure of revolution oriented with the axis in the line of sight. In general, an E$n$ image might represent any nebula whose actual ellipticity is equal to, or greater than, $n$, and which is appropriately oriented. The frequency-distribution of actual ellipticities is a statistical problem and it can be readily solved once the distribution of projected ellipticities is known from the observations. The solution involves the law of orientation of nebular axes, and in practice the reasonable assumption of random orientation is chosen.

The problem has been discussed by various investigators and the results are not entirely consistent. The data are rather scanty for isolated nebulæ, among which the ellipticals are relatively rare. Larger lists can be assembled in the clusters, where ellipticals predominate, but their interpretation is confused by variations in average type from one cluster to another. It seems clear, however, that globular nebulæ are relatively rare as compared with lenticular systems, and that numbers increase along the sequence with increasing ellipticity.

### *Summary.*

The discussion of relative frequencies completes the preliminary examination of the objects under investigation. Order has emerged from the apparent confusion and the planning of further research is greatly simplified. The study of nebular forms, as they appear on direct photographs, leads to the conclusion that nebulæ are closely related members of a single family. They are constructed on a fundamental pattern which varies systematically through a limited range. The nebulæ fall naturally into an ordered sequence of structural forms and are readily reduced to a standard position in the sequence. At such a standard stage, the relation between apparent size and brightness is just that to be expected if the same nebula could be examined from various distances. The dispersion in the apparent characteristics of nebulæ is remarkably small. Therefore, any considerable collection of nebulæ, chosen at random, should be a fair sample. Results of detailed studies of conspicuous objects can be applied to nebulæ in general. And broad statistical investigations may be conducted with some assurance that the material is homogeneous.

CHAPTER III

# THE DISTRIBUTION OF NEBULÆ

## *Nebular Surveys.*

THE distribution of nebulæ, as well as their classification, can be studied in a significant way with no knowledge of actual distances. The distribution is derived from extensive surveys. The important numerical data are the numbers of nebulæ brighter than various limits of apparent faintness. As the limits progress—as the surveys penetrate into greater and still greater depths of space—the numbers of nebulæ increase rapidly and the broad features of the distribution become steadily more pronounced. Surveys to a single limit give the distribution over the sky, while the distribution in depth is found by comparing surveys to successive limits. Thus two problems are presented, and the really significant results are achieved at the very faint limits which furnish the largest numbers of nebulæ.

The interpretation of the data is a statistical problem whose intricacies will not be presented in detail. The principle of the method is simple, however, and may be mentioned briefly before the results of the investigations are described. Suppose that real distances were unknown. Nevertheless, if the nebulæ were all of the same intrinsic luminosity (same absolute magnitude or candle power), the relative distance of each individual object would be indicated by its apparent faintness. Positions in space could then be mapped, on some arbitrary scale, and the general features of the distribution would be clearly revealed.

The actual problem was complicated by the probability (now known as a fact) that the nebulæ were not equally luminous. There might be (and there are) giant nebulæ and dwarf nebulæ and all intermediate grades. Thus apparent faintness alone would not determine relative distances, and maps such as those mentioned above would be misleading.

The difficulty was met in a very simple way. Instead of plotting individual nebulæ, the averages of large groups were mapped. Absolute magnitudes of individual objects might be scattered over a considerable range, but the means of large groups, chosen at random, should be fairly constant. Upon this simple principle are based the powerful statistical methods of investigating the distribution, although they are developed to a degree which permits the simultaneous consideration of all possible groupings of the data.

In studying the distribution, one important assumption was made. Consider all the nebulæ *in a given volume of space.* The relative numbers of giants and dwarfs and normal objects—more precisely, the frequency-distribution of absolute magnitudes (candle powers) among these nebulæ—form the "luminosity-function." It was assumed that the luminosity-function remains constant throughout the regions covered by the surveys—that the function is independent of distance or direction—that the giants do not tend to congregate in one region, and the dwarfs in another region. The assumption has not been fully established by direct observations, but it seems reasonable and it is consistent with all information available at the present time. It will be implied in the chapters which follow, even when it is not specifically mentioned.

On the basis of this assumption, the mean absolute magnitude for a large group of neighboring nebulæ should be much the same as the mean magnitude for a group of distant nebulæ. In a statistical sense, apparent faintness would measure relative distances, although absolute distances were unknown. Nothing was postulated concerning total numbers of nebulæ per unit volume of space. The possible variation throughout the explored region (the density-function) was a problem to be investigated, and the survey would lead to a solution provided the luminosity-function was constant.

The surveys furnished data that were selected on the basis of apparent luminosity. Now it is an interesting fact, which may be mentioned in passing, that the mean candle power of a group of nebulæ selected in this manner —for instance, nebulæ of the fifteenth apparent magnitude—is not the same as the mean candle power of nebulæ in a given volume of space. The two quantities are related, it is true, but the relation involves the precise form of the luminosity-function. The subject will be discussed later, when the effects of selection must be evaluated. It is mentioned here merely as a preface to the statement that apparent magnitudes measure relative distances (in a statistical sense) regardless of the precise form of the luminosity-function, provided the form is constant. Thus nebulæ of the fifteenth magnitude average ten times nearer than nebulæ of the twentieth magnitude, and ten times more distant than those of the tenth magnitude. In general

$$\log d_1/d_2 = 0.2\ (m_1 - m_2)$$

where $d$ and $m$ are mean distances and mean apparent magnitudes of any two groups of nebulæ, selected on the basis of $m$ (as they are in the surveys).

The data available for investigations of distribution are of various kinds. The brighter, conspicuous nebulæ are known individually, although accurately measured magnitudes are rare. The most extensive list with usable magnitudes is the Harvard Survey[1] which is believed to be complete over the entire sky, down to and including the limiting magnitude, $m = 12.9$. Nebulæ to brighter limits can be extracted from this list.

[1] Shapley and Ames, "A Survey of the External Galaxies Brighter than the Thirteenth Magnitude," Harvard College Observatory, *Annals*, 88, No. 2, 1932. Extensive counts of faint nebulæ also have been made at Harvard but, except in a few instances, the data have not yet been published in sufficient detail to permit an examination of the important features of completeness and homogeneity.

More or less complete surveys are still in progress with cameras which record large areas of the sky on single plates, but penetrate only to moderate depths. As larger cameras are used, the area per plate diminishes while the depth of penetration increases. Thus one camera may record the constellation of Orion on a single plate and register nebulæ down to $m = 13$, another may record the bowl of the Big Dipper and reach $m = 16$, and a third may record the bowl of the Little Dipper and reach $m = 18$.

Very faint limits are reached only with large telescopes, which record very small areas on single plates but penetrate to great depths. The usable field of the 100-inch, for instance, is about equal to the area of the full moon. Completeness over the sky is impractical and the surveys proceed on the principle of sampling. Plates are centered mainly on selected areas, uniformly distributed, and the regions are assumed to be fair samples of the entire sky. The deepest survey for which extensive, detailed results are available, made with the large reflectors on Mount Wilson, will be discussed at some length, since it furnishes the most general picture of the background of faint nebulæ.[2]

### *Distribution over the Sky.*

The survey included 1,283 separate samples rather uniformly scattered over 75 per cent of the sky. Two telescopes, the 60-inch and the 100-inch, were used indiscriminately under varying conditions. Then, by corrections largely derived from the data themselves, the numbers of nebulæ counted on the plates (about 44,000 in all) were changed to numbers (about 80,000) representing standard conditions. The limiting magnitude of the corrected counts was $20.0 \pm 0.1$ in the central region of the plates.

Analysis of this homogeneous material indicated that the large-scale distribution of nebulæ over the sky was

[2] Hubble, *The Distribution of Extra-Galactic Nebulæ,* "Mt. Wilson Contr.," No. 485; *Astrophysical Journal,* 79, 8, 1934.

approximately uniform, except for the effects of obscuration arising within the galactic system (Fig. 3).

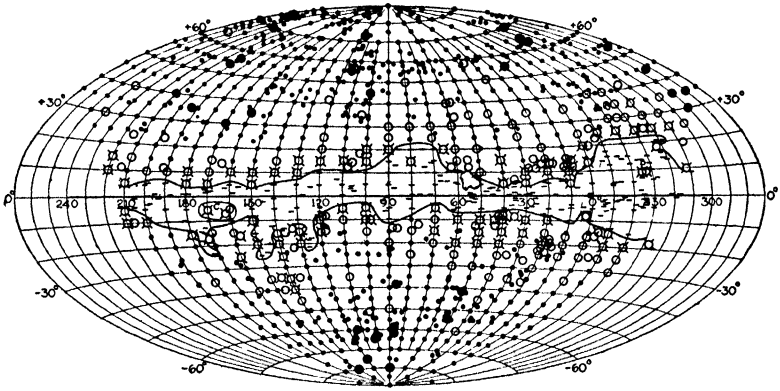



FIG. 3. *Apparent Distribution of Nebulæ, Showing Effects of Galactic Obscuration.*

Positions of samples are plotted in galactic coördinates. The horizontal line, 0°—0°, represents the central plane of the Milky Way. The north galactic pole is at the top. Small dots indicate normal numbers of nebulæ per sample; large disks and circles, excesses and deficiencies, respectively; dashes, samples in which no nebulæ were found.

The zone of avoidance along the Milky Way (regions where dashes predominate) is bordered by partial obscuration (fringes of open circles), beyond which the distribution is approximately uniform.

The blank spaces at the extreme right and left of the diagram represent the southern skies which cannot be observed from the station at which the survey was made.

## *Galactic Obscuration.*

THE evidence for galactic obscuration is as follows.

(*a*) No nebulæ are found along the central region of the Milky Way. The zone of more or less complete avoidance is irregular and unsymmetrical, the width varying from 10° to 40°.

(*b*) Outside the zone of avoidance and the bordering fringe of partial obscuration, the numbers of nebulæ per plate increase with galactic latitude in a manner which closely approximates the trend of a cosecant law (similar to the manner in which stars brighten as they rise from the horizon toward the zenith and are seen through diminishing air-paths).

*Obscuring Clouds.*

Dark, obscuring clouds, ranging from insignificant wisps to vast bodies a hundred light-years and more across, are scattered throughout the stellar system. Some are sensibly opaque, others are semitransparent, and still others are barely perceptible as thin veils. They concentrate markedly toward the galactic plane and are most numerous along the Milky Way, where they are seen as silhouettes against the background of more distant stars. They are most conspicuous in the direction of the center of the galactic system, where they hide the nucleus, but they are found in all directions and, piling up one behind another, effectively hide the rim of the system as well.[8] The dark pattern of this obscuration designs most of the apparent structure of the Milky Way, and isolates many of the so-called star-clouds.

The dark clouds may consist of matter in all forms, but in the sensibly opaque or barely transparent clouds, the major part of the obscuration must be caused by dust. By no other form of material can the obscuration be explained without attributing impossibly large masses to the clouds. Moreover, the stars, and in particular the globular clusters, when they are veiled by heavy, partial obscuration in the clouds, exhibit pronounced color-excesses which indicate selective absorption such as dust would produce. The lighter clouds may have the same composition with smaller densities, or they may be predominantly gaseous.

The apparent distribution of these clouds has much the

[8] The galactic system is a highly flattened swarm of stars—presumably a late-type spiral nebula—in rapid rotation about a polar axis. The sun is near the fundamental plane (the galactic plane), but far out from the center. The system probably has a nucleus, but it cannot be observed because dark clouds lie between the sun and the galactic center.

same pattern as the nebular zone of avoidance—a narrow belt concentric with the galactic plane, from which several great flares sweep out into higher latitudes. Thus the zone of avoidance with its fringe of partial obscuration is readily explained by the presence of obscuring clouds within the stellar system itself. The avoidance is not necessarily complete. Occasional, faint nebulæ are found which are seen through semitransparent paths between sensibly opaque clouds. The object IC 10, at a latitude of $-3°$, is a conspicuous example in which only a portion of a presumably large spiral nebula is visible.[4]

### *The Absorbing Layer.*

Diffuse material in the galactic system has frequently been discussed as though it consisted of a uniform layer of constant depth, centered on the galactic plane. Since the bulk of the obscuration is due to the isolated clouds, such treatment, even as a rough approximation, is likely to be misleading. Disregarding the clouds, however, there is considerable evidence of a tenuous medium, producing some slight absorption, which does approximate the behavior of a uniform layer. This medium may permeate the entire main body of the galactic system, or it may represent a flattened, lenticular cloud so vast that the variations in depth, as observed from the earth, are scarcely perceptible. In either case, the first-order effects would be much the same.

The evidence for such an absorbing layer, derived from the nebular survey in regions free from obscuring clouds, is very clear. The average numbers of nebulæ per plate are largest in the regions of the galactic poles (perpendicular to the plane of the Milky Way), where obscuration due to a uniform layer should be least. From the poles toward the galactic plane, the numbers per plate decrease with latitude in a manner which indicates obscuration proportional to the length of the light path

[4] Mayall, "An Extra-Galactic Object 3° from the Plane of the Galaxy," *Publications of the Astronomical Society of the Pacific*, 47, 317, 1935.

through a uniform layer. In other words, the obscuration, expressed in magnitudes, is C × cosecant β, where C is the obscuration at the pole and β is the latitude. The relation is indicated by the highly simplified diagram in Fig. 4.

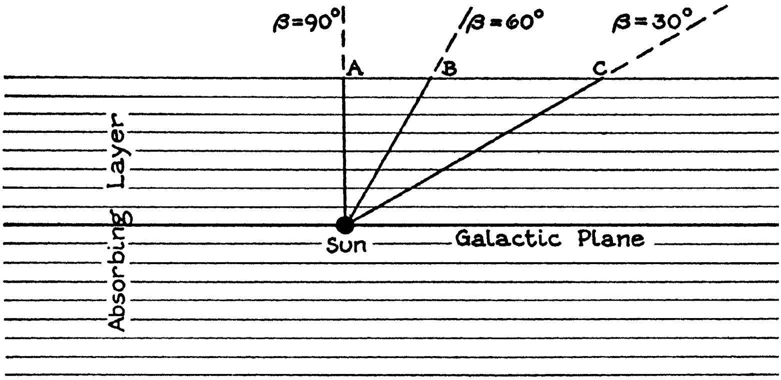



FIG. 4. *Schematic Representation of the Absorbing Layer.*

As observed from the vicinity of the sun, extragalactic objects are obscured according to the lengths of the light-paths through the absorbing layer. The obscuration is least for objects seen in the direction of the galactic poles and increases as the latitudes diminish.

Since the obscuration at $\beta = 30°$ is just twice that at $\beta = 90°$, the difference indicates the actual obscuration at the pole itself. It is about 0.25 mag., and thus the total "optical thickness" of the layer is about 0.5 mag. Numbers of nebulæ per plate are corrected for the latitude effect, and reduced to a homogeneous system representing uniform obscuration equal to that at the galactic pole by the relation (Fig. 5)

$$\log N = \log N_\beta + 0.15 \operatorname{cosec} \beta.$$

Except in the very low latitudes, the data from the surveys give the picture of a uniform layer of tenuous material, extending indefinitely. But near the galactic plane certain systematic departures are found which indicate that obscuration may be greater in the direction of the galactic center than in the direction of the anti-center. The discrepancies are not very important, but they suggest the possibility that the source of the obscuration may

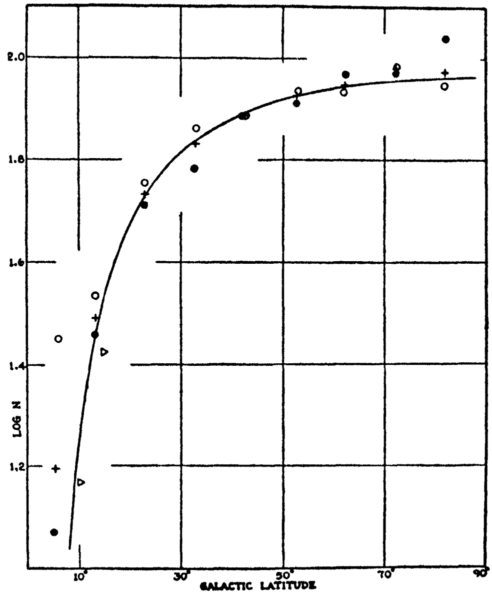



FIG. 5. *Obscuration by the Absorbing Layer as Shown by the Apparent Distribution of Nebulæ.*

The average number of nebulæ per unit area (brighter than a given limiting magnitude) increases with the galactic latitude of the area, in a manner which closely follows the cosecant law,

$$\log N = \text{constant} - 0.15 \text{ cosecant } \beta.$$

Circles and disks represent data from the northern and the southern galactic hemispheres, respectively; crosses are means of the two. The pair of triangles represent supplementary data in low latitudes.

be more reasonably pictured as a highly flattened, lenticular cloud rather than as an indefinitely extended, uniform layer of material. The sun, in this picture, would lie near the median plane of the cloud, but well away from the center. The obscuration would thus be greater in some directions than in others. A few such clouds are actually observed as dim, lens-shaped silhouettes extending for many degrees along the central plane of the Milky Way.

The absorption in the tenuous layer (or cloud), differs from the absorption in the opaque or semitransparent clouds in being essentially nonselective. All colors are absorbed in the same degree (within the uncertainties of the measures), and there is no appreciable variation of nebular colors with galactic latitude. In marked contrast are the color-excesses of globular clusters and early type stars in low latitudes, measured by Stebbins and his associates.[5] These color-effects exhibit a decided correlation with latitude. The more conspicuous cases, however, are within the nebular zone of avoidance, and are rather intimately associated with known, obscuring clouds. The fundamental correlations might well be with positions in the clouds rather than with the latitude.

Investigations of galactic obscuration are still in the formative period. The material includes obscuring clouds, the diffuse medium, and the unknown source of stationary lines in stellar spectra. Preliminary discussions naturally tend to throw these all together and to smooth the various effects into statistical uniformity. Later developments will doubtless emphasize departures from homogeneity, and in this connection the distinction between selective and nonselective absorption has considerable significance.

## *The General Field.*

THE importance of local obscuration in the investigation of nebular distribution is obvious. We are in the midst of obscuring material, and its effects must be eliminated before the true distribution is revealed. The sky may be roughly divided into the galactic belt (latitudes −40° to +40°) and the polar caps (latitudes 40° to 90°). The ga-

[5] Stebbins, "Absorption and Space Reddening in the Galaxy as Shown by the Colors of Globular Clusters," *Proceedings of the National Academy of Sciences*, 19, 222, 1933.

lactic belt, which contains the zone of avoidance and the flares and the bordering fringe of partial obscuration, gives information primarily concerning local obscuration. The polar caps, free from the major effects of local obscuration, give information primarily concerning the distribution of nebulæ.

As the areas of the polar caps are relatively small, the distribution within the caps might not represent the distribution over the sky as a whole. But it is possible, when the effects of the absorbing layer have been removed (Fig. 6), to derive additional information by following the

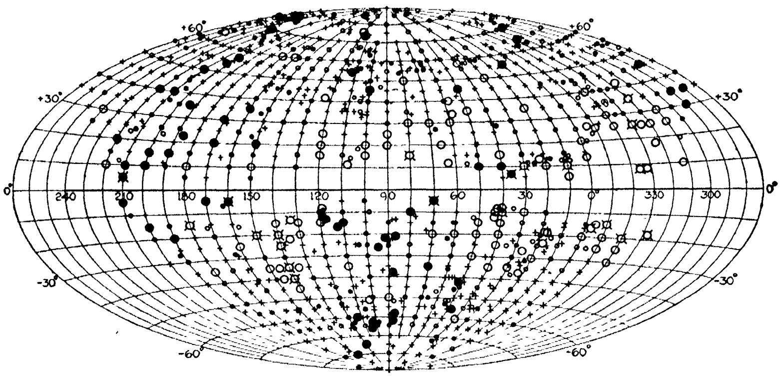



FIG. 6. *Distribution of Nebulæ When Observations Are Partially Corrected for Galactic Obscuration.*

The orientation of the diagram is similar to that in Fig. 3. The nebular counts have been corrected for the latitude effect (absorbing layer), but the effects of clouds, whether opaque or semitransparent, are still present.

Crosses indicate samples with normal numbers of nebulæ (log $N$ differing from the average by not more than 0.15); small disks and circles, moderate excesses and deficiencies; large disks and circles, considerable excesses and deficiencies. Except for the influence of obscuring clouds, there is no evidence of conspicuous systematic variation in the distribution of nebulæ over the sky.

general field of nebulæ down into the galactic belt, between the great flares from the zone of avoidance with their bordering fringes of partial obscuration. In this way the distribution of nebulæ may be traced, over a considerable range in longitude, down to latitudes as low as 15°. The data are individually less precise than those for the polar caps alone, but the general results are thoroughly consistent over as much of the field as can be safely explored. These results are as follows:

(*a*) The two polar caps, northern and southern, are similar, and the distributions agree within the uncertainties of the data.

(*b*) There are no appreciable, systematic variations in distribution within the general field.

(*c*) Individual values of log $N$, where $N$ is the number of nebulæ per plate, are distributed at random about the mean value of them all (frequency-distribution of log $N$ approximates a normal error-curve).

*Large-Scale Distribution over the Sky.*

From the agreement between the two polar caps or, more broadly, between the two galactic hemispheres, it follows that the sun is close to the median plane of the absorbing layer of diffuse material. This plane, moreover, is close to the galactic plane. The agreement between the polar caps, together with the absence of systematic variations either in longitude or in latitude, may be summed up in the statement that the large-scale distribution over the sky is approximately uniform. In technical terms, the distribution is isotropic—the same in all directions.

The conclusion is drawn from only a fraction of the sky. The zone of avoidance with its bordering fringes withdraws large areas from possible exploration and, in addition, about 25 per cent of the sky cannot be efficiently observed from the station at which the survey was made. However, the areas investigated include both galactic poles, the whole of the northern cap, 60 per cent of the southern cap, and the less obscured portions of perhaps two thirds of the galactic belt. The extent and the pattern of these regions would appear to constitute a fair sample of the sky as a whole, and the complete absence of appreciable, systematic variations strongly suggests that no significant, major departures from isotropy should be expected in the unobserved regions.

### *Large-Scale Distribution in Depth.*

DISTRIBUTION in depth is indicated by the rate at which numbers of nebulæ increase with apparent faintness—in other words, by the comparison of various groups of nebulæ with the volumes of space through which the groups are distributed. If the luminosity-function is independent of distance, then, in a statistical sense, apparent magnitudes measure relative distances. Therefore, the number of nebulæ brighter than a given magnitude, $N_m$, represents the number within a sphere of a particular radius. The comparison of numbers contained in spheres of successive radii—more generally, the form of the relation between $N_m$ and $m$—gives the distribution in depth.

Uniform distribution, since the numbers of nebulæ are then proportional to the volumes of space over which the counts extend, is expressed by the simple relation[6]

$$\log N_m = 0.6\, m + \text{constant.}$$

The relation is closely approximated even by casual observational data. For this reason, serious investigations of distribution in depth are largely confined to the search for minor departures from uniformity, and to the precise evaluation of the constant. When the luminosity-function is known, the constant determines the quantitative distribution—the number of nebulæ per unit volume of space.

Preliminary results, establishing approximate uniformity in the immediate neighborhood of the galactic system, were first derived from counts of the brighter nebulæ, for

[6] Let $d$ represent the distance corresponding to apparent magnitude $m$; let $V$ represent the volume of the sphere with radius $d$, and let $C$, with or without subscripts, represent various constants. Then

$$\begin{aligned} \log d &= 0.2\, m + C_1 \\ N_m &= C_2 V = C_3 d^3 \\ \log N_m &= 3 \log d + C_4 \\ &= 0.6\, m + C \end{aligned}$$

which roughly estimated magnitudes were available. Similar results, extending to very faint limits, were derived later by plotting numbers of nebulæ per photographic plate against exposure times of the plates. This relation could be transformed into a relation between $N_m$ and $m$ by means of the known rate at which the limiting magnitudes on the plates varied with exposure times.

The data now available represent several surveys with large reflectors to well-determined, limiting magnitudes ranging from 18.5 to 21. The detailed results will be discussed later in connection with small apparent departures from uniformity, which are interpreted as effects of red-shifts on apparent luminosity. When appropriate corrections for such effects are made, the data indicate uniform distribution (within the small uncertainties of the investigation) out to the practical limits to which surveys can be carried with existing telescopes. The results are summarized in the relation

$$\log N_m = 0.6\ (m - \Delta m) - 9.09 \pm 0.01$$

where $N_m$ is the number of nebulæ per square degree and $\Delta m$ is the effect of red-shifts at magnitude $m$.

In contrasting the uniform distribution of the nebulæ with the thinning out of stars in the stellar system, the power of the telescope approaches the spectacular. The stars form an isolated system, and the star-density diminishes steadily from the nucleus to the border. Therefore, for an observer within the system, the number of stars brighter than a given limiting magnitude, increases with the magnitude, but the rate of increase steadily diminishes. The phenomenon is especially pronounced in the directions of the galactic poles. In these directions, the distances to the boundaries of the stellar system are shortest and the total numbers of stars in the line of sight are least.

At moderately bright, limiting magnitudes, the stars per square degree greatly outnumber the nebulæ, and the rate of increase is nearly the same for both. As successively fainter limits are observed, the nebulæ maintain a constant rate of increase while the rate for stars steadily diminishes. Eventually, the total number of nebulæ per square degree, approaches the total number of stars. In the regions of the galactic poles, equality may be expected at magnitude 21.5 or thereabouts, which is approximately the extreme limit at which nebulæ can be identified under favorable conditions with the 100-inch reflector. The expectations are realized, since the maximum effective exposures in very high latitudes record fully as many recognizable nebulæ as stars. The fact is, as has been said, a rather spectacular tribute to the power of the telescope.

The uniform distribution in depth clearly shows that the nebulæ are not members of the galactic system. The only assumption involved is the reasonable one that the luminosity-function of nebulæ does not vary with distance in a manner which precisely compensates the counts for effects of falling density. Such a variation would be highly artificial and improbable. Thus the realm of the nebulæ, even without further information concerning the scale on which it is constructed, appears as a definite entity, quite distinct from the realm of the stars.

### *Small-Scale Distribution.*

THE small-scale distribution of nebulæ, as derived from the variations among small samples, is conspicuously nonuniform. Nebulæ are found both singly, and in groups of various sizes up to the occasional, great compact clusters of several hundred members each. Only when large samples are compared does the tendency to cluster average out and the distribution approximate uniformity.

The great clusters, which are relatively rare, were excluded from the surveys and will be described later. The present discussion is restricted to the mixture of isolated nebulæ and small groups. The characteristics of small-scale variation, as derived from a particular survey, depend on the average number of nebulæ per sample—in other words, upon the average volume of space per sample, regardless of whether the volume represents a large angular area extending to a moderate depth, or a small area extending to a great depth.

In the survey to the twentieth magnitude, the average sample in the polar cap was of the order of forty-five nebulæ per plate actually identified before reduction to standard conditions. The reduction of the counts and their transformation to numbers per unit area, confuse the simple representation of the variations among samples, but one result is clear and significant.

If the nebulæ were individually distributed at random, the numbers per sample, $N$, would be scattered almost symmetrically about the average, $\bar{N}$. Actually, the observed scatter is unsymmetrical in the sense that there is a conspicuous excess of small samples. The frequency-distribution of $N$ follows a distorted, or "skew," curve.

A small part of the distortion is introduced by the limitations of the data and by the unavoidable errors of observation and reduction. The remainder is presumably associated with the tendency of nebulæ to cluster. An aggregation of nebulæ drawn from the general field would result in one extra-large sample and several extra-small samples. Such a process, operating on a sufficient scale, would account qualitatively for the skew distribution of samples in the survey. Assuming a tendency to cluster, the unsymmetrical distribution of $N$ (if $N$ is moderately small) would follow as a matter of course.

Now it is well known that, in a frequency-distribution of this sort (excess of small samples or positive skewness), the substitution of log $N$ for $N$, tends to restore the symmetry. The remarkable feature of the nebular survey is the fact that the substitution completely eliminates the skewness and precisely restores the symmetry. The results are shown in Fig. 7, where frequency-curves are plotted for both $N$, and log $N$, per sample. The frequency-distribution of log $N$ closely approximates a normal error-curve, and is fully described by the mean value, $\overline{\log N}$, together with the spread or, more precisely, the dispersion, σ.

This feature appears to be a general characteristic of nebular distribution. It is found in all surveys where definite, limiting magnitudes have been consistently main-

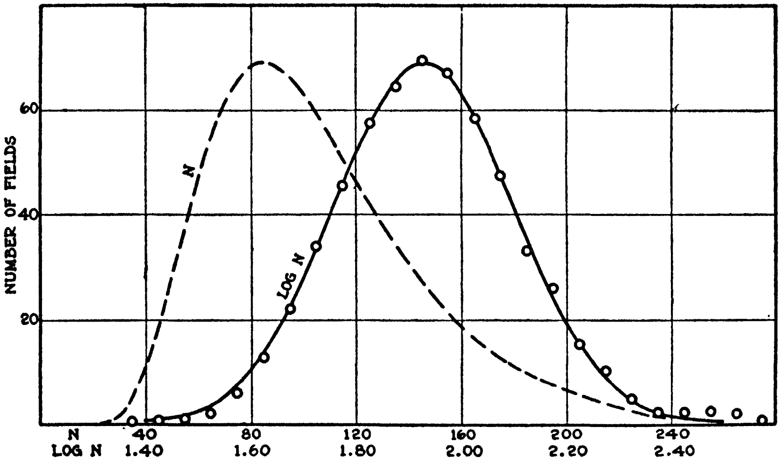



FIG. 7. *Frequency-Distribution of Samples of Various Sizes.*

The circles indicate numbers of samples in which the logarithms of the numbers of nebulæ have various values. The smooth curve drawn through the circles is a normal error-curve. The close fit of the error-curve indicates that the log $N$ per sample are distributed at random around the mean log $N$ for all the samples.

When simple numbers of nebulæ per sample are substituted for the logarithms, the frequency distribution follows the unsymmetrical curve indicated by the dashes.

tained and galactic obscuration has been considered. In each case, the log $N_m$ per sample ($m$ is the limiting magnitude) follow a normal error-curve and the survey is described by the two quantities, $\overline{\log N_m}$ and σ. As the limiting magnitude, $m$, increases (becomes fainter), $\overline{\log N_m}$ increases and σ diminishes. Eventually, for surveys using very large samples, the dispersions represent little more than the incidental errors of investigation. These surveys tend to conform with the theory of random sampling of a homogeneous population, and hence the large-scale distribution of nebulæ is said to be statistically uniform.

A technical point may be mentioned in passing, since it emphasizes the precision with which the scatter of log $N$ per sample is represented by the normal error-curve. Each survey is characterized by $\overline{\log N_m}$ and $\sigma$. However, for the comparison of different surveys, the significant data are not the means of log $N_m$, namely $\overline{\log N_m}$, but the means of $N_m$, namely $\overline{N_m}$. These latter data represent the average numbers of nebulæ per unit area, and hence furnish the total numbers in the sky brighter than the particular limiting magnitudes of the surveys. It is by correlating $\overline{N_m}$, or, for convenience, log $\overline{N_m}$, with $m$, that the distribution of nebulæ in depth, is investigated.

Now the two quantities, $\overline{\log N}$ and log $\bar{N}$ (dropping the subscripts, $m$), have a very simple geometrical relation if the frequency-distribution of the individual log $N$ is a normal error-curve. Then,[7]

$$\log \bar{N} = \overline{\log N} + 1.152\ \sigma^2$$

and the two sides of the equation can be calculated independently from the data for any survey. Such calculations have been made for the five available surveys with large reflectors. A comparison of the two sides of the equation shows an average difference of 0.002 in the logarithms or about half of 1 per cent in the numbers. The result, as has been said, emphasizes the precision with which the condition is fulfilled.

The restoration of symmetry by the transformation to logarithms is so exact that it suggests a property of numbers rather than a property of nebular distribution. However, no satisfactory explanation has been found in the mathematical manipulations, so the latter alternative

[7] This useful relation was brought to the writer's attention by Professor R. C. Tolman of the California Institute of Technology.

seems highly probable. The feature thus serves as a description and a measure of the tendency to cluster.

It is clear that the groups and clusters are not superposed on a random (statistically uniform) distribution of isolated nebulæ, but that the relation is organic. Condensations in the general field may have produced the clusters, or the evaporation of clusters may have populated the general field. Equations describing the observed distribution can doubtless be formulated on either assumption, and, when solved, should contribute significantly to speculations on nebular evolution.

The tendency to cluster appears to operate on a limited scale. No organizations on a scale larger than the great clusters, and no clusters with *as many as* a thousand members, are definitely known. Actually, the maximum population of clusters may be considerably smaller. Samples which, on the average, are large compared with a single cluster, should tend to conform with the theory of random sampling. The frequency-distributions of $N$ per sample should approximate normal error-curves. Smaller samples should give unsymmetrical frequency-distributions of $N$, with an excess of thinly populated fields. In this sense, the average samples in the available surveys are small and the positive skewness is observed.

The largest average sample—that found in the survey to the faintest limit, $m = 21$—is about 200 nebulæ actually identified per plate. The corresponding dispersion in log $N$ is small, $\sigma = 0.084$, and the effects of observational errors are presumably of about the same order as the true dispersion in the nebular distribution. Under these conditions the skewness is not very conspicuous. With still larger average samples, it would probably be inappreciable.

## *Groups of Nebulæ.*

Groups and clusters are important in the study of nebular characteristics as well as in the study of distribution, for each represents a sample collection of objects all of which are at the same distance. Although the distance of a group may be unknown, the relative, apparent dimensions of the members faithfully represent the relative, absolute dimensions.

Double and triple nebulæ are numerous. The whirlpool nebula, M51, is a double. The great spiral in Andromeda, M31, with its two companions, M32 and NGC 205, is a triple system, and the galactic system, with the Magellanic Clouds as companions, is a similar case. All types of nebulæ are represented in such systems, and thus they furnish information concerning the relative dimensions at various stages of the sequence of classification. Moreover, once their distances are known, these very compact systems offer opportunities for deriving the order of nebular masses from statistical investigations of radial velocities. The methods involved are similar to those used for determining masses of double stars from the orbital motions of the components.

Larger groups of nebulæ are also encountered, analogous to the sparser open clusters of stars. The galactic system is a member of such a group and its neighboring fellow members were the first nebulæ whose distances were determined. The most reliable criterion of distance, namely, the study of Cepheid variables, is still confined almost entirely to this local group. Small groups appear to be more numerous than large groups, but the precise manner in which the frequency varies with population has not been determined. Pending definite information, it is supposed that the frequency diminishes as the population increases, over the whole range of groups and loose clusters to the great clusters themselves.

## *Clusters of Nebulæ.*

The nomenclature of the clusters is still arbitrary, and in these discussions the term "cluster" will be restricted to the great clusters alone. The term "group" will be used for all the lesser organizations. The clusters are relatively rare. About twenty are actually known, although

# PLATE III

## *Group of Nebulæ (NGC 3185, 3187, 3190, 3193).*

The group (in the constellation Leo; galactic long., 180°; lat., +56°) exhibits a variety of nebular types—E2 (3193), $S_a$ (3190), $SB_{ab}$ (3185), and $SB_c$ (3187). The apparent magnitudes range from about 12 to 13.5, and the average is 12.65, or, when corrected for local obscuration (latitude effect), about 12.6.

The mean intrinsic luminosity (candle power) of this small sample collection is presumably about the same as that of nebulæ in general. The absolute magnitude of the latter nebulæ (see chap. vii) is $M_0 = -14.2$. Therefore, the distance of the group in Leo, as indicated by the modulus, $m - M = 26.8$, is about 7.5 million light-years.

Humason has measured a red-shift corresponding to a radial velocity of 810 miles/sec. (1300 km/sec.), in one member of the group (3193). This velocity, corrected for the solar motion (see chap. v), indicates a distance of seven million light-years (see chap. vii), in fair agreement with that derived from the luminosities.

The plate was made with the 100-inch reflector, December 24, 1935; north is at the top of the page; 1 mm. = 5″.7.

fragmentary data suggest that in surveys to the twentieth magnitude, perhaps one cluster per fifty square degrees may be expected.

In appearance the clusters are remarkably similar. Each consists of perhaps five hundred members on the average, scattered through a range of about five magnitudes. The form of the frequency-distribution of magnitudes (relative numbers of giants, dwarfs, and normal nebulæ) is difficult to determine, but it appears to be symmetrical about the mean or most frequent magnitude, and roughly to approximate a normal error-curve. The brighter branches of the frequency-curves are the more reliable, since the bright, giant nebulæ stand out conspicuously among the fainter nebulæ in the general field. These branches are so similar in the different clusters that the magnitudes of perhaps the ten brightest members may serve as dependable measures of the apparent characteristics of the clusters themselves.

The mean magnitudes of all members, which in principle offer the most reliable measures, have been directly determined in relatively few cases. Such determinations involve the differentiation of cluster-members from nebulæ in the general field. The problem, although simple in the case of brighter members, is beset with uncertainties in the case of fainter members, since the field-nebulæ of comparable magnitudes are relatively numerous. In general, the mean or most frequent magnitude of a cluster is merely estimated as 2.5 mag. fainter than the brightest member, or 2.1 mag. fainter than the fifth brightest member, or by some similar empirical rule.

The clusters vary through only a moderate range of compactness, and concentration toward the center, while appreciable, is not very conspicuous. In the latter respect, the nebular clusters resemble open, rather than globular, clusters of stars. All types of nebulæ are represented, but in contrast to the general field, the earlier types, and especially the elliptical nebulæ, predominate. In a certain sense each cluster may be characterized by a most fre-

# PLATE IV

## *The Corona Borealis Cluster.*

THE Corona Borealis cluster (R.A. = $15^h$ $19.3^m$, Dec. = +27° 56′, 1930; gal. long. = 10°, lat. = +55°) is a typical example of the great, compact clusters. About four hundred members, most of them elliptical nebulæ, are concentrated in an area of the sky equal to that covered by the full moon. The apparent magnitude of the brightest member is $m =$ 16.5; of the fifth brightest, $m = 16.8$; of the mean of all members, $m =$ 19 (estimated). The faintest members are presumably at the extreme limit of the 100-inch reflector (about $m = 21.5$). Corrections for local obscurations and effects of red-shifts reduce these magnitudes by about 0.25 mag.

Since members of this cluster average about 6.1 mag., fainter than members of the Virgo cluster, they are about 16.5 times as distant. A red-shift corresponding to a radial velocity of 13,100 miles/sec. (21,000 km/sec.) has been measured by Humason in one of the brighter nebulæ of the Corona Borealis cluster. This velocity is 17 times that for the Virgo cluster, in good agreement with the relative luminosities. The adopted distance of the Corona Borealis cluster is derived from the mean absolute magnitude of the fifth brightest nebulæ for clusters in general. Since $M_5 = -16.4$ (chap. vii), the modulus is $m - M = 32.95$, and the distance is 125 million light-years.

The plate was made with the 100-inch reflector, June 20, 1933; north is at the top of the page; 1 mm. = 2″.9.

quent type, although the dispersion around that type is considerable. There are some indications of a correlation between characteristic type and compactness, the density of the cluster diminishing as the most frequent type advances along the sequence of classification. The data are fragmentary, but in connection with the dominance of late types among isolated nebulæ in the general field, they suggest as a possibility that nebulæ may originate in clusters and that the disintegration of clusters may populate the general field. Such speculation, however, is still a subject for conversation rather than a thesis for dissertation, and many more data will be required before it can be seriously considered.

Given the large sample collections of nebulæ present in clusters, the mean, absolute dimensions of the various samples should be fairly comparable. The assumption is strengthened by the apparent characteristics of clusters which, in a general way, resemble those of a single typical cluster as it would appear at selected distances. On this preliminary hypothesis of absolute comparability, which is consistent with all the data available at the present time, the relative distances of clusters were indicated by the mean, apparent luminosities of their members, or, for practical purposes, by the apparent magnitudes of the brighter members. Subsequently, when the absolute distance of the nearest cluster was determined, the absolute distances of all observed clusters were immediately available.

The faintest clusters, since they are the most remote objects to which individual distances can be assigned, are selected when observations at great distances are desired, and the observations are made on the brightest members as a matter of convenience. These brightest nebulæ of clusters represent maximum distances for a specified apparent luminosity.

The results of the surveys may be briefly summarized. The small-scale distribution is irregular, but on a large scale the distribution is approximately uniform. No gradients are found. Everywhere and in all directions, the observable region is much the same.

The nebulæ are not members of the stellar system. The stars form an isolated system, which is embedded in the realm of the nebulæ.

The inhabitants of the realm are scattered singly and in groups. The frequencies of the groups diminish as the sizes of the groups increase. The groups are aggregations drawn from the general field, and are not additional colonies superposed on the field. The largest groups—the great clusters—are curiously similar organizations, and their relative distances are indicated by their apparent dimensions.

CHAPTER IV

# DISTANCES OF NEBULÆ

THE data hitherto discussed concerned the apparent features of nebulæ and their distribution, but gave no indications of absolute distances or dimensions. The investigations were the normal development of a line of research that began long ago with the introduction of photography. Most of the results were within the reach of telescopes of moderate power. The program represented a preliminary phase of the explorations from which there emerged a definite picture of the nebulæ as closely related members of a single family, scattered more or less uniformly through the observable region of space.

The current phase of the explorations is concerned with the interpretation of this pattern. The essential clew was the scale of distance. The mass of nebular data steadily accumulated, but it piled up against the barrier of the unknown quantity. Until distances were available no progress was possible.

The solution of the problem was an achievement of great telescopes. As telescopes and technique improved, they eventually reached a certain critical point and, in due course, the barrier fell. When the breach was open, a wave of exploration swept forward. With distances known, fruitful new methods of research were developed from the knowledge already accumulated. One in particular, derived from red-shifts in the spectra of nebulæ, has led to results that rival in significance the initial solution of the problem of distances.

Investigations of the nebulæ on an absolute scale are based upon two propositions. The first is that distances are indicated by the apparent luminosities of stars involved. The second is that red-shifts are linear functions of distances. These propositions are fundamental, and their development, as well as their applications, will be discussed at some length. The determination of distances came first.

*Development of Criteria of Distances.*

The present phase of nebular research is a very recent development. Three dates are outstanding and any one of them might be selected as an appropriate starting point. The first radial velocity of a nebula was measured in 1912; photographic novæ were discovered in 1917; Cepheids were found in 1924. The second date, 1917, is perhaps the most significant, for the discovery of novæ on photographic plates initiated the study of stars involved in nebulæ. Stars were the clews which led to distances. With distances at last available for a representative collection, the nebulæ were recognized as independent stellar systems, their general characteristics were determined, and the realm of the nebulæ was open to intensive exploration.

The clews have been exploited with extraordinary rapidity. During the course of a single decade—since 1924—a reliable general technique for determining distances has been developed and reconnaissance work has swept out to the very limits of telescopes. The observable region, our sample of the universe, may now be contemplated as a whole.

Precision and finality are beyond the scope of reconnaissance work. Revisions of scale, filling in of details, and especially the recognition of the significance of neglected factors would follow as a matter of course. Nevertheless, the general outlines were sketched with broad strokes. New investigations might be planned and the results interpreted with some preliminary knowledge of their relation to the general scheme.

The situation in 1917 was somewhat as follows. Extragalactic nebulæ (then called spirals, or of the spiral class) were distinguished from the planetaries and the diffuse nebulosities because both of these were recognized as galactic objects. The status of the spirals involved the old controversy about island universes. The wave of speculation generated by the bright nova of 1885 in M31 and, to a less extent, by that of 1895 in NGC 5253, had subsided. The data which were considered of greatest significance were the extraordinary radial velocities, measured by Slipher, and the large angular rotation of M101, measured by van Maanen. The velocities presumably removed the nebulæ from the gravitational control of the galactic system, while the appreciable rotation indicated a moderate distance, presumably leaving M101 well within the system. Thus the evidence was contradictory.

### *Novæ in Spirals.*

IN July, 1917, Ritchey, at the Mount Wilson Observatory, found a previously unrecorded star ($m = 14.6$) on a photograph of the spiral NGC 6946, which, from further data, including a small-scale spectrogram, was identified as a nova.[1] Both Ritchey and Curtis, the latter at the Lick Observatory, immediately examined all duplicate plates of nebulæ in the large collections at their disposal, and found several earlier cases where novæ had appeared in spirals.[2] Two especially interesting objects were identified and followed on a series of plates of M31 which Ritchey had assembled in 1909. The light-curves were definitely of the nova-type, exhibiting the sudden outburst and the slow fading away familiar in the case of galactic novæ, with no subsequent recurrence.

Both objects in M31 had been caught on the rise, and hence their maxima were well determined at about $m = 17$ or 25,000 times fainter than the faintest naked star. Systematic observations during the next two years resulted in the detection of fourteen additional novæ in M31, but none in any other spiral. These novæ were all faint and

[1] Announced in *Harvard College Observatory Bulletin,* No. 641, July 28, 1917.

[2] *Publications of the Astronomical Society of the Pacific,* Vol. 29, 1917. Curtis, *op. cit.,* p. 180; Ritchey, *op. cit.,* p. 210.

September 30, 1925→

# PLATE V

## *Novæ in Messier 31.*

*Left-Hand Plate, September 20, 1925.* Nova No. 54 is the brightest nova that has been observed in M31, with the exception of the great supernova of 1885. On the plate, eight days after maximum ($m = 15.3$), the nova is at $m = 15.7$. It was last seen about one month later.

*Right-Hand Plate, October 4, 1932* (plate by Baade). Nova No. 108 is at $m = 17.0$, about nine days after maximum ($m = 16.0$). A spectrum of this nova, photographed by Humason, is closely similar to spectra of galactic novæ. Nova No. 109 is at $m = 16.7$, about six days after the (unobserved) maximum.

Both plates are with the 100-inch reflector; west is at the top (outer margin of the page); 1 mm. = 7″.0.

October 4, 1932→

represented a homogeneous group, of which Ritchey's two were typical examples. The nova of 1885 in M31 was of a different order. Its luminosity at maximum was a considerable fraction of the total luminosity of the spiral, and in this respect it resembled the photographic novæ detected in other, fainter nebulæ. Two groups of novæ were obviously indicated, one of which was probably several thousand times brighter than the other. The significance of the new data depended upon the answer to the question: Which, if either, of the two groups, dwarfs and giants, was comparable with galactic novæ? The division into groups was suggested by Curtis and others[3] and the problem was formulated in definite terms by Lundmark in 1920.

In 1917, before the distinction was clearly recognized, the available data were thrown together and used indiscriminately. Shapley, as well as Curtis, immediately pointed out that the apparent faintness of novæ in spirals indicated large distances, averaging at least fifty (Shapley) to one hundred (Curtis) times greater than the mean distance of galactic novæ.[4]

Curtis accepted this conclusion as virtually a proof of the island universe hypothesis. Shapley found the evidence inconclusive, and favored the hypothesis that spiral nebulæ were members of the galactic system. Novæ in spirals, he suggested, might be considered as the engulfing of a star by rapidly moving nebulosity. Later (1920),

[3] The early discussions of the implications of novæ in spirals were almost entirely the contributions of three men—Curtis and Shapley in this country, and Lundmark in Sweden. Later, other contributions were made, the most important of which were Luplan-Janssen and Haarh, "Die Parallaxe des Andromeda-Nebels," *Astronomische Nachrichten*, 215, 285, 1922, and Oepik, "An Estimate of the Distance of the Andromeda Nebula," *Astrophysical Journal*, 55, 406, 1922. The former included a comparison of novæ in M31 and the galactic system, using two methods which led to distances of 0.17 and 3.3 million l.y. The latter was a very ingenious use of the (spectrographic) rotation of M31, which, on the assumption of a similar mass-luminosity relation in the spiral and in the galactic system, led to a distance of about 1.5 million l.y. for the spiral.

[4] *Publications of the Astronomical Society of the Pacific*, Vol. 29, 1917. Curtis, *op. cit.*, p. 206; Shapley, *op. cit.*, p. 213.

the two views were more fully developed by their authors for a quasi-debate on "The Scale of the Universe" before the National Academy of Sciences.[5]

In the same year Lundmark[6] published an exhaustive review of available data bearing on the relation of spiral nebulæ to the stellar system, and on the estimation of their distances. His paper, together with the debate, summarized the state of the problem at the time. The novæ clearly furnished a significant criterion of distance, but its application involved the question as to whether the giants or the dwarfs should be identified with the normal novæ in the galactic system. Both Lundmark and Curtis selected the numerous, faint novæ in M31 as the more likely to be comparable with galactic novæ,[7] and estimated the distance of the spiral as of the order of half a million light-years. Curtis concluded that spirals are independent systems comparable with the galactic system, "and indicate to us a greater universe into which we may penetrate to distances of ten million to a hundred million light years." Shapley rejected this conclusion and Lundmark[8]

[5] The two statements, considerably revised, were later published in *Bulletin of the National Research Council,* No. 11, 1921. Preliminary statements are found: Shapley, "On the Existence of External Galaxies," *Publications of the Astronomical Society of the Pacific,* 31, 261, 1919; Curtis, "Modern Theories of the Spiral Nebulæ," *Journal of the Washington Academy of Science,* 9, 217, 1919.

[6] "The Relations of the Globular Clusters and Spiral Nebulæ to the Stellar System," *Kungl. Svenska Vetenskapsakademiens Handlingar,* Band 60, No. 8, 1920.

[7] This identification was later confirmed, and the "dwarfs" and "giants" are now known as "normal novæ" and "supernovæ," respectively.

[8] Lundmark italicized his conclusion (p. 62 of his 1920 paper): "The present investigation has given as the main result that the spiral nebulæ must be considered as situated at considerable distances from the solar system. Whether they are Jeans' star-producing mechanisms or remote galaxies is, on the other hand, more difficult to decide. Possibly we might in the present facts see a suggestion that the latter is the case, but the spiral nebulæ do not, however, seem to be of such dimensions as those that should be ascribed to the galactic system with regard to Shapley's investigations, and much also speaks against regarding the galaxy as having a structure, analogous to that of spiral nebulæ." Although his conclusion was cautiously worded, Lundmark, in his discussions, clearly favored the extragalactic nature of the spirals.

was noncommittal. All three, however, agreed that the new criterion furnished by the novæ placed the spirals at very considerable distances from the solar system.

### *Resolution of Nebulæ.*

THE discovery of novæ led inevitably to the consideration of the more general problem of stars involved in nebulæ and this, in turn, to the definitive solution of the problem of nebular distances. In 1889, Ranyard, editor of *Knowledge,* had reproduced Roberts' photographs of M31, the first to show the spiral structure of the great nebula, and had called attention to numerous stars in the outer regions. The phenomena seemed normal, since current speculation assumed that all the white nebulæ were island universes and would be resolved if only sufficient telescopic power were available.

Roberts himself gave no such clear-cut description of the granulations in the spiral arms. He used the terms "stars," "starlike condensations," "stars surrounded by nebulosity," indiscriminately, both for the nuclei and for the granulations.[9] Suspicions were gradually aroused doubting the stellar character of the condensations, and the suspicions seemed justified when Ritchey, in 1910, described his photographs of large spirals, made with the new 60-inch reflector.[10] These photographs were on a comparatively large scale and were easily the finest and sharpest available. Therefore, when Ritchey stated that "all these [spirals, including M33, 51, 101, etc.] contain great numbers of soft, star-like condensations which I shall call nebulous stars," and referred to 2,400 "nebulous stars" in M33, 1,000 in M101, etc., it was naturally assumed that the condensations did not in general represent individual normal stars.[11] This interpretation of the

[9] Isaac Roberts, *Photographs of Stars, Star Clusters and Nebulæ,* Vol. II, 1900. Attention may be called to pp. 23 and 66. Ranyard's description of M31 is in the February (1889) number of *Knowledge.*

[10] "Mt. Wilson Contr.," No. 47; *Astrophysical Journal,* 32, 26, 1910.

[11] The impression was materially strengthened when Shapley, nine years later, described Mount Wilson photographs in the following terms: "With

photographic images very appreciably delayed investigations of stars involved in nebulæ.

Lundmark, in 1920, stated that an inspection of Ritchey's photograph of M33 revealed "several thousand of the stars that form a part of the vast stellar system." Lundmark, however, had seen only a copy,[12] and, in view of Ritchey's conclusions from the original plates, the new interpretation could scarcely be accepted without further investigation. The new evidence which Lundmark[13] eventually offered as proof was derived from slitless spectra made with a 36-inch reflector and had no direct bearing on the question at issue. The nature of the condensations still remained a matter of speculation.

The solution of the problem emerged several years later from a combination of results from two separate investigations made with a larger telescope, the 100-inch, which was then in operation.[14] One was a study of the photographic images of nebular condensations, using a greater resolving power than had been previously employed; the other was an investigation which led to the recognition of Cepheid variables in nebulæ. A reëxamination of Ritchey's plates of the large nebulæ confirmed the previous conclusion that the images of condensations, although very small, appeared softer than equally faint images on photographs of star-fields. The condensations

one or two possible exceptions, the secondary nuclei in spiral nebulæ are so distinctly nebulous that they cannot be considered individual stars. Even in Messier 33, probably the most conspicuously nucleated of the brighter spirals, it is easy on large-scale plates, to distinguish between the superposed stellar images and the 'softer' nebular condensations." "On the Existence of External Galaxies," *Publications of the Astronomical Society of the Pacific*, 31, 265, 1919. In this paper, Shapley states the arguments which led him to reject "the island universe hypothesis of the spiral nebulæ."

12 Lundmark states that the plate of M33 was placed at his disposal by Professor von Zeipel. This plate was presumably a copy, since the original never left the Mount Wilson Observatory.

13 *Monthly Notices, Royal Astronomical Society*, 85, 890/891, 1925; see also, *Publications of the Astronomical Society of the Pacific*, 33, 324, 1921.

14 The resolution of a spiral, M33, is discussed in "Mt. Wilson Contr.," No. 310; *Astrophysical Journal*, 63, 236, 1926. References to earlier papers are given.

in the nuclear regions, however, were superimposed on relatively dense, unresolved backgrounds; those in the outer regions, where the background was less conspicuous, were distorted by the various aberrations of the telescope. Thus the nonstellar appearance of the images might arise either from the nature of the condensations or from photographic effects under particular conditions.

The latter possibility was investigated in two ways; first, by short exposures centered on the nuclear regions; then by longer exposures centered both on the outer regions of the nebulæ and on neighboring selected areas. In both cases the plates were made with the 100-inch reflector under critical conditions. These plates registered star-images with the smallest angular diameters yet recorded. They fully established the essentially stellar appearance of the photographic images of the great majority of the condensations when effects of photographic smearing were avoided. In M33, for instance, many surface-images, assumed to represent groups of stars, clusters, and occasional patches of nebulosity, were obviously present, but otherwise the images in general were indistinguishable from equally faint star-images on plates centered well away from the nebula.

These results cleared the ground for further investigations. They did not prove that the condensations were stars: they merely proved that the appearance of their photographic images was indistinguishable from that of stellar images. The diameters of the condensations might be any amount less than half a second of arc. But half a second of arc, at great distances, represents large linear diameters. At a million light-years, for instance, an angle of half a second is subtended by about two and a half light-years. A sphere with this diameter might contain many stars, or large masses of nonstellar material.

The definite interpretation of condensations as individual stars was not possible until some of the condensations were identified as Cepheid variables and the range of the light fluctuations was found to be normal. If one star in a condensation representing a group or cluster varied through a certain range, the variation of the condensation as a whole would be much less than that of its single member. The normal range in the Cepheid-like condensations established these condensations as single stars —not even double stars, to say nothing of groups or clusters.

Other types of stars were tentatively identified; the brighter condensations were found to be a predominately early-type (white or blue), implying high luminosities; the faint novæ in M31 were recognized as comparable with galactic novæ; and similar objects were found in M33. The remaining condensations exhibited a frequency-distribution of apparent luminosities resembling, in a general way, that expected among the brighter stars in stellar systems. Hence, when absolute luminosities were established by Cepheids, novæ, and other stars, at certain points in the scale of apparent luminosities, the analogy was complete, and the condensations in general were recognized as individual stars. Further evidence of the consistency of the conclusion is found in the fact that the very brightest stars in the nebulæ whose distances are well determined from Cepheids, are comparable, in their absolute luminosities, with the brightest stars in the galactic system.

## *Cepheids.*

VARIABLE stars in extragalactic nebulæ were first recognized in 1922 when Duncan[15] reported three within the area covered by M33. His data were not sufficient to determine the nature of the variation and he refrained from suggesting any relation between the variables and the nebula. The following year (1923), a dozen variables were found in NGC 6822, an irregular nebula similar to the Magellanic Clouds. Cepheid characteristics were indicated in several of these variables, but were not fully

[15] *Publications of the Astronomical Society of the Pacific*, 34, 290, 1922.

established until the observations had been extended over another year.

The first extragalactic Cepheid was definitely recognized toward the end of 1923 in M31.[16] In the autumn of that year a systematic program of observations had been initiated for the purpose of assembling statistical data on the novæ which were known to appear frequently in the great spiral. The first good plate in the program, made with the 100-inch reflector, led to the discovery of two ordinary novæ and a faint, eighteenth magnitude object which was at first presumed to be another nova. Reference to the long series of plates previously assembled by observers at Mount Wilson in their search for novæ, established the faint object as a variable star and readily indicated the nature of the variation. It was a typical Cepheid with a period of about a month, and hence its absolute luminosity at maximum, as indicated by similar stars in the Magellanic Clouds, was of the order of $M = -4$, or about 7,000 times as bright as the sun. To appear as faint as the observations indicated ($m$ [max] $= 18.2$), the required distance was of the order of 900,000 light-years.

This first definite identification led to an extensive study of the great spiral, using all available material, but based primarily on long exposures with the 100-inch reflector. By the end of 1924, when the first results were published, 36 variables were known, 12 had been recognized as Cepheids, and the order of the distance fully established.[17] In 1929, when the data were published in detail, 40 Cepheids were known and 86 novæ.[18] The light-curves of four Cepheids are shown in Fig. 8.

Meanwhile the investigations had naturally extended

16 The first references to variables in NGC 6822 and in M31, are found in the *Annual Reports of the Mount Wilson Observatory* for the years 1922–1923 and 1923–1924, respectively.

17 The preliminary notice concerning Cepheids in M31 and M33 will be found in *Publications of the American Astronomical Society* (33d meeting), January, 1925. The notice is reprinted in *Observatory*, 48, 139, 1925.

18 "Mt. Wilson Contr.," No. 376; *Astrophysical Journal*, 69, 103, 1929.

August 24, 1925→

# PLATE VI

## *Cepheid Variables in Messier 31.*

*Left-Hand Plate, August 24, 1925* (plate by Duncan); *Right-Hand, November 26, 1924.* The region is centered on an open cluster about 48′ south-preceding the nucleus and close to the major axis of the spiral (see frontispiece). Stars Nos. 43 and 44 are irregular variables; the others are Cepheids. Variation is conspicuous in Nos. 25, 26, and 30, and is appreciable in 37, 39, 43, and 48.

Both plates are with the 100-inch reflector; east is at the top (outer margin of the page); 1 mm. = 5″.0.

The globular cluster marked on the frontispiece is well shown on the left-hand plate (var. No. 30 bright), 15.5 mm. below and 48.5 mm. to the right of upper, left-hand outer corner of the black margin. The large round image of the cluster differs from the comparable images of stars in showing no diffraction rays.

November 26, 1924→

to the large, neighboring spiral, M33. The two brightest of the objects discovered by Duncan were identified as irregular variables, and the faintest, as a Cepheid. By the

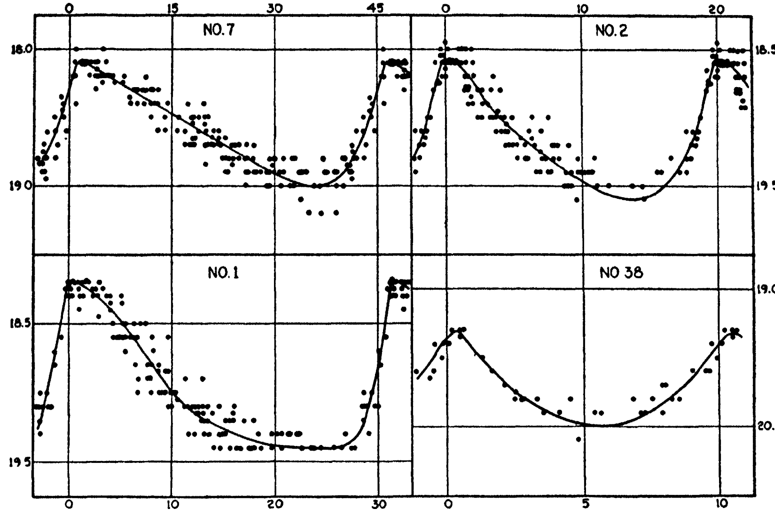



FIG. 8. *Light-Curves of Four Cepheids in Messier 31.*

The vertical scales represent apparent photographic magnitudes; the horizontal, days. The dots represent observations made during many different cycles. The various cycles were superposed and *normal* light-curves drawn through the totality of the data. It will be noticed that the brightest Cepheid, No. 7, has the longest period, and the faintest, No. 38, has the shortest period.

end of 1924, 22 Cepheids were known. By 1926, 35 Cepheids, and 2 novæ were available for discussion.[19] The irregular nebula NGC 6822 had furnished 11 additional Cepheids, and variables of undetermined types had been observed in several other conspicuous nebulæ.[20]

[19] "Mt. Wilson Contr.," No. 310; *Astrophysical Journal,* 63, 236, 1926.

[20] "Mt. Wilson Contr.," No. 304; *Astrophysical Journal,* 62, 409, 1925. After the variables had been found, but before their Cepheid characteristics had been fully established, Shapley published a provisional estimate of the distance of NGC 6822 (order of a million light-years) based on analogies with the Magellanic Clouds. *Harvard College Observatory Bulletin,* No. 796, December, 1923.

The Cepheids appeared to be unequivocal criteria of distances. Light-curves were typical, and periods exhibited the familiar relations with luminosities that had been first established among Cepheids in the Magellanic Clouds. Spectral types were undetermined, of course, but colors, representing integrated spectra, were normal. Distances of spirals could now be derived by precisely the same methods used for investigating the remoter region of the galactic system, namely, the application of criteria of absolute stellar magnitudes. The greatest uncertainty lay in the zero-point of the period-luminosity curve for the Cepheids, a constant necessary for the reduction of apparent magnitudes to absolute magnitudes and hence to distances. The value of the constant was generally accepted as of the correct order, but moderate revisions were expected as new and improved data might be compiled. Meanwhile, the distances of spirals could be rather accurately expressed in terms of the distance of one or both of the Magellanic Clouds as a unit, leaving to future investigations the precise determination of the absolute value of the unit.

### *The Nebulæ as Comparable Galaxies.*

Cepheids were by no means the brightest stars observed in the nebulæ. They were surpassed by novæ, certain irregular variables, and blue giants, and the relative luminosities were all in their normal order as observed in the galactic system. Patches of diffuse nebulosity, giving emission spectra and with blue stars involved (similar to galactic nebulosities) were occasionally found, and, later, objects similar to globular clusters were identified in profusion. The stellar contents presented a consistent analogy with those that would be expected from the Magellanic Clouds or the galactic system if these systems could be studied from very great distances. The evidence of the stars, together with that of radial velocities, was overwhelming, and the theory of island universes seemed to be established beyond reasonable doubt.

The theory had taken two forms. "Island universes" implied merely that the nebulæ were independent stellar systems, scattered through extragalactic space. "Comparable galaxies" carried the additional implication that the dimensions of nebulæ were more or less comparable with those of the galactic system itself. In flat contradiction to both formulations of the theory, there still existed the direct and powerful evidence of large angular rotations.[21] As early as 1916 van Maanen had reported an annual rotation for M101 of the order of 0″.02.[22] Between 1921 and 1923, he published rotations of the same order for six additional spirals and later reported measurements tending to confirm the earlier results.[23]

These large angular rotations implied relatively small distances, a few thousand light-years at most, and hence directly contradicted the evidence of the stars. For instance, in M33, where the linear velocity of rotation was known from spectrograms, the angular rotation indicated a distance of the order of 2,100 light-years as compared with that of 720,000 light-years derived from Cepheids. Lundmark, in 1923, remeasured the pair of plates for M33 and found a rotation in the same direction but numerically so small that it could be considered within the uncertainties of the determination.[24] Otherwise the data on rotations, although completely isolated, were internally consistent but were wholly inconsistent with a theory of island universes.

21 The rotations were derived from the comparison of photographs made several years apart. The relative positions of field-stars and nebular condensations were measured, and systematic differences were found when the measures were compared. These displacements were interpreted as motions in the nebulæ which had occurred during the time-intervals separating the plates—either rotations of the nebulæ or motions of the condensations along the spiral arms.

22 "Mt. Wilson Contr.," No. 118; *Astrophysical Journal,* 44, 210, 1916.

23 A general discussion of the apparent motions in all seven of the spirals is given by van Maanen in the last paper of the series: *Internal Motion of the Spiral Nebula Messier 33, NGC 598,* "Mt. Wilson Contr.," No. 260; *Astrophysical Journal,* 57, 264, 1923.

24 "Mt. Wilson Contr.," No. 308; *Astrophysical Journal,* 63, 67, 1926.

Since the evidence of the stars and of the radial velocities could not be reconciled with that of the angular rotations, it was necessary to reject one of the two sets of data. Because the probability in favor of the first set of data was strong, the rotations were ignored and the field of nebular research was developed in spite of a flat contradiction at its very foundation. The contradiction was removed only in 1935 when investigations of several of the nebulæ by various measurers, using much longer intervals, gave negative results and indicated that the large rotations previously found arose from obscure systematic errors and did not indicate motion, either real or apparent, in the nebulæ themselves.[25]

Another argument, against the theory of "comparable galaxies" rather than that of "island universes," was the very large diameter of the galactic system, 300,000 light-years, which Shapley had derived from his investigations of globular star-clusters.[26] If dimensions of nebulæ were comparable, the distances indicated by apparent diameters would be so great that the novæ would be impossibly bright. The dilemma seemed serious at the time; either the dimensions of nebulæ or the luminosities of the novæ were of a different order from that believed to hold in the galactic system. But novæ offered the more familiar criterion, and the order of distances they suggested was eventually established by the Cepheids, regardless of dimensions. The argument was then restated in the form that, if nebulæ were island universes, the galactic system was a continent.

The discussion eventually settled down to a comparison of the galactic system with M31, which was recognized as an exceptionally large spiral. The large, galactic dimensions had been derived not from the distribution of luminous material in general, which would determine the

[25] Hubble, "Mt. Wilson Contr.," No. 514; *Astrophysical Journal,* 81, 334, 1935. The short paper is followed by a statement by van Maanen.

[26] This value is mentioned by Shapley in the quasi-debate on "The Scale of the Universe," to which a reference has already been made.

surface brightness as seen from a great distance, but from the distribution of several dozen globular clusters. Moreover, the effects of obscuration had been ignored. Many clusters in or near the Milky Way appear faint, not because they are at enormous distances, but because they are obscured by clouds of dust and gas which pervade the low latitudes. When later investigations evaluated these effects, the probable diameter of the system, as outlined by the clusters, was reduced to a half or possibly a third of the original estimate of 300,000 light-years.

The diameter of M31, on the other hand, had been derived from the luminous material in general. Later, when globular clusters were found in M31, the clusters outlined a much larger system, the order of which is comparable with that of the galactic system, although the latter probably represents the less-concentrated type of nebula.[27] Moreover, the original estimates were made from simple inspection of the images on small-scale photographs, and such images can be readily traced with photometers well beyond the limits to which they can be followed by simple inspection. The measured diameters of M31 are now known to be more than double the original estimates, and agree fairly well with the diameters indicated by clusters.[28]

Thus the discrepancy between the dimensions of the spiral and the galactic system has largely vanished. With a better perspective, the continent has shrunk and the island has grown until they can no longer be assigned to different orders. The galactic system may be considered as one of the larger nebulæ. Globular clusters are distributed throughout a great volume of space, but in the outer regions, where an occasional cluster may still be conspicuous, the star-density is probably very low. It is not impossible that the galactic system as viewed from

[27] "Mt. Wilson Contr.," No. 452; *Astrophysical Journal*, 76, 44, 1932.

[28] Stebbins and Whitford, "The Diameter of the Andromeda Nebula," *Proceedings of the National Academy of Sciences*, 20, 93, 1934. See also, later measures by Shapley, *Harvard College Observatory Bulletin*, No. 895, 1934.

M31 covers an area in the sky comparable with that of the spiral as seen from the galactic system.

### *Additional Criteria of Nebular Distances.*

The theory of island universes, even that of comparable galaxies, is now fully established with no outstanding discrepancies. At the time the Cepheids were found, the position was more tentative. Extensive analyses of stars were made in the two conspicuous spirals M31 and M33 and in the irregular nebula NGC 6822. They were clearly independent stellar systems at distances of less than a million light-years. The Magellanic Clouds were then recognized as extragalactic systems at even shorter distances. A small sample collection of nebulæ was thus available as a point of departure for further exploration. The results of the investigation of this group will be presented later, but the general nature of the methods employed may be indicated at this point.

The collection of nebulæ was so small that it could scarcely be considered as a fair sample. But the possibilities offered by the stellar contents were by no means exhausted. Cepheids are not the brightest stars in nebulæ. They are surpassed, as previously mentioned, by normal novæ, by certain irregular variables, and by blue giants such as the O and B stars. Each stellar type furnishes indications of distances, the Cepheids rather accurately, the others only roughly. All are important, since stars alone are the fundamental criteria; other methods of determining nebular distances must, in the end, be calibrated by stars.

With increasing distance we should expect the Cepheids to fade out first, then the irregular variables, then the novæ, then the blue giants, until only the very brightest of all the stars would be seen. Finally, there would remain the millions of nebulæ in which no stars at all, except an occasional supernova, could be detected. The expectations are fulfilled rather precisely by the observations. Moreover, the data, although meager, strongly suggest that the very brightest stars in *late spirals* are of about the same order of absolute luminosity. There appears to be an upper limit of stellar luminosity and this limit, about 50,000 times the luminosity of the sun, is closely approximated in most of the great stellar systems. Thus, when any stars at all can be detected in nebulæ, some rough estimate of distance is possible.

For statistical purposes the method is fairly reliable, and it furnishes a collection of certain types of nebulæ at known distances, large enough to be regarded as a fair sample. The most serious defect in the method is the fact that, in general, stars can be detected only in the later, more open, spirals and in the irregular nebulæ. Fortunately, stars can be detected in some of the spirals which are members of the great Virgo cluster of nebulæ. The other types of nebulæ are well represented among the several hundred members of the cluster, and hence their distances, as well as those of the spirals, are derived from stars. Analysis of the large sample-collection thus available has furnished average characteristics of the nebulæ themselves, which can be used as statistical criteria of distances out as far as the nebulæ can be recorded. Eventually, another criterion was found in the red-shifts, the percentage-accuracy of which increases with the distance.

The exploration of the realm of the nebulæ was carried out with the aid of these criteria. The early work was justified largely by the internal consistency of the results. The foundations were firmly established, but the superstructure represented considerable extrapolations. These were tested in every way that could be devised, but the tests for the most part concerned internal consistency. The ultimate acceptance of the superstructure was due to the steady accumulation of consistent results rather than to critical and definitive experiments.

CHAPTER V

# THE VELOCITY-DISTANCE RELATION

## *Early Spectrograms of Nebulæ.*

SPECTRA of nebulæ were first investigated visually in 1864 by Sir William Huggins (1824–1910).[1] Those of white nebulæ, as extragalactic systems were then called, were apparently continuous, but so faint that no details could be determined with certainty. Prolonged study of the brightest, M31, led to the surmise that both absorption and emission lines or bands were present, and a very faint photograph, achieved in 1888, seemed to confirm the tentative conclusion. No report of the photograph had been published in 1899 when Scheiner settled the question with readable spectrograms of M31.[2] They showed a solar type spectrum with no emission. He concluded that the spiral was probably a stellar system and thus revived the waning interest in the controversy over island universes. Fath and Wolf extended the investigations to other nebulæ with similar results, and eventually the prevalence of solar types among spectra of the brighter spirals, was generally recognized.

## *The First Radial Velocity.*

THE radial velocity of a nebula was measured for the first time in 1912 by V. M. Slipher at the Lowell Observatory.[3] Although the general character of the spectra had been established, the more difficult problem of determining the precise positions of the absorption lines had not been solved. The difficulties arose from the dim surface-bright-

1 *The Scientific Papers of Sir William Huggins* (1909), pp. 101 f.

2 "On the Spectrum of the Great Nebula in Andromeda," *Astrophysical Journal*, 9, 149, 1899.

3 "The Radial Velocity of the Andromeda Nebula," *Lowell Observatory Bulletin*, No. 58, 1914.

ness of the nebular images. Unlike the stars, whose light is concentrated into practically point-images by all telescopes, the nebulæ form relatively large images and the areas increase with the focal length of the telescope employed. Larger telescopes, if the focal ratios are constant, merely spread more light over larger images, leaving the surface-brightness unchanged.

The difficulties are met in direct photography by shortening the focus for a given aperture and thus concentrating the light into smaller images. When the images are photographed through a prism, however, this modification of the telescope offers no advantage. The explanation is simple, but as it involves properties of optical instruments it need not be presented in detail. For large, uniform surfaces, all telescopes are about equally efficient. No advantage can be gained except in the camera behind the prism, which actually photographs the spectra. The rule breaks down in the case of small surfaces, and for the concentrated, semistellar images of the fainter nebulæ the larger telescopes are increasingly efficient. Nevertheless, the most important single factor in the photography of spectra of faint light sources is the speed of the camera.

Slipher exploited this principle and adapted a very fast, short-focus camera to a small-dispersion spectrograph attached to the 24-inch refractor at the Lowell Observatory. With this equipment he was able to record the spectrum of M31 with good definition and on a scale which, although small, was sufficient to show that the absorption lines were not quite in their customary positions. The displacements were toward the violet end of the spectrum, indicating that the radial component of motion was toward the earth. Precise measures revealed that the velocity of approach was about 190 miles (300 kilometers) per second. Four spectrograms secured in the autumn of 1912 gave consistent velocities, and the results could be published with complete confidence in their reliability.

## *Spectra of Nebulæ.*

### A. *Large-Scale Spectra of M32, Compared with the Solar Spectrum.*

THE spectra of the nuclear regions of M31 and M32 are the only absorption spectra of nebulæ that have been obtained on a large scale (original plates, 1 mm. =73 A at 4350 A). They closely resemble the solar spectrum, except that the lines in the nebular spectra are broader, possibly as a result of internal motions in the nebulæ. Dwarf characteristics are conspicuous, and the absolute magnitudes, indicated by the relative intensities of lines, are about the same in all three spectra.

The plate shows the spectrum of M32 below the spectrum of the sun. The comparison spectra are those of an iron arc. Red is to the right; violet, to the left. The last conspicuous absorption line to the right is the $H_\beta$ line of hydrogen. The conspicuous iron lines near the center of the nebular spectra are displaced toward the violet, with respect to the comparison lines, indicating relative motion in the line of sight, toward the observer, amounting to about 120 miles/sec. This motion is largely the reflection of the sun's motion in its orbit around the center of the galactic system. (Plate by Humason.)

### B. *Spectrum of NGC 3115 Showing Evidence of Rotation.*

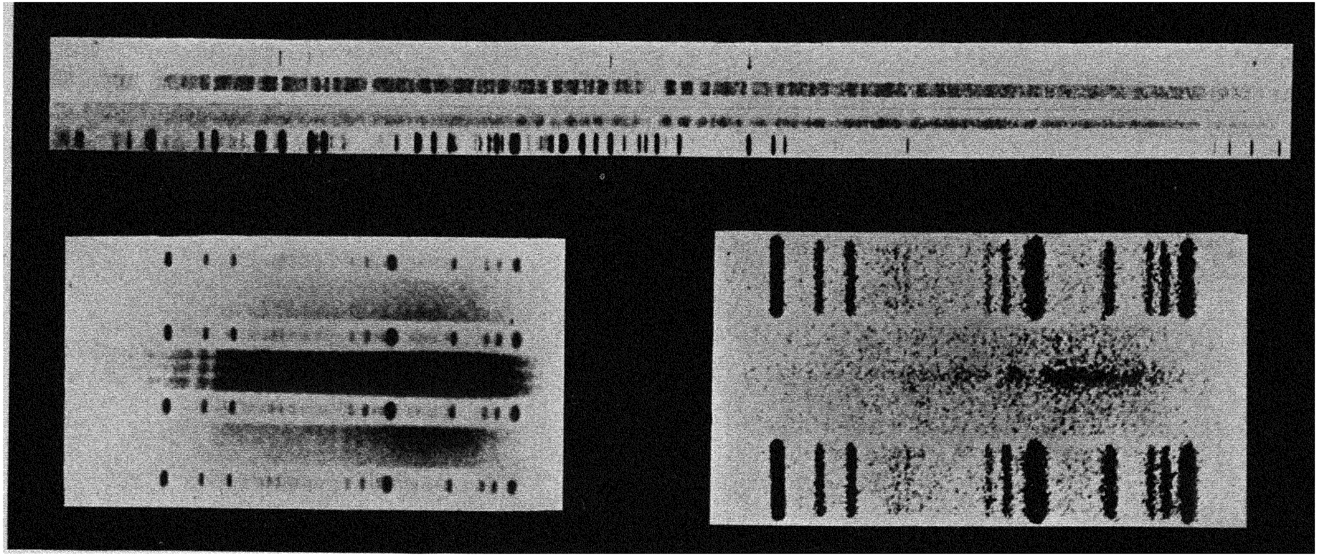
NGC 3115 is the type example of an E7 nebula and is exhibited in Plate I. The slit of the spectrograph has been oriented along the major axis of the spindle-shaped image and, therefore, the upper part of the spectrum represents light from one end of the nebula; the lower part, light from the opposite end; the central strip, light from the nuclear region.

The pair of conspicuous absorption lines near the left (violet) end of the nebular spectrum are the H and K lines of calcium. They are inclined, the top being displaced toward the red, and the bottom, toward the violet, with respect to the central nuclear region. The inclination is interpreted as evidence of rotation of the nebula about the minor axis. One end is receding and the other end is approaching, with respect to the nucleus. The speed of rotation and the manner in which it varies with distance from the nucleus is indicated by the angle of inclination. (Plate by Humason.)

### C. *Spectrum of a Nebula in the Boötes Cluster.*

THE illustration is an unretouched enlargement of a spectrogram (by Humason) on which the scale is, 1 mm. = 875A, at λ4500. It shows the largest well-determined red-shift that has been recorded—an apparent velocity of 24,400 miles/sec. (39,000 km/sec.). Since the spectrogram represents an achievement near the extreme limit of instrumental powers, the significant features are not very conspicuous. However, the H and K lines, blended on the small scale, are readily seen opposite the strongest line in the comparison spectrum (the λ4471 line of helium). The gap in the nebular spectrum, to the left of H and K, is due mainly to a lens-shaped insensitive spot in the coarse-grained emulsion necessarily employed in photographing the very faint spectrum. The normal positions of H and K, in nearby light-sources, are near the second comparison line from the left (see the spectrum of NGC 3115 on this plate, in which H

A



B C

## *Slipher's List of Radial Velocities.*

THE determination of the velocity of M31 has been discussed at some length on the general principle that the first steps in a new field are the most difficult and the most significant. Once the barrier is forced, further development is comparatively simple. But the accumulation of nebular velocities was a slow process and became increasingly laborious after the brightest objects had been observed. Slipher carried on the work almost alone. In 1914 he presented a list of thirteen velocities,[4] and by 1925 the number of his contributions had grown to forty-one.[5] A few of the velocities had been redetermined at other observatories, sufficient to establish the validity of the data beyond any reasonable doubt, but only four new velocities had been added to Slipher's list. In 1925, a total of forty-five nebular velocities was available for discussion.[5]

Although the first velocity was negative, indicating motion toward the observer, positive velocities, indicating motion away from the observer, were found in increasing numbers and soon they completely dominated the list. Moreover, after the most conspicuous nebulæ had been observed, the numerical values of the new velocities were found to be surprisingly large. The complete list ranged from −190 miles/sec. to +1,125 and averaged about +375. The velocities were of an entirely different order from those of any other known type of astronomical body. They were so large that the nebulæ were probably beyond the control of the gravitational field of the stellar system. The nebulæ, it appeared, were independent bodies and this conclusion was consistent with the theory of island universes.

[4] Slipher, "Spectrographic Observations of Nebulæ," Seventeenth Meeting of the American Astronomical Society, August, 1914; reprinted in *Popular Astronomy*, 23, 21, 1915.

[5] The list is published in a paper by Strömberg, *Analysis of Radial Velocities of Globular Clusters and Non-Galactic Nebulæ*, "Mt. Wilson Contr.," No. 292; *Astrophysical Journal*, 61, 353, 1925.

## *Interpretation of the Data.*

### SOLAR MOTION WITH RESPECT TO THE NEBULÆ.

ACTUALLY no other theory was seriously considered in attempting to interpret the data. The stellar system, carrying the sun along with it, was supposed to be moving rapidly through the realm of the nebulæ, which themselves were rushing about with comparable speeds in random directions. Each observed velocity was thus a combination of (*a*) the "peculiar motion" of the nebula, as the individual motion is called, and (*b*) the reflection of the solar motion.[6] If sufficient nebulæ were observed, their random peculiar motions would tend to cancel out, leaving only the reflection of the solar motion to emerge from the totality of the data.

The principle was a familiar one and had worked very well within the stellar system for the determination of the motion of the sun with respect to the stars. It was first applied to the nebulæ by Truman in 1916, when only a dozen nebular velocities were known.[7] Others also solved the equations, including Slipher,[8] when, in 1917, he had twenty-five velocities at his disposal. The numerical results were all rather similar—a solar motion, interpreted as effectively the motion of the stellar system, of the general order of 420 miles/sec., in the general direction of the constellation Capricornus.

It was expected that, when the solar motion was removed, the residual, peculiar motions of the nebulæ would be much smaller than the observed velocities and, furthermore, that they would be distributed at random—that velocities of approach would be as numerous as velocities of recession. Actually, the residual motions were still large and predominantly positive. The unsymmetrical

[6] The solar motion was a combination of the motion of the sun within the stellar system and the motion of the stellar system with respect to the nebulæ.

[7] Truman, "The Motions of the Spiral Nebulæ," *Popular Astronomy*, 24, 111, 1916.

[8] Slipher, "Nebulæ," *Proceedings of the American Philosophical Society*, 56, 403, 1917.

distribution indicated the presence of some systematic effect in addition to the motion of the sun. It was for this reason that Wirtz, in 1918, introduced a seemingly arbitrary K-term—a constant velocity to be subtracted from all observed velocities before the search for the solar motion was undertaken.[9]

The conception of a K-term was not new. It had been used, for instance, in determining the solar motion with respect to the B-stars. In that case it amounted to about four kilometers per second and was supposed to represent some effect of atmospheric pressure, gravitational field, or other condition peculiar to the blue giants. In the case of the nebulæ, however, a term of fantastic dimensions—of the order of a hundred times four kilometers—would be required in order to effect seriously the distribution of residuals. The introduction was a logical step but it requires some boldness to make such a venture.

Wirtz's formulation of the problem included the K-term together with the solar motion as unknowns to be determined from the observational data. He knew of only fifteen velocities at the time of his first solution, but three years later (1921),[10] he repeated the investigation using twenty-nine velocities. The new values were of the same general order as the earlier results. The K-term amounted to about 500 miles/sec. The solar motion was again about 440 miles/sec., but was now in the general direction of the north pole of the heavens. More important, however, the scatter of the residuals or, in other words, of the peculiar velocities of the individual nebulæ, was more or less random. The evidence of systematic effects had almost vanished. The problem was not completely solved—the residu-

9 "Über die Bewegungen der Nebelflecke," *Astronomische Nachrichten*, 206, 109, 1918.

10 Lundmark, in the interim, had made a similar solution, with similar results, using twenty nebular velocities: "The Relations of the Globular Clusters and Spiral Nebulæ to the Stellar System," *Kungl. Svenska Vetenskapsakademiens Handlingar*, Band 60, No. 8, 1920. Wirtz's second paper is "Einiges zur Statistik der Radialbewegungen von Spiralnebeln und Kugelsternhaufen," *Astronomische Nachrichten*, 215, 349, 1921.

als were not wholly satisfactory—but the improvement was so marked that the K-term was accepted as a characteristic feature of nebular velocities. All subsequent discussions of the problem included the K-term as a matter of course.

*The K-Term as a Function of Distance.*

WHEN Wirtz first introduced the K-term, he merely stated that it was necessary because of the preponderance of positive signs and the large numerical values of the velocities. He was fully aware of the consequence that, if the displacements of the spectral lines were literally interpreted as actual velocity-shifts, the K-term must represent a systematic recession of all nebulæ from the vicinity of the galactic system. He did not fully commit himself to this interpretation, but left the question open and used the term as though it were an arbitrary device for "saving the phenomena." The explanation might be found later.

It seemed possible, however, that current theory had already indicated the significance of the K-term. Einstein, in 1915, had formulated his cosmological equation which expressed the relation between the contents of space and the geometry of space, as derived from the theory of general relativity. On the assumption that the universe is static (does not vary systematically with time), he had found a solution to the equation and had, therefore, described a particular kind of universe. de Sitter, in 1916–17, using the same equation, had found another solution. It was later shown that no other solutions[11] were possible on the particular assumptions. The two possible universes were carefully studied in order to see which of them more closely corresponded to the universe we actually inhabit.

[11] "On Einstein's Theory of Gravitation and Its Astronomical Consequences"—Three papers in *Monthly Notices, Royal Astronomical Society,* 76, 699, 1916; 77, 155, 1916; 78, 1, 1917. A third solution, representing a particular case corresponding to special relativity, was possible but it was of no particular interest as an interpretation of the physical universe.

One outstanding difference between the two was the fact that de Sitter's solution predicted positive displacements (red-shifts) in the spectra of distant light sources, which, on the average, should increase with distance from the observer. de Sitter knew of only three velocities[12] at the time and could not make an extensive comparison between theory and observation. Nevertheless, it was clear, as he stated, that the large, positive velocities of the two fainter nebulæ (NGC 1068 and 4594) as contrasted with the negative velocity of M31, the brightest of all the spirals, were consistent with the prediction.

The de Sitter universe is no longer considered as a representation of the actual universe, but at the time it served the important purpose of directing attention to the possibility of a variable K-term. The numerical rate of increase of red-shifts with distance was not predicted by the theory; the rate might be large or small, conspicuous or imperceptible, and the question could be determined only by observations. But among the necessary data were distances of nebulæ, and distances were then unknown. This fact, together with perhaps a natural inertia in the face of revolutionary ideas couched in the unfamiliar language of general relativity, discouraged immediate investigation. It was not until later, when Eddington and others had, as it were, "popularized" the new ideas, that the problem was seriously considered.

If velocities increased with distance, the large constant K-term might represent the velocity corresponding to the mean distance of the particular group of nebulæ that had been observed. This possibility was generally recognized, although no one seems to have made the statement specifically. The problem was formulated as follows: Is the K-term constant for all nebulæ or does it vary with the distance?

12 Slipher's list of thirteen velocities, although published in 1914, had not reached de Sitter, probably as a result of the disruption of communications during the War. For the same reason, Wirtz, in 1918, was probably not aware of de Sitter's papers.

Absolute distances of nebulæ were very uncertain. The only available criteria of relative distances were apparent diameters and apparent luminosities. Neither were reliable because the ranges in intrinsic size and brightness were wholly unknown, and the ranges were believed to be considerable. In the triple system of M31 and its two companions, for instance, the diameters ranged from sixty to one, and the luminosities from a hundred to one. There was no evidence that even these ranges applied to nebulæ at large. Nevertheless, in a general way, the smaller, fainter nebulæ were doubtless at greater mean distances than the larger, brighter objects. The criteria might be useful provided the range in distance covered by the velocities was large compared to the scatter introduced by the criteria.

Wirtz, the leader in the field, made the first attempt, in 1924, to express the K-term as a function of distance using apparent diameters and velocities of forty-two nebulæ.[13] A plausible correlation appeared in the expected direction—velocities tended to increase as diameters diminished. The results, however, were suggestive rather than definitive. Not only were they subject to uncertainties arising from the unknown scatter in real diameters, but they included the effects of an apparent correlation between velocity and concentration. The highly concentrated globular nebulæ, as a class, exhibited the largest mean velocity, and the large, faint irregular nebulæ and open spirals exhibited the smallest mean velocity. In between these limits, the velocities increased with the concentration.

The correlation was generally known, and had inspired unsuccessful attempts to account for the K-term as Einstein-shifts produced by strong, gravitational fields—analogous to the red-shift in the solar spectrum which had served as one of the crucial tests of general relativity. Eventually it was realized that the correlation was

[13] "De Sitter's Kosmologie und die Bewegungen der Spiralnebel," *Astronomische Nachrichten,* 222, 21, 1924.

a simple effect of selection. The concentrated objects, because of their great surface-brightness, were given preference in the laborious task of photographing nebular spectra. So, although these objects are relatively rare, there was a natural tendency to select them for the investigations of faint nebulæ. They represented, on the average, the faintest and most distant of the nebulæ observed and for this reason they exhibited the largest mean velocity. But the explanation came much later. At the time, it was believed that the progression in diameters might signify a progression either in concentration or in distance or in both; therefore the correlation between diameters and velocities was ambiguous.

Furthermore, Wirtz had not used the simple diameters, but their logarithms. The choice was a convenient one, but it led him to express his results as a linear relation between velocities and log diameters or, as he considered, log distances. Such a relation differed in principle from the relation predicted by de Sitter. Hence, in view of the possibility of an alternative explanation as a concentration effect, there was a tendency among astronomers to defer judgment until more information should become available.

Wirtz[14] presented some arguments suggesting that his correlation could not be entirely due to variations either in real diameters or surface-brightness, and Dose,[15] shortly afterwards, showed that a similar, although less-pronounced, correlation existed between velocities and simple diameters. Nevertheless, the later investigations of Lundmark[16] and of Strömberg[17] failed to establish any

14 Wirtz later published a stimulating popular presentation of the investigation and the implications of the results (*Scientia,* 38, 303, 1925) in which he assumed that de Sitter's prediction had been verified.

15 "Zur Statistik der nichtgalaktischen Nebel . . . ," *Astronomische Nachrichten,* 229, 157, 1927.

16 "The Determination of the Curvature of Space-Time in de Sitter's World," *Monthly Notices, Royal Astronomical Society,* 84, 747, 1924.

17 *Analysis of Radial Velocities of Globular Clusters and Non-Galactic Nebulæ,* "Mt. Wilson Contr.," No. 292; *Astrophysical Journal,* 61, 353, 1925.

definite relation between velocity and distance. Lundmark in 1924, using the same nebulæ as Wirtz and employing diameters and luminosities in combination as criteria of distances, concluded somewhat optimistically that "there may be a relation between the two quantities (velocity and distance), although not a very definite one." Strömberg, in 1925, using luminosities alone as criteria of distances, made an especially clear-cut analysis of the data and found "no sufficient reason to believe that there exists any dependence of radial motion upon distance from the sun." This statement, of course, referred to the situation as indicated by the information then at hand. It represented the observer's point of view—that whatever the ultimate truth of the matter might be, the data did not establish a relation. Further discussion would probably contribute little; the important desiderata were additional data and more precise criteria of distances. Strömberg did bring out rather clearly, however, that although the K-term did not seem to vary systematically with distance, it probably varied from nebula to nebula, being small for M31 and the Magellanic Clouds, but large for NGC 584 (for which the largest velocity, +1,125 miles/sec., had been measured).

Shortly afterwards, Lundmark made a final attempt to uncover a variable K-term.[18] He used the same data as before but replaced the constant K-term in the equations by a power series,

$$K = k + lr + mr^2$$

where $r$ was the distance in terms of the undetermined distance of M31 as a unit. The results were disappointing. The constant, $k$, in the series, was found to be 320 miles/sec., somewhat smaller than, but still of the same general order as, the former values of K. The coefficient, $l$, was small and uncertain, about +6 miles/sec., indicating a slight distance-effect (about 8 per cent of the current

18 "The Motions and the Distances of Spiral Nebulæ," *Monthly Notices, Royal Astronomical Society*, 85, 865, 1925.

value), but the coefficient *m*, very small and still more uncertain, was negative, —0.047. Lundmark considered that *m*, although its precise value was uncertain, expressed a real phenomenon which obviously set an upper limit to the velocities of recession that nebulæ could ever attain (aside from their peculiar motions). He concluded that "one would scarcely expect to find any radial velocities larger than +3000 km/sec. among the spirals."

### *The Velocity-Distance Relation.*

HERE the matter rested until 1929. Slipher had turned to other problems and only two or three new velocities had been determined. But new criteria of distances had been developed which were much more reliable than those derived from apparent size and brightness. The new criteria, as described in the preceding chapter, were furnished by stars involved in nebulæ and not by the nebulæ themselves. Nebulæ were now recognized as independent stellar systems scattered through extragalactic space. In a few of the very nearest neighbors, swarms of stars could be photographed and among them various types, well known in the galactic system, could be identified. The apparent faintness of these stars furnished reliable distances of the nebulæ in which they were involved.

Less precise distances were indicated by the apparent faintness of the very brightest stars in the nebulæ. This criterion could be applied as far as the nearest of the great clusters, the Virgo cluster, at a distance of the order of six or seven million light-years. The scatter in the new criterion was reasonably small compared to the range in distance covered by the velocities. The new development led inevitably to a reinvestigation of the K-term as a function of distance.

Although velocities of forty-six objects were available in 1929, the new criteria gave distances for only eighteen of the isolated nebulæ and for the Virgo cluster. Nevertheless, the uncertainties in the distances were so small compared to the range over which they were distributed, that the velocity-distance relation (Fig. 9) emerged from the data in essentially its present form.[19]

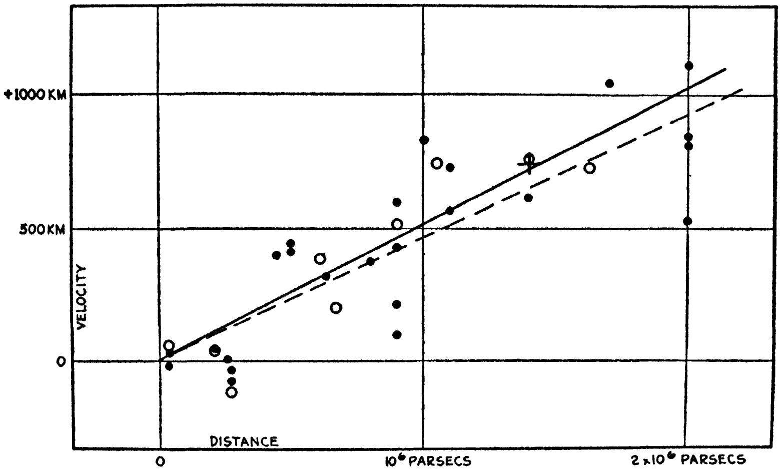



FIG. 9. *The Formulation of the Velocity-Distance Relation.*

The radial velocities (in km/sec.), corrected for solar motion, are plotted against distances (in parsecs) estimated from involved stars and, in the case of the Virgo cluster (represented by the four most distant nebulæ), from the mean luminosity of all nebulæ in the cluster. The black disks and full line represent a solution for the solar motion using the nebulæ individually; the circles and dashed line, a solution combining the nebulæ into groups.

The motion of the sun with respect to the nebulæ was found to be about 175 miles/sec. in the general direction of the bright star Vega. This result is not very different from the solar motion due to galactic rotation—the orbital motion of the sun around the galactic center. The agreement clearly indicates that the motion of the galactic system among the nebulæ must be small. The data are not yet sufficient to determine this motion with any precision.

The K-term was closely represented as a linear func-

[19] Hubble, "A Relation between Distance and Radial Velocity among Extra-Galactic Nebulæ," *Proceedings of the National Academy of Sciences,* 15, 168, 1929.

tion of distance. Velocities, on the average, increased at the rate of roughly 100 miles/sec., per million light-years of distance, over the observed range of about 6.5 million light-years. When the distance effects as well as the solar motion were eliminated from the observed velocities, the residuals, which represented the peculiar motions of the nebulæ, averaged about 100 miles/sec. Moreover, the velocities of approach were about as numerous as the velocities of recession. The observed velocities were thus reduced to order and the distribution of residuals was satisfactory.

The velocity-distance relation, once established, could evidently be used as a criterion of distance for all nebulæ whose velocities were known. The first application of the new criterion was made to the nebulæ in Slipher's list in which no stars could be detected. Observed velocities, divided by the K-term, indicated distances whose only uncertainties were those introduced by peculiar motions. Distances and apparent faintness, taken together, indicated intrinsic luminosities. Intrinsic luminosities derived in this manner were strictly comparable with those of the nebulæ in which stars were observed; both the mean luminosities, and the ranges through which they were scattered, agreed within the uncertainties of the determinations. The nebulæ whose velocities had been measured appeared to form a homogeneous group, and the nebulæ in which stars could be observed, were a fair sample of the group. The consistency of these results was additional evidence of the validity of the velocity-distance relation.

### *Humason's List of Radial Velocities.*

THE data were rediscussed by various authorities with occasional slight modifications, but it was recognized that, in a general way, the linear relation fairly accounted for the observed velocities then available. Yet the data were few and were distributed through a very small portion of the observable region of space. Further progress depended upon the extension of the observations out among fainter, more distant nebulæ. This exacting task was undertaken by Milton Humason at the Mount Wilson Observatory.

Slipher, in his pioneer work, had observed a representative collection of the brighter nebulæ down to about the efficient limits of his 24-inch refractor. Humason, using the great reflectors on Mount Wilson, carried the work far out into unexplored regions. He initiated his program in 1928 and by 1935 had added nearly 150 new velocities, ranging over distances out to thirty-five times the distance of the Virgo cluster.

The new phase in the study of nebular spectra represented a steady advance in technique and instrumental equipment. Where the semistellar images of small, faint nebulæ were concerned, the large reflectors, and especially the 100-inch, had a decided advantage over smaller telescopes. Spectrographs were designed to exploit this peculiar power and were frequently modified as experience suggested.

Development of the most essential equipment, the camera behind the prism, led to the Rayton lens. This lens, designed by Dr. W. B. Rayton of the Bausch and Lomb Optical Company, was constructed on the principle of a reversed microscope objective.[20] The focal ratio of the lens is F0.6—the focal length is a little more than half the aperture—and this great speed enables it to record spectra of exceedingly faint nebulæ.

The success of the Rayton lens led to further experiments which culminated logically in the adaptation of the microscope objective for oil immersion. A focal ratio of F0.35—focal length about one third the aperture—has been reached on this principle, although the lens has not yet been tested at the telescope itself.[21]

20 Rayton, "Two High-Speed Camera Objectives for Astronomical Spectrographs," *Astrophysical Journal*, 72, 59, 1930.

21 This lens was designed by Mr. R. J. Bracey of the British Scientific Instrument Research Association. It is described by Hale, "The Astrophysical Observatory of the California Institute of Technology," *Astrophysical Journal*, 82, 111, 1935.

## *Clusters.*

Humason began his investigations with the spectra of a few bright nebulæ whose velocities were already known. When he was sure that his equipment and technique were reliable he ventured out into the new regions. The first problem was to test the velocity-distance relation over a great range in distance. For this reason, the observations were concentrated on the brightest nebulæ in clusters.

The first velocity, +2,400 miles/sec., for a cluster in Pegasus,[22] was more than double the largest velocity previously known. Thereafter, as fainter and fainter clusters were observed, the absorption lines marched steadily across the spectrum. A velocity of 4,700 miles/sec. was found for the Coma cluster, 9,400 for Ursa Major No. 1, 15,000 for Gemini, and, ultimately, for clusters in Boötes and Ursa Major (No. 2), velocities of 24,500 and 26,000, respectively, which are about one seventh of the velocity of light itself.

The nebulæ in these latter clusters could not be seen at the long (Cassegrain) focus of the telescope. The slit of the spectrograph was set on neighboring stars and was then moved by measured amounts (determined from direct photographs) to the positions of the unseen nebulæ. The observations thus extend nearly to the extreme limit of existing equipment. No very significant advances are expected until larger telescopes are constructed.

The velocities over the entire range increase directly with distances, and the linear relation holds as closely as distances can be estimated.[23] Apparent luminosities of the five or ten brightest nebulæ in clusters would offer fairly reliable criteria of distances except for effects of the

[22] Humason, "The Large Radial Velocity of NGC 7619," *Proceedings of the National Academy of Sciences,* 15, 167, 1929. Humason has since published two considerable lists of velocities, "Mt. Wilson Contr.," Nos. 426 and 531; *Astrophysical Journal,* 74, 35, 1931 and 83, 10, 1936.

[23] Hubble and Humason, *The Velocity-Distance Relation among Extra-Galactic Nebulæ,* "Mt. Wilson Contr.," No. 427; *Astrophysical Journal,* 74, 43, 1931. Subsequent investigations are discussed in chap. vii of this book.

# PLATE VIII

## *The Velocity-Distance Relation.*

THE five examples in the Plate illustrate the empirical law that red-shifts in the spectra of nebulæ increase with the apparent faintness of the nebulæ. Since apparent faintness measures distance, the law can be stated in the form that red-shifts increase with distance. Detailed investigation shows that the relation is linear (red-shifts = constant × distance).

Red-shifts resemble velocity-shifts, and no other satisfactory explanation is available at the present time: red-shifts are due either to actual motion of recession or to some hitherto unrecognized principle of physics. Therefore, the empirical law is generally described as the *velocity-distance relation* (velocity = constant × distance), and is often considered as visible evidence of the expanding universe of general relativity.

The spectra are by Humason. The velocity of NGC 221 is negative—toward the earth—and is a reflection of the sun's motion in its orbit about the center of the galactic system. The other velocities are positive—away from the earth. The distance of NGC 221 (M32) should read 700,000 light years. The revision takes into account the effects of local obscuration.

enormous velocities. Spectra in the fainter clusters are shifted bodily toward the red so far that the distribution of light in the photographic region is sensibly altered. The nebulæ thus appear fainter than normal, and the effect increases directly with the red-shifts. The precise evaluation of the effect is somewhat uncertain and will be discussed later in connection with the interpretation of red-shifts. But the general order of the effect is known and approximate corrections are now applied as a matter of course.

The two faintest clusters in which velocities have been measured are at distances estimated to be of the order of 230 and 240 million light-years (70 and 73 million parsecs), respectively. The linear velocity-distance relation is thus established over an immense volume of space and may be considered as a general characteristic of the observable region itself.

### *Isolated Nebulæ.*

After the preliminary investigations of clusters, emphasis was transferred to isolated nebulæ. Here also, Humason has rapidly enlarged the body of data. The earlier list of eighteen nebulæ in which stars could be detected has grown to thirty-two, and the total number of velocities of isolated nebulæ has passed the hundred mark. Groups also were included in the program, and are now represented by about fifteen velocities in five well-defined groups.

The isolated nebulæ are well represented down to magnitude 13.0, and a considerable number are scattered over the fainter magnitudes. The faintest isolated nebula yet observed is at magnitude 17.5, and has a velocity of +12,000 miles/sec. All the objects appear to conform to the velocity-distance relation determined by the clusters. The relation as derived from the isolated nebulæ is definitely linear. The numerical value of the K-term cannot be determined independently, but when the value for the clusters is introduced, and certain necessary corrections are made for effects of selection, the resulting distances and luminosities of the nebulæ are thoroughly consistent with data derived from other sources. Actually, the velocity-distance relation is so firmly established that it is assumed to hold for all nebulæ, and the observed residuals are analyzed for the information they give concerning the scatter in the intrinsic luminosities or the luminosity-function of nebulæ.[24]

*Significance of the Velocity-Distance Relation.*

As a mere criterion of distance the relation is a valuable aid to nebular research. The only serious errors are those introduced by the peculiar motions. These average between 100 and 150 miles/sec., and are presumably independent of distance. As the K-term increases with distance, the peculiar motion, remaining constant, is an ever-smaller fraction of the K-term. Thus the percentage-accuracy of the determinations increases as the distances increase—a welcome contrast to methods which involve photometric measures.

The velocity-distance relation is not merely a powerful aid to research, it is also a general characteristic of our sample of the universe—one of the very few that are known. Until lately, the explorations of space had been confined to relatively short distances and small volumes—in a cosmic sense, to comparatively microscopic phenomena. Now, in the realm of the nebulæ, large-scale, macroscopic phenomena of matter and radiation could be examined. Expectations ran high. There was a feeling that almost anything might happen and, in fact, the velocity-distance relation did emerge as the mists receded. This was of the first importance for, if it could be fully interpreted, the relation would probably contribute an essential clew to the problem of the structure of the universe.

[24] Hubble and Humason, "The Velocity-Distance Relation for Isolated Extra-Galactic Nebulæ," *Proceedings of the National Academy of Sciences*, 20, 264, 1934. See also, chap. vii of this book.

Observations show that details in nebular spectra are displaced toward the red from their normal positions, and that the red-shifts increase with apparent faintness of the nebulæ. Apparent faintness is confidently interpreted in terms of distance. Therefore, the observational result can be restated—red-shifts increase with distance.

Interpretations of the red-shifts themselves do not inspire such complete confidence. Red-shifts may be expressed as fractions, $d\lambda/\lambda$, where $d\lambda$ is the displacement of a spectral line whose normal wave-length is $\lambda$. The displacements, $d\lambda$, vary systematically through any particular spectrum, but the variation is such that the fraction, $d\lambda/\lambda$, remains constant. Thus $d\lambda/\lambda$ specifies the shift for any nebula, and it is the fraction which increases linearly with distances of the nebulæ.[25] From this point, the term red-shift will be employed for the fraction $d\lambda/\lambda$.

Moreover, the displacements, $d\lambda$, are always positive (toward the red) and so the wave-length of a displaced line, $\lambda + d\lambda$, is always greater than the normal wave-length, $\lambda$. Wave-lengths are increased by the factor $(\lambda + d\lambda)/\lambda$, or the equivalent $1 + d\lambda/\lambda$. Now there is a fundamental relation in physics which states that the energy of any light quantum, multiplied by the wave-length of the quantum, is constant. Thus

$$\text{Energy} \times \text{wave-length} = \text{constant.}$$

Obviously, since the product remains constant, red-shifts, by increasing wave-lengths, must reduce the energy in the quanta. Any plausible interpretation of red-shifts must account for the loss of energy. The loss must occur either in the nebulæ themselves or in the immensely long paths over which the light travels on its journey to the observer.

Thorough investigation of the problem has led to the following conclusions. Several ways are known in which red-shifts might be produced. Of them all, only one will produce large shifts without introducing other effects

25 The apparent radial velocity of a nebula is, to a first approximation, the velocity of light (186,000 miles/sec.) multiplied by the fraction, $d\lambda/\lambda$.

which should be conspicuous, but which are not observed. This explanation interprets red-shifts as Doppler effects, that is to say, as velocity-shifts, indicating actual motion of recession. It may be stated with some confidence that red-shifts are velocity-shifts or else they represent some hitherto unrecognized principle in physics.

The interpretation as velocity-shifts is generally adopted by theoretical investigators, and the velocity-distance relation is considered as the observational basis for theories of an expanding universe. Such theories are widely current. They represent solutions of the cosmological equation, which follow from the assumption of a nonstatic universe. They supersede the earlier solutions made upon the assumption of a static universe, which are now regarded as special cases in the general theory.

Nebular red-shifts, however, are on a very large scale, quite new in our experience, and empirical confirmation of their provisional interpretation as familiar velocity-shifts, is highly desirable. Critical tests are possible, at least in principle, since rapidly receding nebulæ should appear fainter than stationary nebulæ at the same distances. The effects of recession are inconspicuous until the velocities reach appreciable fractions of the velocity of light. This condition is fulfilled, and hence the effects should be measurable, near the limits of the 100-inch reflector.

The problem will be discussed more fully in the concluding chapter. The necessary investigations are beset with difficulties and uncertainties, and conclusions from data now available are rather dubious. They are mentioned here in order to emphasize the fact that the interpretation of red-shifts is at least partially within the range of empirical investigation. For this reason the attitude of the observer is somewhat different from that of the theoretical investigator. Because the telescopic resources are not yet exhausted, judgment may be suspended until it is known from observations whether or not red-shifts do actually represent motion.

Meanwhile, red-shifts may be expressed on a scale of velocities as a matter of convenience. They behave as velocity-shifts behave and they are very simply represented on the same familiar scale, regardless of the ultimate interpretation. The term "apparent velocity" may be used in carefully considered statements, and the adjective always implied where it is omitted in general usage.

CHAPTER VI

# THE LOCAL GROUP

THE previous chapters have described the apparent features of nebulæ and of their distribution, and the development of methods for investigating their intrinsic characteristics. The remaining chapters will present some of the results derived from the application of these methods, first to a group of nebulæ in our immediate neighborhood, then to the remoter nebulæ scattered through the general field, and finally to the realm of the nebulæ as a whole.

Apparent dimensions will no longer be of primary interest except as they lead to absolute dimensions. Linear distances, in general, will be expressed in terms of light-years (l.y.), and luminosities, in terms of absolute magnitudes ($M$). It may be repeated that absolute magnitudes are merely the apparent magnitudes which objects would exhibit, if they were at a certain standard distance—10 parsecs or 32.6 l.y. The sun, at this distance, would be just comfortably visible to the naked eye—the photographic luminosity would be $M = +5.6$. Supergiant stars would rival Venus, and the brightest of them would be readily visible in broad daylight. The faintest nebulæ would be somewhat fainter than the full moon, and the brightest, perhaps a hundred times brighter than the moon.

## *Members of the Local Group.*

SURVEYS of the sky show that nebulæ are scattered singly and in groups of various sizes up to the occasional great clusters. The small-scale distribution resembles that of stars in the stellar system. Analogies among the nebulæ are readily found with individual stars, doubles, triples, multiples, sparse clusters, and open clusters. Globular clusters alone seem to have no counterpart in the realm of the nebulæ.

The galactic system is a member of a typical, small group of nebulæ which is isolated in the general field. The known members of the "local group" are the galactic system with the Magellanic Clouds as its two companions; M31 with M32 and NGC 205 as its companions; M33, NGC 6822 and IC 1613. The three nebulæ, NGC 6946, IC 10 and 342, may be members, but they are so heavily obscured that their distances are indeterminate.

The known members of the local group are fairly accessible. Cepheids have been observed even in the most distant members, and stellar contents have been studied in some detail. These neighboring systems furnished a small sample collection of nebulæ from which criteria were developed for exploring the remoter regions of space.

The nearer nebulæ in the general field are far beyond the limits of the group. Investigations of their stellar contents are so difficult that very little definite information has been gathered. No Cepheids[1] have been identified as yet—distances are estimated from less precise criteria derived from the local group. The fact that the galactic system is a member of a group is a very fortunate accident. The assembling of a sample collection of nebulæ from the general field would probably have been delayed until telescopes were built much larger than those now in operation.

The local group contains two triple nebulæ. The galactic system and M31 are each accompanied by a pair of satellites, so close that their outermost regions probably mingle with those of the primary body. The extragalactic character of the Clouds was not fully recognized until comparatively late. Because of their proximity and their

[1] In M101, one of the nearest of the field-nebulæ, Cepheid characteristics are presumed in the case of a few variables with short intervals between observed maxima, but the light curves and hence the identifications have not been established.

unusual types—they are highly resolved irregular nebulæ —there was a tendency to regard the Clouds as possibly local aggregations in the galactic system. Although they were really the most accessible of the nebulæ, they were passed over, in a sense, and the first definite conquests in extragalactic space were made among more distant systems.

TABLE V.

*Members of the Local Group.*

| | | | | Unit = 1,000 Light-Years | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Galactic | | Dis- | Diam- | | | Veloc- |
| Nebula | Type | λ | β | tance | eter* | M(neb) | M(stars)† | ity‡ |
| LMC | Ir | 247° | −33° | 85 | 12 | −15.9 | −7.1 | 0 |
| SMC | Ir | 269 | −45 | 95 | 6 | 14.5 | 5.8 | +60 |
| M31 | $S_b$ | | | | 40 | 17.5 | 6.0 | −30 |
| M32 | E2 | 89 | −21 | 680 | 0.8 | 12.6 | | |
| NGC 205 | $E5_p$ | | | | 1.6 | 11.5 | | |
| M33 | $S_c$ | 103 | −31 | 720 | 12 | 14.9 | 6.3 | −180 |
| NGC 6822 | Ir | 354 | −20 | 530 | 3.2 | 11.0 | 5.6 | −30 |
| IC 1613 | Ir | 99 | −60 | 900 | 4.4 | −11.2 | 5.8 | |
| Mean | | | | | | −13.6 | −6.1 | |
| *Possible Members.* | | | | | | | | |
| IC 10 | $S_c$? | 87 | −3 | | | | | |
| IC 342 | $S_c$ | 106 | +11 | | | | | +150 |
| NGC 6946 | $S_c$ | 64 | +11 | | | | | +110 |

* Diameters refer to the main bodies of the nebulæ.

† M(stars) refer to the mean of the three or four brightest nonvariable stars in each nebula.

‡ The radial velocities are corrected for the rotation of the sun around the center of the galactic system. The velocities thus represent combinations of the individual (peculiar) motions of the nebulæ with respect to the galactic system, and the distance-effects.

The known and the possible members of the local group are listed in Table V, with certain of their apparent characteristics and their intrinsic dimensions. The known members are distributed through an ellipsoidal volume of space whose longest diameter is about a million light-years. Distances of the possible members are indeterminate, but they could be placed in the same ellipsoid, or in one slightly larger, without unduly straining the observational data.

The galactic system is well out toward one end of the group. The long axis of the ellipsoid stretches away in the general direction of M31. Although the data are much too

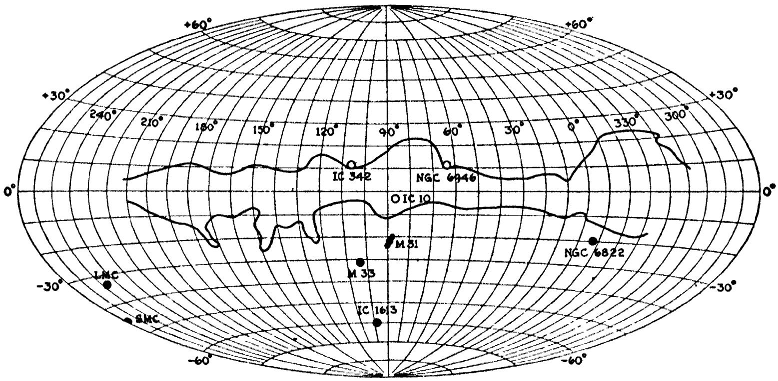



FIG. 10. *Apparent Distribution of Members of the Local Group.*

Positions are plotted in galactic coördinates. The central, horizontal line, 0°—0°, represents the galactic plane, along which lies the irregular zone of avoidance (see Fig. 3.). Disks indicate known members of the local group; circles, possible members. The obscuration of the possible members is conspicuous, especially in the case of IC 10.

scanty for precise statements, the galactic coördinates of the axis may be tentatively adopted as $\lambda = 80°$ to $90°$; $\beta = -25°$. The eccentric position of the galactic system is indicated by the fact that the known and possible members of the group are all within 40° of the axis, with the exception of the Magellanic Clouds and NGC 6822 (Fig. 10). Since the former are satellites of the galactic system itself, and the latter is the nearest of the independent members, their directions are the least significant in the list.

The nebulæ as a group favor the low southern latitudes.

IC 1613, at $\beta = -60°$, represents the limit in one direction and the possible members, IC 342 and NGC 6946, both at $\beta = +11°$, represent the other limit. The distribution is somewhat unfortunate for nebular research, since the prevailing low latitudes lead to conspicuous galactic obscuration. Uncertainties are thus introduced in the study of the very nebulæ which must serve to calibrate methods for further exploration. The obscuration is especially heavy in the case of the three doubtful members; it renders their distances, and hence their membership in the group, indeterminate. Other members may be completely hidden behind the great clouds in the Milky Way, especially in the general direction of the galactic center, where the zone of avoidance is many degrees wide. Moreover, the nebulæ in low latitudes between longitudes 200° and 300° cannot be observed with the great reflectors in the northern hemisphere. Little is known of their stellar contents or their spectra. Several of them, for example NGC 4945 and 5128, must be carefully investigated before they can be definitely excluded from the list of possible members.

The group contains samples of several different types. Barred spirals and early-type normal spirals are missing, but ellipticals, intermediate, and late-type normal spirals, and irregular nebulæ are included. The irregular nebulæ, although relatively rare in the general field, are especially well represented by four members. The Large Magellanic Cloud is a giant nebula and the Small Cloud is about normal. The two remaining irregulars, NGC 6822 and IC 1613, are intrinsically the faintest of all nebulæ whose luminosities have been determined. In view of the wide range in dimensions thus represented, the collection offers an opportunity for the study of the general characteristics of the irregular nebulæ as a class.

The local group has been investigated for two purposes. In the first place, the members have been studied individually, as the nearest and most accessible examples of their particular types, in order to determine the internal structures and stellar contents. The significant data are the forms and structural patterns, the luminosities, dimensions, and masses, and in particular the types and luminosities of the stars involved. In the second place, the group may be examined as a sample collection of nebulæ, from which criteria can be derived for further exploration.

## *The Galactic System.*[2]

THE galactic system is a highly flattened swarm of stars, dust, and gas, in rapid rotation about an axis perpendicular to the galactic plane. The sun is a star nearly in the galactic plane, but far out, possibly 30,000 l.y., from the center of rotation. Details of form and structure are difficult to determine, partly because of the observer's interior position, but more especially as a result of the obscuration produced by the dust. Nevertheless, it is possible, by a combination of facts, analogies, and speculations, to construct a reasonable working hypothesis.

The great, conspicuous dust-clouds[3] are sensibly opaque, obscuring the view over large areas along the Milky Way. The galactic center, presumably the nuclear region of the nebula, is completely hidden. The dust-clouds account for most of the apparent irregularities in the Milky Way by framing normal areas with bordering fields of obscuration. Yet, when full allowance has been made for effects of obscuration, there still remain regions of high local density which are called star-clouds.

[2] For a more extensive discussion, see J. S. Plaskett, "The Dimensions and Structure of the Galaxy" (the Halley Lecture at Oxford, 1935). The lecture is a simple, clear, and authoritative statement of the present conception of the galactic system.

[3] The opaque or semiopaque clouds may consist of material in all forms, but the greater part of the obscuration must be attributed to dust (particles with diameters comparable with the wave-length of light). With material in other forms, impossibly large masses would be required to account for the observed obscuration. Tenuous clouds do exist, which veil rather than obscure, and these might consist of either dust, in smaller quantity, or gas (molecules, atoms, and electrons). Between these limits all degrees of obscuration are found.

The star-clouds, together with other evidence of clustering, produce a certain patchiness in stellar distribution which suggests the observed texture of open spirals or irregular nebulæ such as the Magellanic Clouds. The analogy with spirals is preferable, since the rapid rotation of the galactic system implies a symmetry of form which is lacking in the irregular nebulæ. Finally, the very dim surface-brightness of the galactic system (as it would appear from a great distance), the extent of the resolution, and the prevalence of blue giants and emission-nebulosity, all suggest that the spiral is of the late-type, $S_c$. It is possibly similar to the "massive" spiral, M33.

The star-density presumably diminishes outward from the nucleus to undefined borders. Estimates of dimensions are arbitrary unless a particular density is specified, to which the estimates extend. To an observer far out in space, the surface-brightness of the galactic system might be inconspicuous at forty thousand light-years from the nucleus, although individual giant stars and clusters could be detected to much greater distances. The main body may be pictured as a lenticular mass, possibly seventy to eighty thousand light-years across, and perhaps ten thousand light-years thick at the center. This body consists of a very tenuous medium, strewn with stars, within which the ill-defined spiral arms wind outward from the nuclear region. Star-clouds are found along the vague spiral arms, and obscuring clouds are scattered throughout the fundamental plane.

The rotation of the system is determined from the motion of the sun with respect to the other stars, some of which are nearer the nucleus and others are more remote. At the distance of the sun from the center, the period of rotation is of the order of 220 million years. This period suggests a total mass for the galactic system of the general order of 200,000 million times the mass of the sun.

The brightest stars (photographically) are blue giants (O-type). Other types follow in much the same order as that observed in neighboring systems. Cepheids are conspicuous among the four or five brightest magnitudes. Novæ flare out at the rate of several per year. Patches of emission-nebulosity (Orion nebula, etc.), and globular clusters, are conspicuous features.

## *The Magellanic Clouds.*

THE Magellanic Clouds, because of their proximity, offer exceptional opportunities for detailed study of nebulæ as stellar systems. They are southern objects (Dec. −69° and −73°) and for this reason they have not been analyzed with great reflectors. Radial velocities of numerous patches of emission-nebulosity in the Clouds were determined at the southern station of the Lick Observatory.[4] Otherwise, most of the recent data have been furnished by a 24-inch camera at the southern station of the Harvard College Observatory.[5] With this camera, all objects in the Clouds brighter than about absolute magnitude zero, or perhaps one hundred times brighter than the sun, are readily observed. More detailed information has been obtained from the Clouds than from any other nebulæ in the sky.

Both systems, as previously mentioned, are typical irregular nebulæ—highly resolved, with no nuclei and no conspicuous evidence of rotational symmetry. The stellar contents, down to the observable limits, are remarkably similar to those of the galactic system. The relative, apparent luminosities of corresponding objects furnish many independent estimates of distance. The Cepheids, of course, indicate the most precise distances, but the other criteria are valuable for confirming the order and establishing the general consistency of the results.

The Clouds are readily visible to the naked eye. Their

[4] Wilson, *Publications of the Lick Observatory*, 13, 185, 1918.

[5] Wattenburg, *Astronomische Nachrichten*, 237, 401, 1930, summarizes investigations previous to that data. Some data are found in Shapley, *Star Clusters* (1930). Subsequent investigations are scattered through the *Annals*, *Bulletins*, and *Circulars* of the Harvard College Observatory. In addition, total magnitudes of the Clouds were measured by Van Herk, *Bulletin of the Astronomical Institutes of the Netherlands*, No. 209 (1930).

apparent, photographic magnitudes are about 1.2 and 2.8. The distances derived from Cepheids, according to Shapley, are about 85,000 and 95,000 l.y., the Large Cloud being the nearer. Since they are approximately 23° apart in the sky, their absolute separation is of the order 0.4 times their mean distance from the earth, or perhaps 35,000 l.y.

The Large Cloud, at latitude −33°, is more seriously affected by galactic obscuration than is the Small Cloud (latitude −45°). The differential effect, of the order of 0.1 mag., should reduce the relative distance of the Large Cloud by about 5 per cent. The actual distance of the Small Cloud might also be revised, but uncertainties in the photometric data are probably comparable with the effects of obscuration and the latter may be ignored for the present.

Total absolute magnitudes corresponding to Shapley's distances are −15.9 for the Large Cloud and −14.5 for the Small Cloud—approximately 400 million and 100 million times the luminosity of the sun, respectively.

The main bodies of the clouds are roughly circular, with diameters of about 11,000 and 6,000 l.y. Each contains an elongated, central core whose dimensions are roughly ½ by ¼ the larger diameter. Individual objects, such as variable stars and clusters, can be traced well beyond the main bodies. The latter term, as previously mentioned, is used loosely for the regions readily seen on well-exposed photographs. Diameters have no precise significance unless the density limits are specified.

Radial velocities of +276 and +168 km/sec., for the Large and the Small Cloud, respectively, were derived from patches of emission-nebulosity. The velocities are largely accounted for as reflections of the solar motion, and, therefore, the peculiar motions of the nebulæ in the line of sight are very small.

A single nova has been recorded in each Cloud. The estimated maxima were about $M = -5$ and $-6$, so the objects were similar to the normal novæ found in M31, M33, and the galactic system.

More than 3,000 variable stars have been found in the two Clouds and the list is reported to be incomplete. The majority of the variables are presumably Cepheids,[6] although detailed investigations have been published for only about 200 in the Small Cloud and 40 in the Large Cloud. The former list established the period-luminosity relation in its current form. The list, it may be mentioned, includes the relatively few Cepheids used by Miss Leavitt in her original formulation of the relation.

Long-period variables, irregulars, and eclipsing binaries are occasionally represented, but the analysis of data is incomplete and the relative frequencies are still uncertain. The brightest variable of all is the irregular S Doradus, a P-Cygni type of star involved in the enormous diffuse nebulosity catalogued as NGC 2070. This mass of nebulosity, contained in the Large Cloud, spreads over an area about 200 light-years in diameter. The mean intrinsic luminosity of S Doradus is listed as about $M = -8.3$, which is the brightest luminosity that has been assigned to any individual star. It is about 350,000 times the luminosity of the sun.

The Large Cloud is exceptionally rich in supergiants. Nonvariables range up to perhaps $M = -7.2$ as compared to about $-6.0$ in the Small Cloud. Otherwise the relative frequencies of stars of different luminosities are much the same as those found in the galactic system.

Open clusters and patches of emission-nebulosity are numerous in both Clouds. More than thirty globular clusters are known in the Large Cloud, although only a few have been found in the Small Cloud. The estimated magnitudes of the clusters have been revised from time to

[6] Shapley reports that about 2.5 per cent of all stars in the Small Cloud, brighter than $m = 16.8$ ($M = -0.5$), are variable stars, presumably almost wholly of the Cepheid class; *Proceedings of the National Academy of Sciences*, 22, 10, 1936.

time. The most recent estimates, although admittedly uncertain, suggest that these clusters are closely comparable with similar objects in M31, but are systematically fainter than globular clusters in the galactic system.

### *Messier 31.*[7]

M31, the great spiral in Andromeda, is fairly conspicuous to the naked eye as an elongated cloud about half the size of the full moon, and about as bright as a star of the fourth or fifth magnitude. The spiral structure has never been seen with any telescope, although it is readily photographed with small cameras.

The nebula is a typical $S_b$ spiral with a relatively large, unresolved nuclear region (the portion visible to the naked eye) and fainter arms. The outer parts are well resolved into stars. Novæ, Cepheids, early-type giants, and globular clusters are found in considerable numbers, with relative luminosities much the same as those of similar objects in the Magellanic Clouds. The Cepheids, of which forty have been studied in detail, furnish a fairly reliable distance and the order is confirmed by the other types of objects.

The Cepheids in the spiral appear systematically fainter than Cepheids in the Small Magellanic Cloud (the standard system) by about 4.65 mag. Part of the difference must be attributed to differences in galactic obscuration, for the spiral is only 21° from the galactic plane, while the Cloud is 45°. The appropriate correction reduces the difference in magnitudes arising from relative distance, to about 4.3 mag. Therefore the spiral is 7.25 times as far away as the Cloud, or about 680,000 l.y. (210,000 parsecs).

The spiral has the usual lenticular form with a ratio of axes probably 6 or 7 to 1, but is oriented so that the ratio

[7] For a general discussion of M31, including references to earlier investigations, see Hubble, *A Spiral Nebula as a Stellar System, Messier 31,* "Mt. Wilson Contr.," No. 376; *Astrophysical Journal,* 69, 103, 1929.

in the projected image is about 3 or 4 to 1. The major diameter of the bright nuclear region is of the order of 3,000 l.y. and the spiral arms are readily traced to a diameter of 40,000 l.y. Fainter extensions can be detected by refined methods to perhaps double this diameter, and globular clusters are distributed over the larger area.[8] The dimensions thus appear to be of the same general order as those of the galactic system, although M31 is a more compact type of nebula.

The absolute luminosity is rather uncertainly estimated as $M = -17.5$ (1,700 million times as bright as the sun). The mass, as derived from the rotation of the nuclear region (from inclination of lines in the spectrum), is possibly 30,000 million times the mass of the sun.

One supernova has been observed—S Andromeda (1885). It reached a maximum absolute luminosity of about $M = -14.5$ (100 million times the sun), which is brighter than most nebulæ. Normal novæ flare up at the rate of 25 or 30 per year (115 have been recorded). The light curves, luminosities, and spectra[9] are similar to those of galactic novæ. The absolute magnitudes at maximum are symmetrically distributed around the mean value, $M = -5.5$, with a dispersion of about 0.5 mag. The brightest observed maximum is $M = -6.7$, about 85,000 times as bright as the sun.

The same limit is reached by the brightest variable, an early-type star with an irregular light curve, but with some indications of a five-year period. A half dozen other irregulars are known, one of which is red and may be comparable with the galactic star Betelgeuse. The Cepheids at maximum range from $M = -4.1$ to $-2.7$, depending on their periods. Nonvariable stars brighter than $M = -5$

[8] For the diameter from globular clusters, Hubble, "Mt. Wilson Contr.," No. 452, *Astrophysical Journal,* 76, 44, 1932; from photoelectric measures, Stebbins and Whitford, *Proceedings of the National Academy of Sciences,* 20, 93, 1934; from micro-photometer measures, Shapley, *Harvard College Observatory Bulletin,* No. 895, 1934.

[9] Humason, "The Spectra of Two Novæ in the Andromeda Nebula," *Publications of the Astronomical Society of the Pacific,* 44, 381, 1932.

are not numerous, and the upper limit appears to be near $M = -6$.

An object catalogued as a separate nebula, NGC 206, is really a typical star-cloud in the outer region of M31. The dimensions of the cloud are about 1,300 × 450 l.y. Some 90 stars are brighter than $M = -3.5$, and the numbers increase steadily with increasing faintness until the individuals are lost in the unresolved background. The brighter stars are early types, although their precise classification has not been established. No patches of emission-nebulosity have been identified in the cloud, nor elsewhere in the spiral.

A few open clusters are known. A typical example is found about 48′ south-preceding the nucleus, along the major axis. It is perceptibly elongated, with a major diameter of the order of about 50 l.y. The cluster is partially resolved and a few individual stars can be detected on the border. The spectrum is A-type, and the color-index is considerably less than those of the globular clusters.

About 140 globular clusters are known, and the list is still incomplete for the extreme outer regions of the nebula.[10] Their forms, structures, colors, and spectra are similar to those of globular clusters in the galactic system. The luminosities range from $M = -4$ to $-7$, and the diameters range from about 12 to 50 l.y. The clusters in M31 are thus comparable with those in the Magellanic Clouds, and are systematically smaller and fainter than those in the galactic system.

The distribution of the clusters follows that of the luminosity in the spiral, and is definitely not the spherical distribution found in the galactic system. The clusters, since they can be distinguished with some confidence from the numerous faint nebulæ in the general region of M31, serve to define the maximum extensions of the spiral

[10] Hubble, *Nebulous Objects in Messier 31 Provisionally Identified as Globular Clusters*, "Mt. Wilson Contr.," No. 452; *Astrophysical Journal*, 76, 44, 1932.

beyond the main body. The distribution of clusters suggests a maximum diameter of the order of 100,000 l.y. The search for clusters has also shown the feasibility of determining the opacity of M31 from the manner in which the very faint field-nebulæ fade out as the nucleus is approached. The data, however, are still incomplete and no definite result can be stated.

### *Messier 32.*

M32, the closer and brighter of the two statellites of M31, is the type example of elliptical nebulæ of the class E2 (ratio of axes 8 to 10). As seen in projection, it is superimposed on an arm of the great spiral, about 25′ south of the nucleus. The least possible separation of the two nebulæ (nucleus to nucleus) is thus 5,000 l.y. If M32 lies in the plane of the spiral, the separation is 12,000 l.y. The larger distance may be used for speculations, although M32 may lie anywhere along the line of sight within rather wide limits.

The nebula is highly concentrated, and the luminosity falls rapidly away from the nucleus to undefined borders. Isophotal contours (contours of equal luminosities) are similar ellipses. Diameters, and hence total luminosities, increase with exposure times as far as exposures have been pushed. The largest recorded major diameter is about 8′.5, or 1,700 l.y., and catalogued values range down to about 2′. The great range emphasizes the difficulties of using diameters or luminosities without specifying the condition to which they refer. The absolute luminosity of M32, out to a major diameter of about 4′, or 800 l.y., is of the order of $M = -12.6$ (20 million times the sun). The inclusion of exterior nebulosity does not materially alter this value.

The nucleus, when defined as the least trace of an image that can be recorded on photographs, appears as a semistellar disc about 2″ in diameter. The apparent magnitude is about $m = 13.4$, considerably brighter than the corre-

# PLATE IX

## *Messier 31.*

THE composite picture is made from three plates with the 100-inch reflector, using the Ross zero-power correcting lens at the Newtonian focus. (Plates by Duncan, August 19, 20, 1933.) The nuclear region of the great spiral is unresolved—the stars are too faint to be recorded individually—although novæ flare up in the region at frequent intervals. The outer arms are well resolved and bright giant stars can be studied in detail. The lower right-hand portion (south-preceding the nucleus) is shown on larger scales in the Frontispiece and in Plate V, and there the resolution is conspicuous.

Messier 32, the brighter, closer satellite of the spiral, appears directly below (south of) the nucleus, on the edge of the spiral arms. The brightest star on the plate is below and to the right of Messier 32. NGC 205, the fainter, more distant satellite, is found in the upper right-hand corner of the central plate (north-preceding the nucleus of the spiral).

The picture shows most of the main body of the spiral but fainter extensions can be traced with microphotometers to diameters at least double those of the main body. At the Newtonian focus of the 100-inch reflector, the image of the main body is about two feet long; at the Cassegrain focus, about six feet long.

sponding image in M31. Further analysis, below the limits of photographic methods, can be made visually. Sinclair Smith,[11] who investigated the nucleus of M32 with an interferometer on the 100-inch reflector, found no fringes and concluded that no stellar center was involved. Under critical conditions, he was able to see a stable nuclear image with a diameter of about 0″.8 or 2 l.y., from which the nebulosity falls away in all directions.

The spectral-type of M32 is dG3, the dwarf characteristics being conspicuous. The color-class is g8 on the scale of giant stars. Neither characteristic varies with distance from the nucleus. No polarization has been detected.

The texture of the image is smooth and featureless. There is not the slightest suggestion of resolution and hence stars brighter than, say, $M = -2$, are definitely absent. The conception of M32 as a stellar system presents less serious inconsistencies than other current theories of its structure, but it does not account for the considerable color-excess (difference between color-class and spectral-type). This phenomenon is a characteristic feature of elliptical nebulæ in general.

## *NGC 205.*

NGC 205, the fainter companion of M31, is an abnormal elliptical nebula classed as $E5_p$. Its position, as seen in projection, is about 37′ north-preceding the nucleus of M31, close to the minor axis of the spiral. The minimum separation is thus 7,500 l.y. The precise location in the line of sight is unknown, but for purposes of speculation it may be assumed to lie in the plane of the spiral. In this case the separation is about 30,000 l.y.

The nucleus resembles that of M32, but is considerably fainter and is embedded in relatively fainter nebulosity.

[11] Smith, *Some Notes on the Structure of Elliptical Nebulæ,* "Mt. Wilson Contr.," No. 524; *Astrophysical Journal,* 82, 192, 1935. For investigations of the extranuclear regions, see Hubble, *Distribution of Luminosity in Elliptical Nebulæ,* "Mt. Wilson Contr.," No. 398; *Astrophysical Journal,* 71, 231, 1930.

The luminosity fades away to undefined borders and the isophotal contours are similar ellipses with a ratio of axes of about 5 to 10. The more conspicuous portion of the nebula is about $8' \times 4'$ (1,600 × 800 l.y.). A major diameter of 12′ (2,400 l.y.) has been measured on a photograph with moderate exposure[12] and no doubt still larger dimensions could be derived from longer exposures. The total luminosity is uncertainly estimated as $M = -11.5$ (7 million times the sun), and thus the nebula is a very faint dwarf.

The region immediately surrounding the nucleus exhibits some structure, which is apparently the effect of numerous small, rather sharply defined, patches of obscuration. Very faint stars are more numerous than would be expected for foreground stars alone, and some of them may be associated with the nebula. Several globular clusters are also concentrated within the area, and they may be associated with NGC 205 rather than with the great spiral. These various features, together with an exceptionally low luminosity-gradient, are so unique that the nebula is classified as peculiar. The early spectral-type, F5, is also abnormal.

The radial velocities for the components of the triple system (M31, M32, and NGC 205) are of the same order[13] and are largely accounted for as reflections of the solar motion. The velocity of NGC 205 agrees with that of M31, but the velocity of M32 differs by about 35 km/sec. The latter difference is small, but since it is derived from large-scale spectra, it is probably real. It suggests the possibility of orbital motion of the satellite (M32) around the

[12] Reynolds, "Photometric Measures of the Nebula NGC 205," *Monthly Notices, Royal Astronomical Society*, 94, 519, 1934.

[13] Velocities of M31 and M32, derived from moderately large-scale spectrograms obtained with the same equipment, are −220 and −185 km/sec. The velocity of NGC 205, necessarily derived from very small-scale spectra, is −300 km/sec. and agrees with the velocity of M31 as measured in spectra obtained with the same spectrograph. The difference in the velocities of M31 in the two cases, represents a relatively small systematic variation with scale, the effects of which are known and are applied to nebular velocities as a matter of routine.

primary (M31). NGC 205, if it revolved about M31 in the plane of the spiral, would show no radial component of motion. It lies near the minor axis of the projected image of the spiral, and the orbital motion would be wholly across the line of sight.

The orbital motion of M32, if it were in the plane of the spiral and at a distance of 12,000 l.y. from the nucleus, would be about 105 km/sec. This velocity corresponds to a mass for M31 of the order of $10^{10}$ suns—a not unreasonable value. The satellite (M32), however, would be moving in a direction opposite to that of material in the spiral arms, as indicated by the radial velocities in M31. The discrepancy seems serious and is probably fatal to the assumption that M32 rotates in the plane of the spiral, or is now located in that plane. Still, the dynamical problems of the triple system are in the stage of speculation and much additional information must be assembled before they can be discussed in a definitive manner.

### *Messier 33.*[14]

M33 is an $S_c$ spiral of the massive type, tilted so that the ratio of axes in the projected image is about 2 to 3. The main body has a major diameter of about 1° (12,000 l.y.) and fainter extensions have been followed over nearly twice this diameter. The nucleus resembles in appearance a giant globular cluster, although no evidence of resolution is found. It is semistellar, $M = -8$, spectral type F5, color-excess appreciable, radial velocity, −320 km/sec., as derived from moderately large-scale spectra.

The nuclear region exhibits a background of unresolved nebulosity with a vague, spiral structure and numerous patches of obscuration. Stars are thickly scattered over this background. With increasing distance from the nucleus, the unresolved nebulosity fades and the spiral arms

[14] For a general discussion of M33, including references to earlier investigations, see Hubble, *A Spiral Nebula as a Stellar System, Messier 33,* "Mt. Wilson Contr.," No. 310; *Astrophysical Journal,* 63, 236, 1926.

# PLATE X

## *Messier 33.*

The plate shows the central region, as photographed with the 100-inch reflector. The nebula probably resembles the galactic system but is much smaller.

M33 is a late-type spiral at about the same distance (approximately 700,000 l.y.) as the intermediate type spiral, M31, with its elliptical companions (see Frontispiece and Plate X). A comparison of these nebulae furnish some detailed information on the systematic variation in stellar content along the sequence of classification. For instance, the resolution in M33, unlike that in M31, extends inward to the nucleus. Moreover, an appreciable percentage of the total luminosity of M33 is contributed by blue, supergiant stars; the corresponding percentage in M31 is much smaller, and in the elliptical nebulae, no such stars are found. Data of this kind are accumulating slowly but eventually they may throw some light on the evolution of stars and of stellar systems.

The photograph of M33 was made with the 100-inch reflector on November 30, 1935; south is at the top of the page; 1 mm. $=5''.5$.

become more conspicuous. The arms are broad and highly resolved into stars, clusters, and clouds.

The distance, as indicated by 35 Cepheids, is 720,000 l.y. (220,000 parsecs). The Cepheids appear systematically brighter than those in M31, by 0.1 mag. The difference was at first attributed to relative distance alone and M33 was believed to be slightly nearer than M31. The order is now reversed, for the difference in apparent luminosity is more than balanced by the difference in galactic obscuration—about 0.2 mag., in the opposite direction. However, the obscuration (latitude-effect) is determined in a statistical manner and local variations are possible, especially near the edge of the Milky Way where M31 is situated. This possibility introduces some uncertainty in the relative distances. Perhaps no more precise conclusion is warranted than the remark that both spirals are about equally remote (order of 700,000 l.y.). Equality of their distances is significant in view of their small, angular separation in the sky—about 15°. The linear separation is less than 200,000 l.y., which is a relatively short distance on the cosmic scale.

M33 is somewhat brighter than the average nebula. The total luminosity is about $M = -14.9$ or 160 million times the luminosity of the sun. The mass, as suggested by spectrographic rotation, is probably of the order of 1,000 million suns.

Six apparently normal novæ have been recorded. The brightest variable is irregular and, in 1925, reached a maximum at $M = -6.35$. The spectrum was then of an early type with faint emission lines, presumably the Balmer series of hydrogen. The color-index was negligibly small.

The upper limit for nonvariable stars is about $M = -6.4$. The brighter stars are blue, and colored stars brighter than the Cepheids are very rare. The relative frequency of stars brighter than $M = -3$ is similar to that found in the galactic system. Some small clusters in M33 resemble, in appearance and color, the globular clusters in M31, but are systematically fainter by a magnitude or more. Their real nature, therefore, is somewhat doubtful.

Patches of emission-nebulosity with blue stars involved are numerous, and several are catalogued as separate nebulæ. The most conspicuous, NGC 604, is slightly elongated, and roughly 230 l.y. in diameter. The spectrum closely resembles that of emission-nebulosity in the galactic system, for example, the Orion nebula. A small cluster of stars is involved in NGC 604, the fifteen or twenty brightest of which range from $M = -5$ to $-6.2$. From colors and faint traces of their spectra, the stars are uncertainly identified as O, and B0 types. The relation between the luminosities of the stars and the extent of the nebulosity is consistent with the general relation established within the galactic system.

### *NGC 6822.*[15]

NGC 6822 is an irregular nebula similar to the Magellanic Clouds but much smaller and fainter. It is close to the Milky Way (lat. −20°) and in the general direction of the galactic center (long. 354°), where the great obscuring clouds are most conspicuous. For this reason, the correction for galactic obscuration is rather uncertain.

The apparent magnitudes of 12 Cepheids indicate a distance of 6.7 times the distance of the Small Magellanic Cloud and the normal latitude-correction reduces the absolute distance to about 530,000 l.y. (164,000 parsecs). The main body of the nebula is elongated, with diameters of 3,200 and 1,600 l.y. ($20' \times 10'$). There is a central core about $1{,}250 \times 470$ l.y. ($8' \times 3'$), similar to the cores in the Magellanic Clouds. The possibility of very faint extensions has not been investigated. The nebula is a very faint dwarf, one of the faintest known, with an absolute magni-

[15] For a general discussion, including references to earlier investigations, see Hubble, *NGC 6822, a Remote Stellar System,* "Mt. Wilson Contr.," No. 304; *Astrophysical Journal,* 62, 409, 1925.

tude of the order of $M = -11$ (5 million times the luminosity of the sun).

No novæ have been observed. Several irregular variables are found, none of which is brighter than the brightest Cepheid. A few clusters, presumably globular, are involved, but they are relatively faint and resemble those in M33 rather than those in M31. There are five conspicuous patches of emission-nebulosity, the largest of which resembles a ring, about 130 l.y. in diameter, centered on a small group of bright stars. The upper limit of stellar luminosity is about $M = -5.6$, a low value that may reflect the limited stellar contents of the nebula. A radial velocity of −150 km/sec., derived from a patch of emission-nebulosity, agrees closely with the reflection of the solar motion.

### *IC 1613.*

IC 1613, like NGC 6822, is a small, faint, irregular nebula. The high latitude, −60°, frees it from the grosser effects of local obscuration and thus simplifies the interpretation of the photometric data. Baade, who discovered the nature of the system, has made an exhaustive analysis of the stellar contents with the 100-inch reflector, and the final results will shortly be published.[16] He has found many variable stars, the majority of which are Cepheids. The distance is about 900,000 l.y. The main body is roughly circular, about 4,400 l.y. in diameter, and perhaps five million times as bright as the sun. The stellar content resembles that of NGC 6822, although patches of nebulosity are not so conspicuous.

### *Possible Members of the Local Group.*

THE nine systems previously described are known to be members of the local group. Three other nebulæ, NGC 6946, IC 342, and IC 10 may be considered as possible

[16] Some of the results are mentioned in the *Annual Report of the Mount Wilson Observatory* for the year 1934–35.

# PLATE XI

## *IC 1613.*

THE nebula is irregular and resembles the Magellanic Clouds, although it is much smaller and fainter (about five million times the luminosity of the sun). In comparison with the Large Cloud, IC 1613 is about 75 times fainter and 10.6 times farther away (distance 900,000 l.y.); therefore, it appears about 8500 times fainter than the Large Cloud.

IC 1613 is the most distant of the known members of the local group. Since the galactic latitude is high ($\beta = -60°$), the foreground stars are not numerous. For this reason the differentiation of stars belonging to the nebula is a comparatively simple problem. The stellar contents is similar to those of the Magellanic Clouds. Minor differences are found but they are accounted for by the fact that the Clouds offer much larger sample collections of stars.

The photograph was made (by Baade) with the 100-inch reflector on November 14, 1933; east is at the top of the page; 1 mm. $= 5''.9$.

members. The first two[17] are large, faint $S_c$ spirals found on the northern edge of the Milky Way, both at latitude of +11°. They lie in the zone of partial obscuration which borders the opaque clouds in the Milky Way, and may be heavily obscured in addition to the normal latitude effect. The total obscuration may be any amount from one to three magnitudes, and the photometric data are correspondingly indeterminate as criteria of distance. Stars, none of them variable, are found in both nebulæ, but they merely set upper limits to possible distances. If the obscuration amounts to two or three magnitudes, the nebulæ are members of the local group; if it amounts to one magnitude, they are not members.

Radial velocities furnish additional, independent information, but are also indeterminate. The velocities, corrected for solar motion, are +110 and +150 km/sec., for NGC 6946 and IC 342, respectively. The values represent the sums of the peculiar motions and possible distance-effects. By suitable combinations of these quantities, both of which are unknown, the nebulæ could be placed at will either within or without the local group.

The third nebula, IC 10, is one of the most curious objects in the sky. Mayall, at the Lick Observatory, was the first to call attention to its peculiarities.[18] It lies well within the borders of the Milky Way, at latitude −3°. The longitude, 87°, is about 122° from the galactic center. The texture is apparently that of an extragalactic nebula with some indications of resolution. The photographs are difficult to interpret fully, but they suggest that a portion of a large, late-type spiral is dimly seen between obscuring clouds. The radial velocity is not known. The possibility

[17] IC 342 was recognized as a possible member of the local group in Hubble and Humason, "The Velocity-Distance Relation for Isolated Extra-Galactic Nebulæ," *Proceedings of the National Academy of Sciences*, 20, 264, 1934. The spiral was later described in *Harvard College Observatory Bulletin*, No. 899, 1935. The similar rôle of NGC 6946 has not been commented upon.

[18] "An Extra-Galactic Object 3° from the Plane of the Galaxy," *Publications of the Astronomical Society of the Pacific*, 47, 317, 1935.

NGC 6946→

# PLATE XII

## *NGC 6946 and IC 342.*

THESE large, faint late-type spirals are possible members of the local group. They are seen close to the Milky Way ($\beta = +11°$), and for this reason the foreground stars are numerous (compare Plate XI at $\beta = -60°$ and Plate XIII at $\beta = +75°$). They lie on the edge of the zone of avoidance (see Fig. 10) and are subject to heavy local obscuration. Since the precise amount of the obscuration is unknown, the distances, and hence membership in the local group, are indeterminate.

Both nebulae are partially resolved but, with one exception, no stellar types have been definitely recognized. In 1917, a nova was found in NGC 6946 (see page 85), and the discovery initiated the line of investigations which led to the determination of nebular distances.

The photograph of NGC 6946 (by Humason) was made with the 100-inch reflector on June 19 and 20, 1921; south is at the top (outer margin of page); 1 mm. = 6″.4. The photograph of IC 342 was made with the 60-inch, on November 16, 1933; west is at the top; 1 mm. = 12″.1.

IC 342→

of membership in the local group rests entirely on the excessive obscuration that would be expected at the very low latitude. No more definite statement can be made until further information becomes available.

## *Summary.*

ABSOLUTE magnitudes of group members range from $M = -11$ to $-17.5$, with a mean of $-13.6$. These values may be compared with the corresponding values later found in the general field and in clusters, which range from about $-11.6$ to $-16.8$, with a mean about $-14.2$. The slight discrepancies are largely accounted for by the presence of three very faint dwarfs in the local group—IC 1613, and NGC 6822 and 205. The results suggest that there might be, in the general field, many similar dwarfs so faint that they would be overlooked in general surveys. A careful reëxamination of the surveys demonstrates that such nebulæ would be detected if they existed in considerable numbers, and, therefore, they must be regarded as relatively rare objects. Their presence in the local group appears to be a unique feature of the group, and they detract from its significance as a fair sample of nebulæ in general.

The galactic system was not considered in deriving the means of M(neb) and M(stars). The system is a prominent member of the group and its luminosity, although unknown, is presumably of the same order as that of M31. If this hypothetical value were included in the list, the mean M(neb) would be about $-14.0$ and the small sample collection would agree very closely with the larger sample furnished by field-nebulæ and clusters. The mean M(stars) would probably not be changed by more than 0.1 mag.

One supernova, that of 1885 in M31, has certainly been observed in the group. Any one of several galactic novæ —in particular, the naked-eye nova of 1572—might be placed in the class of supernovæ if their distances were known. The nova of 1917, in the possible member NGC

6946, was presumably a supernova, discovered at some unknown interval after maximum.

Normal novæ have been observed in four members, in addition to the galactic system—one in each of the Magellanic Clouds, six in M33 and 115 in M31. The giant $S_b$ spiral is clearly a favored system, and the $S_c$ spiral is more favored than the irregular nebulæ. The galactic system (probably $S_c$) lies between M31 and M33. The scanty data suggest that the frequency of normal novæ may depend upon the nebular type (degree of concentration of material within the system), and, among nebulæ of a given type, upon total luminosity (stellar contents).

Stars are not found in the typical elliptical nebula, M32, and their presence in the peculiar system, NGC 205, is questionable. All the other members of the group are well resolved. The brightest stars in the Large Magellanic Cloud and in M31 are variables (irregular variables), but are nonvariable in the other resolved members. The Large Cloud appears to be abnormally rich in supergiant stars, even when variables are excluded. The upper limit of stellar luminosity, of the order of $M = -7.2$, appears to be an outstanding exception. The corresponding limits in the other members range from about $-5.5$ to $-6.5$ with a mean around $-6.0$. A review of these data, together with a few additional data from the nearer nebulæ in the general field, has led to the adoption of $M = -6.1$, for the brightest stars in resolved nebulæ. This quantity will be used, in the next chapter, as a criterion of distance.

Patches of emission-nebulosity are found in the late-type spirals and irregular nebulæ, but not in the earlier types. This distribution seems to hold generally. Blue giants and emission-nebulosity are characteristic features of high resolution.

The radial velocities of group members are largely accounted for as simple reflections of the solar motion. The residuals represent combinations of peculiar motions and distance-effects, and the distance-effects are necessarily positive. Since the residuals are small—the mean is a small negative quantity—the distance-effects must be lacking, or the peculiar motions must be large and systematically negative. The latter alternative, which implies that the group is contracting or that the galactic system is approaching the center, does not seem to be supported by the observed velocities of the nearer field-nebulæ. The data are few and their interpretation is not entirely clear, but they suggest, although they do not establish, the hypothesis that the velocity-distance relation does not operate within the local group.

CHAPTER VII

# THE GENERAL FIELD

## *Criteria of Distances.*

THE local group is a sample collection of nebulæ whose distances have been derived by the familiar methods used for investigations within the galactic system. The sample is small but it is a very welcome, and almost a necessary, connecting link with nebulæ in the general field. The brightest class of objects in members of the local group can still be seen in the nearer field-nebulæ that have been resolved. These brightest objects furnish a common denominator by means of which the two sets of nebulæ can be compared. The comparison shows that resolved field-nebulæ are similar, in a general way, to resolved nebulæ in the local group. This conclusion is of fundamental importance.

The common denominator is the "brightest star," defined as the mean of the three or four brightest individual stars in a nebula. These "brightest stars," although not precisely equal, are of the same order of intrinsic luminosity in all of the resolved nebulæ. The mean luminosity was determined largely from the members of the local group, where distances were known from Cepheids. Once the calibration was established, the brightest stars indicated the distances of all resolved nebulæ in the general field. The distances, however, were statistical—subject to errors in individual cases, although reliable in the mean.

The brightest stars might be used to explore the inner fringe of the general field, in much the same way that Cepheids were used to explore the local group. A few of the more conspicuous field-nebulæ might be described in some detail, just as the neighboring systems were treated in the previous chapter. But, for purposes of general exploration, the criterion is used in another way. The resolved field-nebulæ are sufficiently numerous—about 125 arė known—to constitute a fair sample of later-type nebulæ in general. This sample collection of resolved nebulæ, as a group, serves to calibrate a new criterion of distance, namely, the intrinsic luminosity of the nebulæ themselves, which applies to all nebulæ, regardless of type or apparent faintness.

The field of the new criterion is the realm of the nebulæ—the whole of the observable region of space. The criterion is again statistical, but now the dispersion is large. It is essential that the calibration should include not only the mean luminosity of nebulæ but also the scatter among the individual luminosities. The requisite data are contained in the luminosity-function, which, as previously mentioned, is the technical term for the frequency-distribution of absolute magnitudes among the nebulæ in a given volume of space. The characteristics of the luminosity-function are (*a*) the form of the frequency-curve, (*b*) the mean or the most frequent magnitude, and (*c*) the dispersion. When these are known the function is completely described.

The form of the frequency-curve, together with the dispersion, has been determined from apparent magnitudes in various sets of data. The results are fairly consistent: the curve, to anticipate, approximates a normal error-curve, and the dispersion is somewhat less than a magnitude. The value of the mean absolute magnitude, $M_0$, as it is termed, can be derived only from nebulæ whose distances are known. Members of the local group are too few for the purpose and so the determination of $M_0$ depends primarily on the sample collection of resolved nebulæ in the general field.

The gradual development of distance criteria was inevitable but is none the less impressive. The field of exploration was enlarged step by step. Each criterion furnished a sample collection of objects which calibrated a new criterion, less precise but ranging to greater distances. The first step was within the solar system—the distance of the sun, which is the astronomical unit. Then followed, within the stellar system, parallaxes of stars, stellar motions, spectrographic parallaxes, and Cepheids. This last criterion furnished the link with extragalactic regions. It led to brightest stars in nebulæ, luminosities of nebulæ, and finally, as will be seen, brightest nebulæ in clusters.

Light from the sun reaches the earth in about eight minutes (more precisely, 500 seconds); from the most distant cluster that has been measured, the journey requires 240 million years. The multiplication factor is about $1.5 \times 10^{13}$, yet the distance of the cluster is known with an uncertainty that is probably not more than 15 per cent. The comparison emphasizes the high order of consistency encountered in the realm of the nebulæ. Such precision can be expected only when statistical methods are applied to large bodies of closely comparable data.

*Brightest Stars.*

The study of the general field begins with the comparison between members of the local group and the nearer field-nebulæ. Cepheids are by no means the brightest objects that can be recognized in the members of the local group. They are surpassed by certain types of irregular variables, normal novæ, brightest stars, globular clusters, open clusters, and patches of emission-nebulosity. With increasing distance, these objects should fade out class by class, and the actual observations do, in fact, present a series of nebulæ in just the expected sequence. Eventually, when the last detail of stellar contents is lost to the telescope, there remain only the total luminosities of the nebulæ and the red-shifts in their spectra, as possible criteria of distances.

The only exceptions to this statement are the exceedingly rare supernovæ which appear in any one nebula at average intervals of the order of 500 to 1,000 years. From the scanty data available, they are believed to reach, at maxima, a fairly uniform intrinsic luminosity which is comparable with the average luminosity of the nebulæ themselves. Supernovæ can be detected at immense distances and, in principle, they are a criterion of distance about as reliable as that of total luminosities of the nebulæ. Actually, however, the maxima are so seldom observed and the novæ themselves are so rare that they contribute very little to the present problem.

A few irregular variables and normal novæ have been identified in three or four of the nearest field-nebulæ, but not in sufficient numbers to furnish very precise distances. Open clusters and patches of emission-nebulosity are more frequent, but they are difficult to identify with certainty, and their group characteristics are not well enough known to inspire confidence in their validity as criteria. The problem at present seems to be the determination of their group characteristics on the basis of distances derived from other independent sources.

Globular clusters also present anomalies to be investigated. Stars can be readily distinguished from typical clusters down to apparent magnitude 19, and, under the best conditions, to perhaps 19.5. A survey of nebulæ in which stars are found brighter than these limiting magnitudes indicates that globular clusters, if they are generally present, must vary widely from system to system. The brightest cluster in a system seldom surpasses the brightest star, and even then the difference is small.[1] Near the limits of photographic plates, occasional clusters

[1] Within a given nebula, the globular clusters appear to range through about four or five magnitudes. The brightest clusters are slightly brighter than the brightest stars in M31 and M101, but are considerably fainter than the brightest star in M33, NGC 6822 and other nebulæ, probably including the Magellanic Clouds. The relation in the galactic system is rather uncertain, although available data suggest that the clusters are exceptionally luminous as a group. The brighter ones, if current estimates of their intrinsic luminosities are reliable, would be conspicuous objects in members of the local group and would be readily identified in the nearer field-nebulæ and even in members of the Virgo cluster. The observed absence of such objects in the nearer nebulæ suggests that globular clusters in the galactic system are unique or that the estimates of their intrinsic luminosities are excessive.

may be mistaken for brightest stars, but the effect of the confusion is not serious in deriving average results from considerable bodies of data.

The brightest stars thus appear to be the last useful criteria in the stellar contents of nebulæ to fade out with increasing distance. These stars are believed to be about equally luminous in all the resolved nebulæ. The assumption is supported by two arguments, the one theoretical, the other, empirical. There are theoretical reasons, as Eddington has shown, for supposing the existence of a fairly definite upper limit to the luminosity which normal stars ever attain.[2] If such a limit exists, it would probably be approached by some objects in any normal sample collection containing many millions of stars. The nebulæ, or at least the later-type spirals, represent normal sample collections of the requisite order, and therefore each of them should include some stars which are close to the limit.[3]

But aside from theory, it is found, as an empirical fact, that the brightest stars are fairly comparable in resolved nebulæ whose distances are known from other criteria. Such nebulæ are not numerous. The list consists of the galactic system, six members of the local group and three of the nearest nebulæ in the general field.[4] The term

[2] The argument is that luminosity is a direct function of mass (mass-luminosity relation) and that the mass cannot exceed a certain limit (of the order of a hundred suns) because the radiation-pressure would then become so great that the star would be unstable. See Eddington, *The Internal Constitution of the Stars* (1926), chaps. i, vi, and vii; "Stars and Atoms" (First Lecture), 1927.

[3] The absence of very bright stars in ellipticals and early spirals suggests that these nebulæ represent abnormal sample collections of stars and may have evolutional significance.

[4] The field-nebulæ are M101, M81, and NGC 2403, in which distances are indicated by novæ, irregular variables, and globular clusters, in addition to brightest stars. Certain variables which may be Cepheids are found in M101, but the identification is not yet established. The distances are not so reliable as those for members of the local group, but the addition to the list of separate nebulæ is believed to more than balance the uncertainties. See Hubble and Humason, *The Velocity-Distance Relation among Extra-Galactic Nebulæ*, "Mt. Wilson Contr.," No. 427; *Astrophysical Journal*, 74, 43, 1931.

"brightest star" has been arbitrarily defined as the mean of the three or four brightest stars in a nebula. These stars in general are about equally luminous, so the use of several merely reduces the effects of exceptional cases or misidentifications without seriously restricting the application of the criterion. In this sense, the brightest stars in the ten systems range through about two magnitudes. The mean for all ten systems, together with the dispersion, σ, which measures the scatter, is

$$M_s = -6.1$$
$$\sigma = 0.41.$$

The brightest stars thus average about 48,000 times brighter than the sun.

These numerical values represent the yardstick by which distances of resolved nebulæ are measured. The calibration is made from very scanty data and will be revised when reliable distances of more field-nebulæ are available. A few may be determined, in time, with the 100-inch reflector, but significant additions are not expected until the 200-inch reflector, now under construction, can be applied to the problem.

It has already been mentioned that the brightest stars in different nebulæ are not all equally luminous. Exceptionally bright stars cannot be distinguished from the exceptionally faint, unless distances are already known from other sources. In large groups of nebulæ, the exceptional cases tend to cancel out, but, for an individual nebula, there is necessarily an uncertainty since the brightest star may be normal or abnormal. When the numerical value of the dispersion is known, the uncertainties can be calculated for individual nebulæ or for groups of any size.

Dispersion in such criteria of distance introduces effects of selection in statistical surveys. The effect is small when brightest stars furnish the criterion, but later, when total luminosities of the nebulæ themselves are employed, it becomes important and will be discussed in some detail.

Meanwhile it will be stated, without further explanation, that, when nebulæ are selected on the basis of the apparent magnitudes of their brightest stars, the mean absolute magnitude of these brightest stars is not that derived in the calibration, namely, $M_s = -6.1$, but is brighter by the amount, $1.382\ \sigma^2$, where $\sigma$ is the dispersion. Therefore, in the survey of resolved nebulæ, the value $\overline{M}_s = -6.35$ must be used as the yardstick for statistical distances. The corresponding luminosity is about 60,000 times that of the sun.

### *Uncertainties in the Criterion of Brightest Stars.*

Because of the great significance of the criterion of brightest stars, its uncertainties and limitations will be reviewed before it is actually applied to the general field. Stars are observed only in certain types of nebulæ, namely, the intermediate- and late-type spirals and the irregulars. Among these nebulæ, the upper limit of stellar luminosity seems to vary systematically with the type. The variation has not been precisely determined but it is known to be small. However, since the great majority of nebulæ in which stars can be detected are of a single type, $S_c$, the variation will affect the dispersion of residuals in statistical investigations rather than the mean results.

The objection can be raised that images identified as individual stars might represent groups or clusters. This criticism seems well founded because, in a distant nebula, the contents of a large volume of space would be indistinguishable from a single star. However, a study of clusters and groups in the galactic system, as they would appear from great distances, indicates that the uncertainties in distinguishing the brightest individual stars would not be serious. The conclusion is confirmed by a similar study of clusters in members of the local group. These neighboring nebulæ are especially significant. The Cepheids are obviously single stars. Objects selected as brightest stars are brighter than Cepheids, and the ratio is about the same as that found between corresponding objects in the galactic system. For this reason, it is highly probable that the objects selected as brightest stars in members of the local group are actually individual stars. Finally, as will appear later, the relative luminosity of nebulæ and their brightest stars is much the same in the local group as in the field-nebulæ.[5] Therefore, the brightest stars in the field-nebulæ also are probably individual stars.

Moreover, regardless of their real nature, the objects selected as brightest stars appear to represent strictly comparable bodies. They are selected in the examination of homogeneous material—comparable photographs—and their luminosities, relative to the nebulæ in which they lie, do not vary systematically with apparent faintness. This evidence of homogeneity is important since the observations, in general, are carried out near the limits of the telescopes, where systematic errors are difficult to avoid. Other uncertainties are of minor importance. The elimination of field-stars is a simple statistical problem which may introduce slight accidental, but no appreciable systematic, errors. The occasional confusion of stars and globular clusters has already been mentioned.

The criterion appears to function very well for purposes of preliminary reconnaissance and independent evidence of consistency is furnished by the velocity-distance relation.

## *Application of the Criterion of Brightest Stars.*

### (*a*) THE LUMINOSITY-FUNCTION FOR RESOLVED NEBULÆ.

THE criterion of brightest stars has been applied to three major problems, all of which contribute to the formulation of the general luminosity-function. The problems are (*a*) the luminosity-function of resolved field-nebulæ, (*b*)

[5] The mean difference $m_s - m_n$, uncorrected for nebular type, is 7.93 for the six resolved nebulæ which are definitely known to be members of the local group. Reduced to the type, $S_c$, the mean is 8.28. If the doubtful cases IC 342 and NGC 6946 are included, the mean is 7.96. These values are of the same order as that for the field-nebulæ, namely, 7.85.

the distance of the Virgo cluster, and (*c*) the numerical calibration of the velocity-distance relation.

The luminosity-function, as previously mentioned, is technically the frequency-distribution of absolute magnitudes among nebulæ in a given volume of space. It is impossible to assemble just these data, but the same frequency-distribution can be derived from any large collection of nebulæ which are selected at random and whose individual distances are known. It is essential, however, that the dispersion in the criterion of distance should be reasonably small. The resolved field-nebulæ furnish the only considerable list of data which approximate the specifications.

On photographs made with the large reflectors on Mount Wilson, stars can be identified with some confidence in about 125 nebulæ.[6] Apparent magnitudes of the brightest stars, $m_s$, and of the nebulæ themselves,[7] $m_n$, have been measured or estimated, and the difference $m_s - m_n$, has been calculated for each nebula. The differences measure the luminosities of the nebulæ in terms of the luminosities of their brightest stars. A difference of five magnitudes, for instance, means that a nebula is 100 times brighter than its brightest star; a difference of ten magnitudes, 10,000 times brighter. If the stars were all precisely equal, the differences would directly indicate the absolute magnitudes of the nebulæ. The frequency-distribution of the differences would then completely determine the luminosity-function. Actually, of course, the dispersion among the brightest stars complicates the problem, although the effects are small and can be calculated.

6 Preliminary lists of such data are given by Hubble (''Mt. Wilson Contr.,'' No. 324; *Astrophysical Journal*, 64, 321, 1926) and by Hubble and Humason (''Mt. Wilson Contr.,'' No. 427; *Astrophysical Journal*, 74, 43, 1931). The more extensive list discussed here includes additional data assembled by the writer, which will be published in a forthcoming ''Mt. Wilson Contr.,'' No. 548.

7 The nebular magnitudes, in general, represent actual measures by Stebbins and his associates, using a photoelectric cell directly at the focus of the telescope, or magnitudes listed in the Harvard survey of bright nebulæ, reduced to the system of the photoelectric measures.

The frequency-distribution of the differences, $m_s - m_n$, is shown in Fig. 11. The points are well represented by

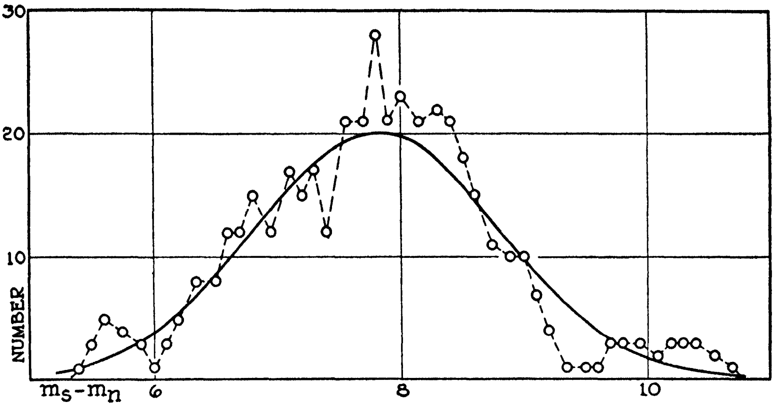



FIG. 11. *Frequency-Distribution of* $m_s - m_n$.

$m_n$ and $m_s$ are the apparent magnitudes of nebulæ and of their brightest stars. Since the intrinsic luminosity (candle power) of the brightest stars is approximately the same in all nebulæ, the differences, $m_s - m_n$, indicate the intrinsic luminosities of the nebulæ in terms of the brightest star as the unit. Therefore, when the absolute magnitude of the unit, $M_s = -6.35$, is subtracted from the values, $m_s - m_n$, the diagram represents the relative frequency of various absolute magnitudes of the nebulæ.

About 125 nebulæ are included in the diagram and the observed distribution (dash line) has been smoothed by the use of overlapping sums of three successive values. The full line is a normal error-curve with a dispersion of $\sigma = 0.9$.

the normal error-curve which has been adjusted to fit the data. The mean or most frequent difference is 7.84 mag., and the dispersion is 0.94 mag.[8] The dispersion includes that of the brightest stars as well as that of the nebulæ. When the former is removed, the dispersion for the nebulæ alone is found to be about $\sigma = 0.84$.

[8] The simple differences range from 5.2 to 11.3, with a mean about 7.9 mag. The frequency-distribution is symmetrical, approximating a normal-curve, with a dispersion slightly less than 1 mag.

Of the 125 nebulæ, 11 are irregulars, 16 are intermediate spirals, and the

The luminosity-function for resolved nebulæ is thus a normal error-curve with

$$M_0 = -6.35 - 7.84 = -14.2$$
$$\sigma = 0.84.$$

The mean luminosity of the nebulæ is about 85 million times as bright as the sun. The dispersion indicates that about half of the nebulæ are within the narrow range from one half to twice the mean luminosity. The generalization of the function to include all types follows from a comparative study of the different types as they appear in clusters, and in the velocity-distance relation.

### (*b*) DISTANCE OF THE VIRGO CLUSTER.

ALL types of nebulæ are found in clusters, although ellipticals and early spirals predominate. A study of apparent magnitudes within single clusters shows that the mean luminosities, and the dispersions, of the various types are comparable throughout the sequence of classification. If systematic variations occur, they are too small to be established by the data at hand. Therefore, the luminosity-function for resolved nebulæ may be provisionally applied to all types, provided the cluster-nebulæ, as a group, are comparable with field-nebulæ.

The question of comparability can be answered only

remainder, 98 in all, are late spirals. The mean differences for the separate types are as follows:

| Type | No. | $m_s - m_n$ |
|---|---|---|
| $S_b$ | 16 | $8.57 \pm 0.20$ |
| $S_c$ | 98 | $7.84 \pm 0.06$ |
| Ir | 11 | $7.15 \pm 0.25$ |

These results indicate a systematic decrease in luminosities of nebulæ, or a systematic increase in luminosities of brightest stars along the sequence of classification. Data from other sources suggest that both phenomena may be involved, but that the variation in brightest stars is the more important. In any case, the results justify the reduction of the intermediate spirals and the irregulars to the system of the late-type spirals. The corrections do not materially affect the mean difference or the form of the frequency-distribution, but they reduce the dispersion to the value mentioned above, namely, 0.9 mag.

when distances of clusters are available. The clusters, as previously mentioned, are very similar among themselves. Their relative distances are so well known that the absolute distance of any one cluster will determine the absolute distances of them all. Fortunately, the distance of the nearest cluster can be estimated by means of the same criterion that has been used for the resolved field-nebulæ, namely, by the brightest stars.

The Virgo cluster, which is the nearest, is comparatively rich in spirals. Stars can be identified in most of the late-type spirals ($S_c$), but in very few of the earlier types.[9] The brightest stars actually observed range from about $m = 19$ to $m = 21$, with a mean value between 20 and 20.5. The true mean, including the unresolved spirals as well as the resolved, must be estimated from the frequency-distribution of the observed magnitudes. The adopted value, $m = 20.6$, is probably of the right order. The corresponding absolute magnitude of these stars is $M_s = -6.1$, for the effect of selection does not apply to a cluster in which the members are all at about the same distance from the observer. Thus the difference between apparent and absolute magnitude, known as the modulus of the distance, is

$$m - M = 26.7$$

and the distance[10] is about seven million light-years.

Since the observations are incomplete—since all the spirals in the clusters have not been resolved—the results are not final. Nevertheless the general order of the distance is established, and a large sample collection of clus-

9 This conclusion is based on unpublished data assembled by the writer. References to fragmentary data are found in Hubble and Humason, "Mt. Wilson Contr.," No. 427; *Astrophysical Journal*, 74, 43, 1931; also, in Hubble, *Red-Shifts in the Spectra of Nebulæ*, being the Halley Lecture (Oxford), 1934.

10 The distance is given by the formula

$$\log d \text{ (l.y.)} = 0.2\,(m - M) + 1.513$$

which follows directly from the definition of *M*, namely, that $M = m$ when $d = 32.6$ l.y. $= 10$ parsecs. Shapley has derived a distance for the Virgo cluster of 10.5 million l.y. *Harvard College Observatory Bulletin*, No. 873, 1930.

M90→

# PLATE XIII

## *Nebulæ in the Virgo Cluster (M90 and M100).*

THE Virgo Cluster (distance = seven million l.y.) is the nearest of the great clusters of nebulæ and is exceptional in containing a considerable number of intermediate and late spirals in addition to the early spirals and ellipticals which dominate the normal clusters. For these reasons, the Virgo Cluster furnishes an opportunity for comparing the stellar contents of nebulæ at various stages in the sequence of classification.

Most of the $S_c$ nebulae in the cluster are partially resolved with the 100-inch reflector, and the apparent faintness of the brightest stars indicates the order of the distance. The brightest stars of all are found in M100 (shown in the Plate).

Only a few of the $S_b$ nebulae can be resolved, and their brightest stars are systematically fainter than those in the $S_c$ nebulae. M90 is a borderline example; a few very faint objects have been uncertainly identified as individual stars. Cluster members of earlier types are, in general, unresolved.

The photographs were made with the 100-inch reflector—that of M90, on December 21, 1935; that of M100, on January 21, 1925. In both cases, north is at the top (outer margin of the page), and the scale is 1 mm. = 4″.25.

M100→

ter-nebulæ of all types is available for comparison with resolved nebulæ in the general field. Apparent magnitudes of the cluster-nebulæ range from $m = 10.2$ to $m = 15$ or fainter, and the most frequent magnitude is probably about 12.7 ±. The brighter limit is well determined, but the fainter limit, where the occasional cluster-members are difficult to distinguish from field-nebulæ, is uncertain. The corresponding absolute magnitudes, derived from the modulus, $m - M = 26.7$, are: range, −16.5 to −11.7 or fainter; most frequent, −14 ±. The cluster-nebulæ are thus comparable with resolved field-nebulæ. The luminosity-function for the latter can be safely generalized to include nebulæ of all types.

### (c) VELOCITY-DISTANCE RELATION.

THE third problem to which the criterion of brightest stars has been applied is the numerical formulation of the velocity-distance relation. The relation, when calibrated, indicates the distances, and hence the intrinsic luminosities, of all nebulæ whose velocities are known. Since all types of nebulæ are included, the information is an important contribution to the problem of the luminosity-function.

The velocity-distance relation is derived from very simple data, namely, red-shifts in nebular spectra and apparent magnitudes of nebulæ or of their brightest stars. The velocities (simple multiples of red-shifts) are corrected for solar motion, but, of course, include the unknown peculiar motions of the nebulæ. The peculiar motions appear as accidental errors in the distance effect and, although they are unknown for individual nebulæ, their influence on statistical averages can be estimated and partially eliminated.

Magnitudes are corrected for galactic obscuration and also for certain effects due to the red-shifts, which will be more fully discussed in Chapter VIII. For the present, the term "corrected" magnitude, $m_c$, will be used without further explanation, instead of the "observed" magnitude $m_0$. The relation is

$$m_c = m_0 - \Delta m_0$$

where $\Delta m_0$ is the effect of red-shifts. The numerical values of the corrections, $\Delta m_0$, increase with red-shifts, but do not become important until the shifts correspond to velocities of 3,000 miles/sec., and greater.

Correlations between logarithms of velocities, log $v$, and apparent magnitudes, $m_c$, have been derived from three independent sets of data. The first correlation is between velocities of resolved field-nebulæ[11] and magnitudes of the brightest stars (29 cases); the second is between velocities of field-nebulæ of all types, whether resolved or unresolved, and magnitudes of the nebulæ themselves (103 cases); the third is between velocities of clusters (each cluster-velocity representing the mean of all velocities observed within the cluster), and the magnitudes of the fifth brightest nebulæ in the clusters (10 cases).[12]

The three correlations,[13] which are shown graphically in Figs. 12, 13, 14, are represented by the formulas

$$\begin{aligned} \log v &= 0.2\, m_c - 1.197 \text{ (brightest stars)} \\ &= 0.2\, m_c + 0.553 \text{ (field-nebulæ)} \\ &= 0.2\, m_c + 0.818 \text{ (clusters)} \end{aligned}$$

where the velocity, $v$, is expressed in miles per second. These formulas represent precisely the same velocity-distance relation. The differences in the constants refer to differences in the absolute magnitudes of the objects

[11] Members of the local group were not included. They are so near that peculiar motions may be large compared to distance effects and, moreover, it is not yet certain whether distance-effects exist within the group.

[12] The great clusters are remarkably similar, and significant correlations can be derived by using any one of say the first ten magnitudes, or the mean of the first several. Actually, the fifth brightest magnitude differs systematically from the mean of the ten brightest, by less than 0.05 mag. Without regard to sign, the average residual is less than 0.1 mag.

[13] The correlations represent analyses of all data available at the end of 1935. Details will be published in "Mt. Wilson Contr.," No. 549.

used as criteria of distances. The mean absolute magnitude of one criterion, namely, that of the brightest stars, is already known, therefore the direct relation between velocities and distances can be expressed numerically, and the absolute magnitudes of the other criteria can be determined. Both results are of major importance.

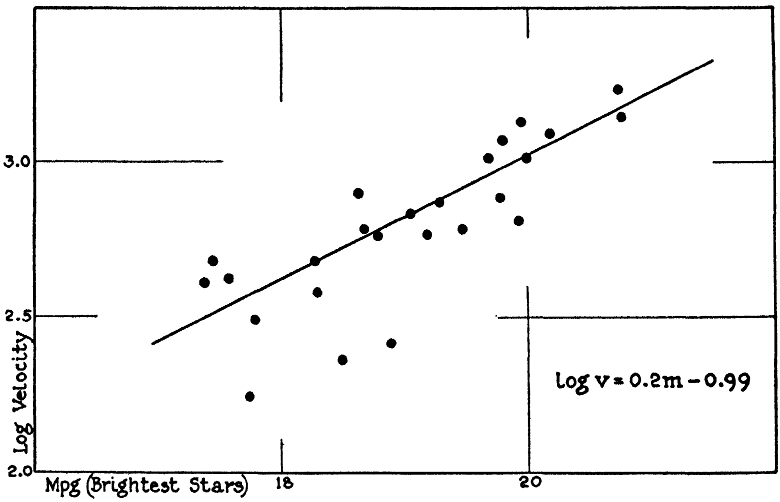



FIG. 12. *Velocity-Distance Relation from Brightest Stars.*

Logarithms of velocities (in km/sec. and corrected for solar motion) are plotted against apparent magnitudes of brightest stars in nebulæ (corrected for local obscuration). Nebulæ in the local group are omitted. The three points below the main body of the correlation diagram may be accounted for by peculiar motions.

## *Calibration of the Velocity-Distance Relation.*

THE calibration of the velocity-distance relation is as follows. The expression for distance (in light-years) is

$$\log d = 0.2\,(m_c - M) + 1.513$$

where $M$ is the absolute magnitude of the object whose

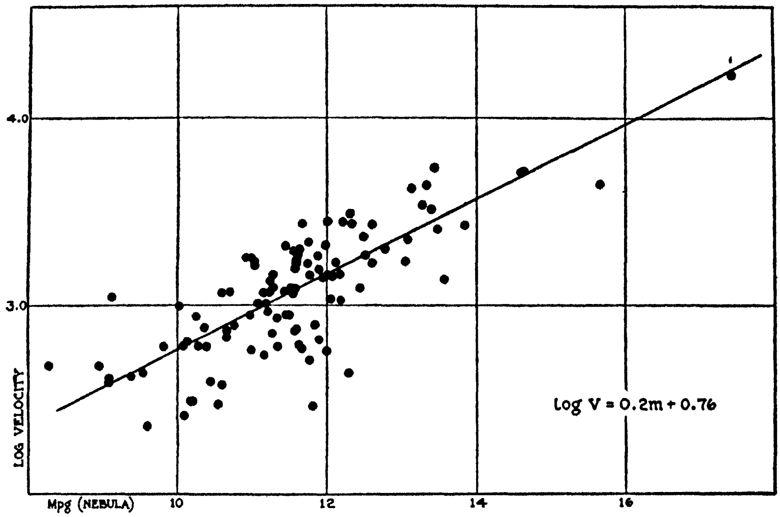



FIG. 13. *Velocity-Distance Relation for Field-Nebulæ.*

Logarithms of velocities (in km/sec. and corrected for solar motion) are plotted against apparent magnitudes of nebulæ (corrected for local obscuration). Effects of peculiar motions are appreciable among the brighter (nearer) nebulæ.

apparent magnitude is $m_c$. Therefore, in the case of the brightest stars, where $M = -6.35$,

$$0.2\, m_c = \log d - 2.783.$$

When this value of $0.2\, m_c$ is substituted in the correlation for brightest stars,

$$\log v = \log d - 3.98$$
$$v = 0.000105\, d$$
$$d = 9550\, v.$$

Thus the apparent velocity of a nebula is about 105 miles/sec. for each million light-years of distance (550 km/sec. per million parsecs).

Another method of analyzing the data leads to slightly smaller, but more probable, values of the coefficients. The

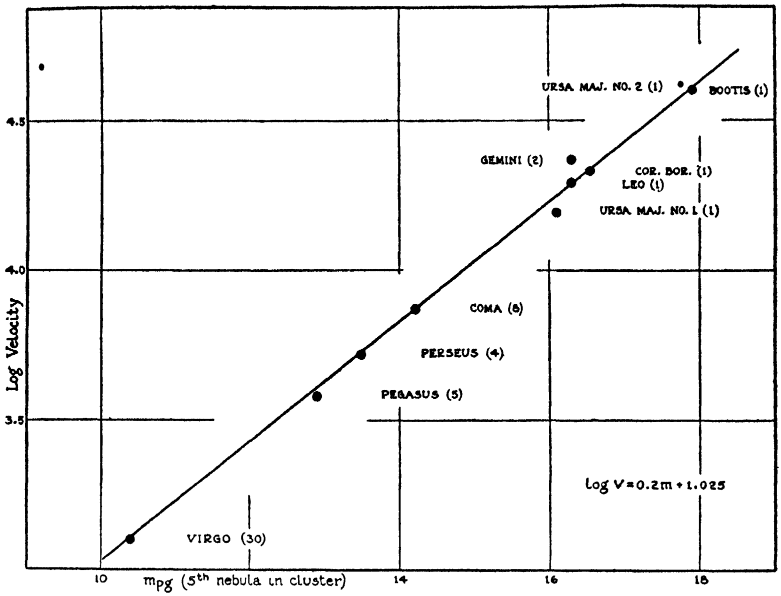



FIG. 14. *Velocity-Distance Relation for Clusters.*

Logarithms of velocities (in km/sec. and corrected for solar motions) are plotted against apparent magnitudes of the fifth brightest nebulæ in clusters (corrected for galactic obscuration). Each cluster-velocity is the mean of the various individual velocities observed in the cluster, the number being indicated by the figure in brackets.

correlations between log $v$ and $m_c$, which are more precisely described as velocity-magnitude rather than velocity-distance relations, are subject to uncertainties arising from the unsymmetrical distribution of residuals due to peculiar motions. The lack of symmetry is difficult to evaluate and correct, especially in the case of the resolved nebulæ for which the distance-effects are necessarily small. Nevertheless, these nebulæ are the only ones whose distances are derived from a fundamental criterion, and they must be used for the numerical calibration.

The uncertainties in the velocity-magnitude relation may be avoided by using simple velocities rather than log

$v$, and simple distances calculated from the moduli, $m_c - M$. The residuals due to peculiar motions are then symmetrically distributed and tend to cancel out when mean velocities of large numbers of nebulæ are considered. The mean distances, on the contrary, will be systematically in error[14] but the necessary corrections are small and can be readily calculated. In the case of the resolved nebulæ, where the dispersion in the absolute magnitudes of the brightest stars is 0.4 mag., the mean distance must be increased by only 1.7 per cent.

The velocity-distance relation starts from the origin of coördinates (the observer) and is known to be approximately linear (from observations of clusters and remote field-nebulæ). Now the mean velocity and the mean distance of the resolved nebulæ, as a group, furnish a point through which the relation must pass. This point determines the slope of the relation (the rate of increase of velocity with distance). The data now available—for the 29 resolved field-nebulæ—lead to the values which are given below,

$$\begin{aligned} \bar{d} &= 4.07 \text{ million l.y.} \\ \bar{v} &= 412 \text{ miles/sec.} \\ v/d &= 101 \text{ miles/sec. per million l.y.} \\ &= 530 \text{ km/sec. per million parsecs.} \end{aligned}$$

The value of $\bar{d}$ includes the correction of 1.7 per cent which was mentioned in the preceding paragraph. When the correction is omitted, the constant in the velocity-magnitude relation can be computed on the assumption that peculiar motions are negligible. The value, constant

[14] Since the frequencies of the absolute magnitudes of the distance-criteria are symmetrically distributed about the mean value (luminosity-functions are normal error-curves), frequencies of the logarithms of distances, for objects of a particular apparent magnitude, are also distributed normally. Therefore, the simple distances will show a skew frequency-distribution. However, the relation between the means is very simple;

$$\log \bar{d} = \overline{\log d} + 1.152\,\sigma^2$$

where $\sigma$, the dispersion in log $d$, is one-fifth of the dispersion in $M$.

= −1.204, is closely comparable with the value, −1.197, previously adopted for the correlation.

Similar calibrations of the velocity-distance relation may be made from the field-nebulæ in general and from the clusters. The results are comparable with those from resolved nebulæ but they are not independent because the distances, or, more precisely, the absolute magnitudes of the distance-criteria, are derived from the fundamental criterion of brightest stars. Therefore, the agreement in the various calibrations emphasizes the consistency of the data, but contributes very little to the absolute scale of the distance-effects.

The calibration of the distance-effect directly from the velocity-distance relation permits an estimation of the dispersion among peculiar motions of nebulæ. The total dispersion of residuals from the relation, expressed in terms of velocities, is about 155 miles/sec. The peculiar motions contribute about 125 miles/sec. and the remainder is divided between accidental errors and the scatter in the absolute magnitudes of brightest stars.

Although the coefficient, 101 miles/sec. per million l.y., is the more reliable measure of the absolute distance-effect, the coefficient derived from the velocity-magnitude relation serves well enough for the comparison of distance criteria in the three correlations. The formula might be applied to each nebula and cluster whose velocity is known. Velocities, in miles/sec., divided by 105, give distances in millions of light-years. The distances, together with the apparent magnitudes, determine absolute magnitudes. Analysis of the list of absolute magnitudes thus assembled would determine the luminosity-function, including the form of the frequency-distribution, the mean magnitude, and the dispersion.

The information, however, can be derived in a much simpler way, directly from the constants and the residuals in the three correlations. The differences in the mean absolute magnitudes of the various criteria of distances are evidently five times the differences in the constants appearing in the correlation formulas.[15] Thus the field-nebulæ are brighter than the brightest stars, by the amount $5 \times (1.197 + 0.553) = 8.75$ mag. The fifth brightest nebulæ in clusters differ from the brightest stars by $5 \times (1.197 + 0.818) = 10.05$ mag. The absolute magnitudes of the three criteria are therefore

$M_s = -6.35$ (brightest stars)
$\overline{M} = -15.1$ (field-nebulæ)
$M_5 = -16.4$ (fifth brightest nebulæ in clusters).

## *Luminosities of Cluster-Nebulæ.*

THE fifth brightest nebulæ in clusters average half a magnitude, or less, fainter than the brightest nebulæ. Therefore, the mean absolute magnitude of the latter is about −16.9. The difference between the brightest and the mean, or the most frequent, magnitude in a cluster is not precisely determined, but the approximate value, 2.5 mag., has long been used. The mean absolute magnitude of cluster-nebulæ[16] thus appears to be of the order of −14.4, as compared with the value, −14.0, previously derived from the brightest stars in cluster members. The mean of these

[15] When $m_c$ is eliminated from the two equations

$$\log v = 0.2\, m_c + \text{constant}$$
$$\log d = 0.2\, m_c - 0.2\, M + 1.513$$

then

$$\log v - \log d = 0.2\, M + (\text{constant} - 1.513)$$

and from the calibration by brightest stars,

$$0.2\, M + (\text{constant} - 1.513) = -3.98.$$

Therefore, differences in the constants appearing in the three correlations equal one fifth the differences in $M$ for the various criteria.

[16] The agreement is probably better than the comparison suggests. The ten clusters from which the difference between the first and fifth brightest nebulæ is derived, include two exceptional cases. One (Perseus Cluster) contains an outstanding giant member (NGC 1275) and the other (Pegasus), is a very sparse cluster with two unusually bright members. If these clusters are omitted, the normal difference between the first and fifth nebulæ would be about 0.3 mag., the brightest nebulæ about −16.7, and the mean or most frequent magnitude about −14.2, in excellent agreement with the field-nebulæ.

two approximate values, namely −14.2, agrees with that found for resolved nebulæ in the general field, where brightest stars are used as the criterion of distance.

*Effect of Selection on Statistical Criteria of Distances.*

These values refer to the mean of all nebulæ *in a given volume of space,* which is designated by the symbol $M_0$. Another symbol, $\bar{M}$, has been employed for the mean absolute magnitude of the field-nebulæ used in the correlation. $\bar{M}$ refers to the mean of all nebulæ *of a given apparent magnitude.* The two quantities would be the same only if the intrinsic luminosities of all nebulæ were precisely equal. Actually, some nebulæ are as much as ten times brighter than the average and others are ten times fainter. A list of nebulæ of a particular *apparent* luminosity contains a mixture of nebulæ of different *intrinsic* luminosities, scattered over a great range in distance. Some are faint and near, others are bright and remote (Fig. 15).

Now the distribution of nebulæ in space is approximately uniform, and the statement applies to nebulæ of any particular intrinsic luminosity as well as to the nebulæ of all luminosities. The uniform distribution, together with the scatter in intrinsic luminosities, leads to a curious result. Consider for a moment, the list of nebulæ with the same apparent luminosity. The intrinsically faint nebulæ, since they are near, are distributed through a relatively small volume of space, while the intrinsically bright objects, since they are remote, are distributed through a relatively large volume of space. Therefore, the intrinsically bright nebulæ greatly outnumber the faint ones among nebulæ with a given *apparent* luminosity. Evidently $\bar{M}$ will be brighter than $M_0$. The situation arises whenever photometric criteria of distances, in which the dispersions are appreciable, are used in statistical investigations.

The complete solution of the problem can be stated as follows.[17] If the space distribution is uniform and if the luminosity-function (the frequency-distribution of abso-

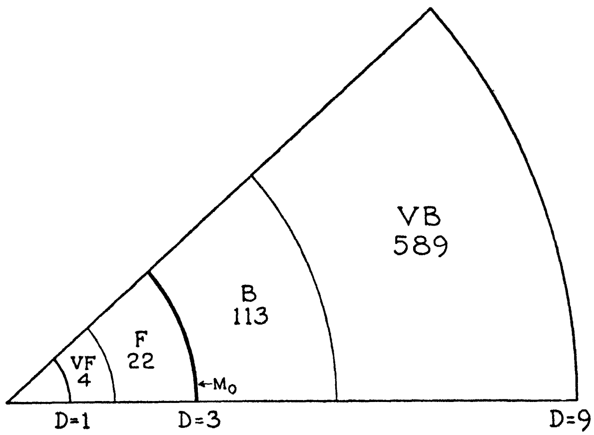



FIG. 15. *Distribution in Space of Nebulæ with the Same Apparent Luminosity.*

Among nebulæ which *appear* equally bright are included objects of various intrinsic luminosities, scattered over a wide range in distance. Thus, if the normal nebula ($M_0$) is at the distance, $D = 3$, the intrinsically faint ($F$) and very faint ($VF$) objects will be at smaller distances (down to about $D = 1$), while the bright ($B$) and very bright ($VB$) objects will be spread over greater distances (out to about $D = 9$). The relative volumes of space through which the four grades ($VF$ to $VB$) are scattered, are indicated by the numbers below the symbols.

Since the distribution of nebulæ is approximately uniform, there will be many more intrinsically bright, than intrinsically faint nebulæ, among those which appear equally luminous. Therefore, $\overline{M}$, the mean absolute magnitude of nebulæ *with a given apparent magnitude,* will be brighter than $M_0$, the mean absolute magnitude of nebulæ *in a given volume of space.*

lute magnitudes *in a given volume of space*) is a normal error-curve with a maximum, $M_0$, and a dispersion, $\sigma$, then the frequency-distribution of absolute magnitudes

[17] Malmquist, "On Some Relations in Stellar Statistics," *Arkiv för Mat., Astr., och Fysik,* 16, No. 23, 1921.

*for a given apparent magnitude* will be a normal error-curve with the same dispersion, $\sigma$, but with a maximum, $\bar{M}$, which is brighter than $M_0$ by the amount

$$M_0 - \bar{M} = 1.382\,\sigma^2.$$

The difference is the effect of selection previously mentioned in connection with the dispersion among brightest stars. $\bar{M}$ must be used when nebulæ are selected on the basis of apparent magnitudes, as they are in the velocity-distance relation for field-nebulæ, and in surveys to successive limiting magnitudes. $M_0$ applies to nebulæ selected on the basis of stars involved, and to clusters.

### *Luminosities of Field-Nebulæ.*

THE value, $\bar{M} = -15.1$, is derived from the correlation for field-nebulæ of all types, and the value, $M_0 = -14.2$, is derived from resolved field-nebulæ selected on the basis of brightest stars. The dispersion for the latter group is $\sigma = 0.84$. Therefore, the resolved field-nebulæ furnish the value

$$\begin{aligned}\bar{M} &= M_0 - 1.382\,\sigma^2 \\ &= -14.2 - 0.93 \\ &= -15.13\end{aligned}$$

which is about that actually observed among the field-nebulæ of all types. The dispersion for the latter group can be calculated directly from the residuals in the velocity-distance relation. When effects of peculiar motions are removed, the dispersion in $M$ is of the order of $\sigma = 0.85$ or slightly less, in substantial agreement with that for the resolved nebulæ. The dispersion among cluster-nebulæ has not been precisely determined, but it appears to be of the same general order as that for the field-nebulæ. The numerical results are thus thoroughly consistent and no conspicuous differences are found among the various groups.

A more detailed analysis of the data used in the velocity-distance correlation for field-nebulæ, yields the following values of $\bar{M}$ for various individual types.

TABLE VI.

*Absolute Magnitudes of Various Types of Nebulæ.*

| Type | No. | $\bar{M}$ |
|---|---|---|
| E0–E2 | 11 | −15.3 |
| E3–E7 | 12 | 15.2 |
| $S_a$–$SB_a$ | 23 | 15.2 |
| $S_b$–$SB_b$ | 27 | 15.1 |
| $S_c$–$SB_c$ | 25 | 15.1 |
| Ir | 5 | −14.4 |

The values are all comparable, except that for the five irregular nebulæ, and the numbers are too limited to establish fully the slight systematic differences between the early and the late types, which are suggested by the data. The abnormally low luminosity of the irregular nebulæ appears to be real since it is confirmed by the mean magnitude of the four additional irregulars in the local group.

The two sample collections of 125 resolved nebulæ and 103 field-nebulæ, have 29 objects in common, but otherwise they are independent. Analyses of the two lists and of the ten clusters lead to the conclusion that all nebulæ are comparable systems, whether they are members of clusters or isolated in the general field, and regardless of type. The only exception appears to be that the rare, irregular nebulæ may average about half as bright as other nebulæ. If systematic differences in the mean luminosities of other types exist, they are so small that they will emerge only from large bodies of very accurate data.

The realm of the nebulæ, or at least the portion that has been explored, thus appears as a vast region of space in which comparable systems are uniformly distributed. The scale of nebular distances is known. Intrinsic luminosities have been discussed at considerable length, but the results may be summarized for convenience. The nebulæ average about 85 million times as bright as the sun. The brightest are about 10 times brighter than the average and the faintest are about 10 times fainter, but about half of the nebulæ are within the narrow range from one half to twice the average luminosity of them all. Other general features concerning which information can be assembled are absolute dimensions and masses.

*Dimensions of Nebulæ.*

LINEAR dimensions are derived from the diameter-luminosity relations presented in the discussion of classification (Chap. II). At any stage in the sequence of classification, the relation is

$$m + 5 \log D_a = \text{constant}$$

where $m$ is the apparent magnitude of a nebula and $D_a$ is the angular diameter in minutes of arc. The constant increases continuously through the sequence, from globular nebulæ to open spirals, and it has been evaluated for the various standard stages or types. Therefore, when the intrinsic, or linear, diameter is known for any one type, it can be readily calculated for other types.

The reduction from apparent to intrinsic luminosities and dimensions is made by means of the relations

$$\log d \text{ (l.y.)} = 0.2\,(m - \overline{M}) + 1.513$$

from which

$$m = 5 \log d - 22.665$$

and

$$D_a = 3438 \times \frac{\text{linear diameter}}{\text{distance}}$$

the last of which is merely a definition. Introducing these expressions into the diameter-luminosity relation, the linear diameter (in l.y.), termed $D_l$, is found to be

$$\log D_l = 0.2 \times \text{constant} + 0.997.$$

Numerical values of $D_l$ follow at once from the previously determined values of the constant for the various types of nebulæ (Chap. II).

It may be emphasized that the angular diameters of nebular images are rather arbitrary quantities. They vary widely with exposure-conditions and methods of measurement, and the values of the constants vary accordingly. Any homogeneous body of data furnishes fairly reliable relative values of the constants along the sequence, but the zero-point of the scale depends upon the specifications (exposure and method of measurement). With these reservations, the approximate[18] linear diameters of the main bodies for various types of nebulæ are listed in Table VII. They correspond to the constants previously given for a particular body of data.

TABLE VII.

*Diameters of Nebulæ.*

| Type | Diameter | Type | Diameter |
|---|---|---|---|
| E0 | 1,900 l.y. | (S0) | (5,300) l.y. |
| E3 | 2,800 | $S_a$ | 6,000 |
| E7 | 4,800 | $S_b$ | 7,600 |
| $SB_a$ | 5,500 | $S_c$ | 9,500 |
| $SB_b$ | 6,300 | Ir | 6,300 |
| $SB_c$ | 8,700 | | |

## *Masses of Nebulæ.*

THE determination of reliable masses is an outstanding problem in nebular research. Two methods have been employed, leading to widely different results. Nevertheless, when considered together, they suggest a general order of

18 The linear diameters are not strictly accurate because the effects of dispersion have been neglected. The numerical values correspond to $\overline{\log D}$, the mean of the logarithms of the angular diameters, and not to $\log \overline{D}$, the logarithm of the mean of the diameters. The two terms are not necessarily the same and the difference is a function of the dispersion in $D$. The calculation of dispersions for various types, and the necessary corrections, are deferred until a body of definitive data is available for investigation.

The dispersion in distances corresponding to a given apparent magnitude is also involved, although that quantity is known and the effects can be determined.

mass which may be tentatively adopted pending a more satisfactory solution of the problem.

One method is based upon spectrographic rotation.[19] The nebulæ, in general, are lens-shaped systems in rapid rotation about their minor axes. In a few cases, where the nebulæ are oriented nearly edge-on, radial velocities have been measured at various points along the major axes, and the nature of the rotation has been determined. The nuclear regions, out to considerable distances, appear to hold their forms and to rotate like solid bodies. The outer regions, however, lag behind, and the lag increases with distance from the nucleus. Interpretation of this behavior on simple dynamical principles is not very clear. It suggests uniform density throughout a large nuclear region, in sharp contrast with the observed luminosity-gradient.

Apart from uncertainties in the dynamical picture, the orbital motion of a point in the equatorial plane of a nebula should be determined by the mass of material inside the orbit. That mass can be calculated in much the same way in which the mass of the sun is found from the orbital motion of the earth (or of the other planets). The exterior mass in the nebula, beyond the orbit under consideration, can only be estimated. If the most distant points with known velocities are used for the calculations, the exterior mass should be a relatively small fraction of the interior mass, and errors in the estimates should be still smaller.

Spectrographic masses of four nebulæ, together with that of the galactic system are listed in Table VIII. They range from about 1,000 million to 200,000 million times the mass of the sun, and average about 50,000 million suns. These nebulæ, however, are much larger and brighter than average and doubtless their masses also are abnormally large. The appropriate correction is speculative, but, as a first approximation, we may assume that masses vary directly with luminosities. In this case, the

19 First used by Oepik, "An Estimate of the Distance of the Andromeda Nebula," *Astrophysical Journal*, 55, 406, 1922.

average nebulæ would be about 2,000 million times as massive as the sun.

The second method of estimating nebular masses was recently employed by Sinclair Smith.[20] Analysis of radial velocities of 32 members of the Virgo cluster, suggested a velocity of escape for the cluster of about 1,500 km/sec. This quantity measured the gravitational field and, therefore, the total mass of the cluster. The total mass, divided by the number of members, gave the average mass per nebula. The latter was of the order of $2 \times 10^{11}$ suns, or about 100 times the order of mass suggested by spectrographic rotations. Internebular material in the cluster was neglected. Such material may exist, but observations give no reason to suppose that the amount is large compared with the amount of material concentrated in the nebulæ.

TABLE VIII.

*Spectrographic Masses of Nebulæ.*

| Nebula | Type | Mass | Luminosity | $M^*/L$ |
|---|---|---|---|---|
| M33 | $S_c$ | $10^9$ suns | $1.45 \times 10^8$ suns | 7 |
| M31 | $S_b$ | $3 \times 10^{10}$ | $1.7 \times 10^9$ | 18 |
| NGC 4594 | $S_a$ | $3.5 \times 10^{10}$ | $1.5 \times 10^9$ | 23 |
| NGC 3115 | E7 | $9 \times 10^9$ | $1.6 \times 10^8$ | 56 |
| Gal. Sys. | ($S_c$?) | $2 \times 10^{11} \pm$ | | |
| Mean | | $5 \times 10^{10}$ | | 26 |

* *M* represents mass and not absolute magnitude. The variation of the ratio, *M/L*, with nebular type is suggestive, but the data are too few to establish it as a general characteristic. The results for NGC 3115 are derived from unpublished measures of rotation by Humason.

The discrepancy seems to be real and important. The dynamical problem presented by the clusters appears to be simpler than that of nebular rotation, and to this extent the cluster-masses may carry the greater conviction. The cluster-masses are, in a sense, upper limits, while the

[20] *The Mass of the Virgo Cluster,* "Mt. Wilson Contr.," No. 532; *Astrophysical Journal,* 83, 23, 1936.

rotational masses, because of assumptions concerning material in the extreme outer regions, may possibly be regarded as lower limits. However, the results must be considered as unsatisfactory until the discrepancy is very materially reduced.

The discussion of masses completes the preliminary reconnaissance of the nebulæ. The investigations are beset with uncertainties, and the numerical results are mainly estimates which will be revised when more elaborate techniques and larger telescopes have been applied to the problems. Nevertheless, valuable information has been assembled concerning the scale of nebular distances, the general features of nebulæ, such as their luminosities, dimensions, and masses, their structure and stellar contents, their large-scale distribution in space, and the curious velocity-distance relation. These data serve to sketch the broad features of the realm of the nebulæ. The outlines are perhaps not yet in a definitive form, but they are sufficiently clear to permit individual problems to be investigated with some knowledge of their relation to the general scheme.

CHAPTER VIII

# THE REALM OF THE NEBULÆ

THE preceding chapters have described the apparent features of nebulæ and their distribution, the development of methods for investigating their intrinsic characteristics, and the nature of the results to which the new methods have led. It is now possible to consider the realm of the nebulæ as a unit, and to discuss the observable region as a sample of the universe.

The explorations started from within an isolated stellar system—a nebula. They penetrated through the swarm of stars into a vast region of space thinly populated with other nebulæ. The nebulæ are all curiously similar—members of a single family. Since their intrinsic luminosities are known, their distances can be determined and their distribution mapped. They are found singly, in groups, and, occasionally, in great clusters, but when very large regions are compared, the tendency to cluster averages out and one region is very much like another.

Preliminary reconnaissance indicated approximate uniformity throughout the whole of the observable region. Obviously, the next step was to follow the reconnaissance with a careful survey and to interpret the results in the light of all available data concerning the nebulæ themselves. As information accumulated, the results could be reinterpreted and the surveys repeated with greater precision. Thus, by successive approximations, it would be possible to approach a comprehensive knowledge of the sample of the universe open to our inspection. Only then could extrapolations be pushed beyond the reach of telescopes, and conclusions drawn that were more significant than mere speculations.

## *Surveys to Successive Limits.*

A STEP in the slow advance is represented by the surveys mentioned in the discussion of apparent features of nebular distribution (Chap. III). Five surveys with large reflectors have been completed, to limiting magnitudes 18.5, 19.0, 19.4, 20.0, and 21.0, respectively. A casual examination of the results suggests that the distribution in space is roughly uniform. Nevertheless, detailed analysis reveals an apparent thinning out. The departures from uniformity, although small, increase systematically with the distances to which the surveys extend.

Now red-shifts are known to introduce these very effects. By reducing the apparent luminosities, red-shifts increase apparent distances; therefore, the fainter nebulæ appear to be scattered through a larger volume of space than actually happens. The numerical values of the expected departures from apparent uniformity differ slightly according to the particular interpretations of red-shifts that are used in the calculations. On all interpretations, however, they are of the same order as the observed departures. Therefore, after the luminosities in the surveys have been corrected for red-shifts, the nebular distribution again appears to be uniform and now to a very close approximation. As a final step, the argument is reversed. The distribution is assumed to be strictly uniform and the observed departures are used to test the various interpretations of red-shifts.

The essential data from the surveys are the average numbers of nebulæ per unit area, $\overline{N}$, equal to, or brighter than, various apparent magnitudes, *m*. These numbers are denoted by $\overline{N}_m$, where the unit area is one square degree (about five times the area of the full moon). The nebular distribution is indicated by the manner in which $\overline{N}_m$ increases with *m*.

One of the five surveys, namely, that to $m = 19.0$, was made by Mayall with the 36-inch reflector at the Lick Observatory.[1] The other surveys were made with the 60-inch and the 100-inch reflectors at Mount Wilson.[2] The data represent counts of nebulæ in about nine hundred fields well distributed over the whole of the northern polar cap, and over one half or more of the southern cap. The galactic belt was avoided because of the uncertainties introduced by local obscuration.

The numbers of nebulæ actually identified were reduced to standard conditions by corrections for variation in definition, atmospheric extinction, and galactic obscuration. The corrected counts, representing more than a hundred thousand nebulæ, were transformed to numbers per square degree in order to facilitate the comparison of mean results obtained from the different telescopes. Investigations of possible sources of errors suggest that the corrected counts are probably satisfactory as regards completeness, but that the very faint, limiting magnitudes are necessarily subject to some uncertainties.

The surveys individually exhibit the same general features of distribution over the sky that were described in Chapter III. Mean results for the two polar caps are similar; there are no systematic variations over the face of the sky; the scatter among the individual samples diminishes as the average size of samples increases. The large-scale distribution over the sky, as derived from each survey, is approximately uniform.

### *Distribution of Nebulæ in Depth.*

EACH survey furnishes the number of nebulæ within a certain sphere whose radius is indicated by the limiting magnitude of the survey. The apparent distribution in depth may be derived by comparing the numbers of nebu-

1 "A Study of the Distribution of Extra-Galactic Nebulæ Based on Plates Taken with the Crossley Reflector," *Lick Observatory Bulletin*, No. 458, 1934.

2 The surveys to $m = 19.4$ and $m = 20.0$ are found in Hubble, *The Distribution of Extra-Galactic Nebulæ*, "Mt. Wilson Contr.," No. 485; *Astrophysical Journal*, 79, 8, 1934. The surveys to $m = 18.5$ and $m = 21$ are unpublished but will appear in a forthcoming "Mt. Wilson Contr."

læ in the successive spheres. Specifically, the problem involves the determination of $\bar{N}_m$ as a function of $m$.

If the distribution were uniform, the numbers of nebulæ would be proportional to the volumes of space through which they are scattered. The data would then be represented by the linear relation[8]

$$\log \bar{N}_m = 0.6\, m + \text{constant.}$$

This straight line, together with the relation actually observed, is shown graphically in Fig. 16. The two relations are obviously not parallel. The numbers of nebulæ increase less rapidly than the corresponding volumes of space; in other words, the nebulæ appear to thin out more and more as the distance increases. The departures from uniformity are already appreciable in the shallowest survey (the slope of the observed relation is less than the theoretical slope), and they increase steadily with the limiting magnitudes.

The source of the departures is evidently in $N$ or in $m$, for no other quantities are observed in the surveys. If $N$ alone were involved, the observed relation would represent the true distribution. The galactic system would then be considered as lying near the center of a large and more or less spherical system of nebulæ which thinned out in all directions. On the other hand, if $m$ were involved, the apparent luminosities of nebulæ would be fading with distance faster than could be accounted for by the familiar law of inverse-squares. It would then be necessary to discover the cause of the fading, and to eliminate the effects, before the true distribution could be examined. With this end in view, the departures may be expressed as incre-

[8] The volume of space (and hence the number of nebulæ) out to a given limit, is proportional to the cube of the distance. To the limiting magnitude, $m$, the distance, $d$, is

$$\log d = 0.2\, m + \text{constant.}$$

Hence $$\log V = 0.6\, m + \text{constant}$$

and $$\log N = \log V + \text{constant.}$$

ments to the magnitudes, $\Delta m$, and the observed relation represented by the formula

$$\log \bar{N}_m = 0.6\,(m - \Delta m) + \text{constant}.$$

The problem now is to examine all known effects which might reduce the apparent luminosities, and to evaluate the probable departures, $\Delta m$, arising from these sources. If the known effects do not completely account for the observed departures, the residuals must be attributed either

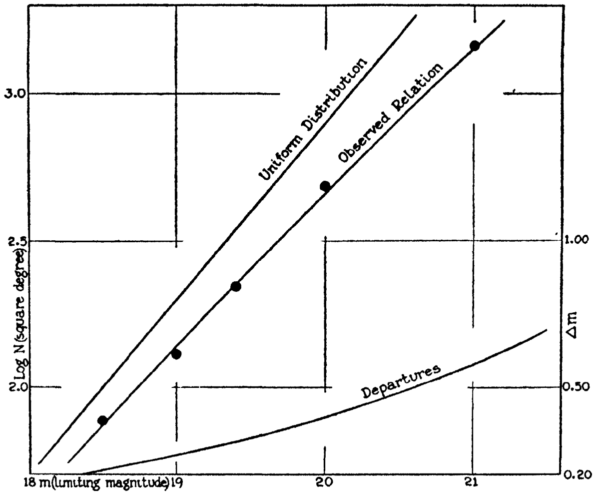



FIG. 16. *Apparent Distribution of Nebulæ in Depth.*

Each point on the Observed Relation represents the average number (actually log $N$) of nebulæ per square degree which are equal to, or brighter than, a particular apparent magnitude, as determined by an entire survey. The line through the points (Observed Relation) is a least square solution in the form, $\log N = 0.6\,(m - \Delta m) + \text{constant}$, derived from the assumption that $\Delta m$ is a linear function of distance.

In the vicinity of the galactic system, $\Delta m$ should be negligible, and the Observed Relation should coincide with a Uniform Distribution indicated by the straight line. As the surveys are extended to greater distances (fainter limiting magnitudes), $\Delta m$ increases, and the Observed Relation departs from the line representing Uniform Distribution. The departures, $\Delta m$ (horizonal displacements between the two lines), are plotted against $m$ (limiting magnitudes of the surveys) in the lowest curve. The departures are interpreted as effects of red-shifts.

to real deviations from uniformity, or to unknown sources of fading. The investigation proves to be unexpectedly simple. Red-shifts reduce apparent luminosities, and the effects increase with distance. The phenomena will be discussed later in considerable detail, but it will be convenient to anticipate one of the conclusions. Within the uncertainties of the data, red-shifts completely account for the observed departures.

Other possible sources of fading may be ignored because, if they made appreciable contributions, they would *overcorrect* the observed departures. The density of the nebular distribution would then increase radially in all directions—a conception so highly artificial that it would be seriously considered only as a last resort to save the phenomena.

The only known mechanism for increasing the apparent luminosities (and thus counteracting space-absorption or other sources of fading) is the possibility that abnormally high intensities in the ultraviolet regions of nebular spectra (conceivably arising from giant blue stars) might be moved into the photographic region by large red-shifts. This possibility has been investigated in various ways—for instance, by examining ultraviolet spectra of bright neighboring nebulæ and the colors of spirals in very distant clusters— and the effects of blue giants appear to be unimportant.

It does not seem possible, with the information so far obtained, to *undercorrect* the observed departures. Thus, in order to avoid the unwelcome idea of a spherically symmetrical, increasing density, uniform distribution must be postulated and space-absorption neglected. The observed departures then stand out as effects of red-shifts alone, and serve to test their interpretations.

### *Quantitative Description of the Distribution.*

UNIFORM distribution thus appears to be the most plausible interpretation of the nebular counts. At any rate, it may be stated with confidence that the distribution is uniform within the uncertainties of the data, and that the uncertainties are small. This conclusion is represented by the straight line in Fig. 16 and by the relation

$$\log \overline{N}_m = 0.6\ (m - \Delta m) - 9.09.$$

The value of the constant, as derived from the five surveys, is consistent with data from other sources. The most important of such data are those in the Harvard survey[4] of the brighter nebulæ, which is believed to be complete over the entire sky to about $m = 12.9$. Since effects of red-shifts, $\Delta m$, are negligible to this limit, the data can be used directly to determine the constant. When the great Virgo cluster is eliminated and the galactic belt is avoided, the data furnish the value −9.10, which is of the same order as that given above. When various corrections are applied, reducing the data to the scale of the deeper surveys, the agreement is less precise, but is still satisfactory in view of the limited numbers of nebulæ involved, and the various uncertainties.

The numerical value of the constant indicates the number of nebulæ brighter than any given apparent magnitude. The actual space-distribution—the number of nebulæ per unit volume of space—is readily derived with the help of the mean absolute magnitude for nebulæ of a given apparent magnitude, $\overline{M} = -15.1$, which was derived in Chapter VII. There is, on the average, about one nebula per $5 \times 10^{18}$ cubic light years.[5] The average distance

[4] Shapley and Ames, "A Survey of the External Galaxies Brighter than the Thirteenth Magnitude," *Harvard College Observatory Annals*, 88, No. 2, 1932.

[5] The distribution density of nebulæ, the number per unit volume, is

$$\log \rho = \log N - \log V.$$

Introducing the number of square degrees over the sky,

$$\log N = 0.6\ (m - \Delta m) - 9.09 + 4.62.$$

The volume of the sphere in cubic parsecs is

$$\log V = \log 4/3\,\pi + 0.6\ (m - \Delta m) - 0.6\,\overline{M} + 4.539.$$

Hence $\log \rho = 0.6\,\overline{M} - 9.62.$

Since $\overline{M} = -15.1$

$$\rho = 2 \times 10^{-19} \text{ (nebula per cu. l.y.)}$$

$$= 7 \times 10^{-18} \text{ (nebula per cu. parsec).}$$

between a nebula and its nearest neighbor is of the general order of two million light-years. The smaller separations among the immediate neighbors of the galactic system emphasize the relative isolation of the local group.

Since the average type among isolated nebulæ is probably about $S_b$, the average diameter is of the order of ten thousand light-years and the distribution may be roughly represented by nebulæ scattered at random with an average separation of about two hundred times the diameters. Tennis balls fifty feet apart suggest the relative scale.

The average mass of nebulæ is uncertain, but the two values previously mentioned, $2 \times 10^9$ suns, from spectrographic rotations, and $2 \times 10^{11}$ suns, from radial velocities in the Virgo cluster, may be tentatively used as minimum and maximum estimates. Introducing these values into the expression for space-distribution, the smoothed-out density of nebular material in space, in grams per cubic centimeter, is found to be

$$\varrho = 10^{-30} \text{ (minimum)}$$

or

$$\varrho = 10^{-28} \text{ (maximum)}.$$

### *Internebular Space.*

These values include all the material that can be actually observed with present methods. The question of matter in internebular space is wholly speculative. The only observational evidence which bears on the question is the complete absence of any sensible obscuration out to the limit of the deepest survey. Space-absorption, if it exists, is probably less than 0.1 mag. (about 10 per cent), in a light-path of the order of a hundred million parsecs (326 million l.y.).

Obscuration by diffuse material varies enormously with the form of the material. Three general forms may be distinguished, following H. N. Russell of Princeton, as gas (molecules, atoms, and electrons), dust (particles with diameters comparable with the wave-length of light) and chunks (diameters large compared with the wave-length of light).

For a given *mass* of material, dust is highly efficient in obscuring distant light sources, gas is less efficient and chunks are highly inefficient. The presence of internebular dust would be readily detected if its mass were only a few per cent of the mass of material concentrated in the nebulæ. Therefore, dust, in the optimal form, does not contribute materially to the smoothed-out density of matter in space. Gas might be present in large quantities. Free electrons would be detected only if the density were a hundred times that of nebular material. In other forms, still larger quantities would be required. Chunks, such as dark stars and meteors, might exist in almost any numbers. Masses which were many thousands of times the total mass of the nebulæ, would produce no appreciable obscuration. Thus the density of matter in space cannot be completely determined by photometric methods alone.

Nevertheless, it is possible, from an examination of the galactic system itself, to set a reasonably definite, upper limit to the mean density of internebular space, regardless of the form of the material. The total density of both stars and interstellar material within any nebula, including the galactic system, is obviously greater than the density in outer space. Moreover, there is a density-gradient outward from the nucleus of a nebula, and the nebular material fades away to undefined borders.

Within the galactic system, the sun is situated well out from the nucleus, in an unusually dense region known as the "local system." Since the local density is high, and the general system continues to thin out for great distances toward the boundaries, it is probable that the system can be traced to densities as low as one per cent of the local density. The current value of the local density, derived largely from dynamical considerations, is of the order of $10^{-23}$, hence the value $10^{-25}$ should fairly represent the border density. This value is the extreme upper limit of the possible density in internebular space.

## *The Observable Region.*

THE observable region is thus homogeneous and isotropic—much the same everywhere and in all directions—and the smoothed-out density of matter in space is greater than $10^{-30}$ and less than $10^{-25}$. There is no observational evidence for supposing the density to be greater than $10^{-28}$.

The size of the observable region is largely a matter of definition. Within the limits of the deepest survey, $m = 21$, there should be some 60 million nebulæ, but a considerable fraction of the total number are lost in galactic obscuration. The nebulæ at the limiting magnitude represent an average distance of the general order of 400 million light-years. Some of these must be dwarf nebulæ much nearer than the average, while others are giants at still greater distances. Since dwarfs and giants cannot be distinguished from one another, distances, in those remote regions, can be used only in a statistical sense.

The figures just given refer to a systematic survey and do not represent the extreme limit of telescopic power. With long exposures under good conditions, the 100-inch reflector records nebulæ which can be distinguished from stars, as faint as magnitude 21.5. In the directions of the galactic poles, where obscuration is least, the recognizable nebulæ average about 2,400 per square degree and are more numerous than the stars. The limiting magnitude represents an average distance of the order of 500 million light-years, and about 100 million nebulæ may be expected within a sphere of this radius. Images of still fainter nebulæ, which cannot be distinguished from stars, are recorded on the photographic plates, and among them, no doubt, are extremely bright giants. Nevertheless, it is improbable that any object has been recorded whose distance is double the average distance mentioned above.

## *Effects of Red-Shifts on Apparent Luminosity.*

THE features of the observable region which have just been described follow directly from the conclusion that

# PLATE XIV

## *The Depths of Space.*

THE illustration is an enlargement from a photograph made with the 100-inch reflector of a region including the north galactic pole. The exposure was 200 min., on a special emulsion furnished by Dr. C. K. Mees of the Eastman Kodak Company. This emulsion (marked I′O) had the highest threshold sensitivity of all that have been used with the 100-inch, and, therefore, the photograph shows the faintest objects that have been recorded with telescopes now available.

The arrow points to a fair example of the faintest objects that can be recognized as nebulæ. Such nebulæ (apparent magnitude estimated as about 21.5) are at an average distance of the order of 500 million light-years.

The plate records fully as many recognizable nebulæ as stars, and the equality is a spectacular description of the penetrating power of the telescope. The plate, made on March 8, 1934, is centered about 6′ north of the star BD $+28°.2145$. The negative print is reversed so that, while east is at the top, north is to the left; 1 mm. $= 2''.35$.

the nebular distribution is approximately uniform. The final step in the analysis boldly assumes that the distribution is precisely uniform, and that apparent departures represent no more than the combined effects of red-shifts and observational errors. The effects of red-shifts are calculated on the alternative assumptions that (*a*) they represent motion (are velocity-shifts) and (*b*) they do not represent motion. Since the numerical results are not the same, the observed departures may be used to identify the correct interpretation. The differences between the two sets of calculated effects are small quantities and may be lost among the small observational errors. Nevertheless, an investigation of this dim threshold region is justified by the fundamental importance of the question at issue, although the final conclusion must be phrased with proper reservations.

Radiation from a nebula may be pictured as light—quanta-parcels of energy—streaming out in all directions. Apparent luminosity is measured by the rate at which the quanta reach the observer, together with the energy in the quanta. If either the energy or the rate of arrival is reduced, the apparent luminosity is diminished. Red-shifts reduce the energy in the quanta whether the nebulæ are stationary or receding. Thus an "energy-effect" may be expected, regardless of the interpretation of red-shifts. The rate of arrival (i.e., the number of quanta reaching the observer per second) is reduced if the nebulæ are receding from the observer, but not otherwise. This phenomenon, known as the "number-effect," should in principle furnish a crucial test of the interpretation of red-shifts as velocity-shifts.

### *The Number-Effect.*

The operation of the number-effect may be described as follows. Consider two similar nebulæ at the same distance, one stationary with respect to the observer, and the other receding with velocity *v*. Both nebulæ radiate the same number of quanta per second in the direction of the observer. At the end of a second, the quanta from the stationary nebula are scattered over a path whose distance is $c$, where $c$ is the velocity of light; the quanta from the receding nebula are scattered over the path $c + v$, which is longer than the other path by the factor $(1 + v/c)$. The density of the stream of quanta from the receding nebula is evidently less than the density of the stream from the stationary nebula. Consequently, the observer receives fewer quanta per second, and the receding nebula appears fainter than the stationary object. The reduction factor for the apparent luminosity is that given above, $(1 + v/c)$, which, for our purpose, is equivalent to $(1 + d\lambda/\lambda)$.

The number-effect is nonselective—is the same for all wave-lengths—and adds the same magnitude-increment, $\Delta m$, to magnitudes on any system—the bolometric for instance, or the photographic. The increment is

$$\Delta m \ (N.E.) = 2.5 \log \ (1 + d\lambda/\lambda)$$

where *N.E.* signifies number-effect.

### *The Energy-Effect.*

The energy-effect, which operates regardless of whether or not the nebulæ are receding, is evaluated from the fundamental relation

$$\text{energy} \times \text{wave-length} = \text{constant}$$

which holds for all quanta. Red-shifts, since they increase the wave-lengths, must decrease the energies if the product remains constant. The factor is the same as that for the number-effect, $(1 + d\lambda/\lambda)$, but the energy-effect is selective.

If the total radiation over all wave-lengths could be measured outside the earth's atmosphere, the apparent luminosity, known as the bolometric luminosity, would be reduced by the factor $(1 + d\lambda/\lambda)$. Thus the increment to the bolometric magnitude is

$$\Delta m_b \ (E.E.) = 2.5 \log (1 + d\lambda/\lambda)$$

where *E.E.* signifies energy-effect.

The effect, since it is selective, must be traced through the atmosphere (selective absorption), and through the telescope (selective reflection), up to the photographic plate (selective sensitivity), before the increment to the photographic magnitude can be evaluated. The procedure is complicated and will not be described in detail. It will be assumed, instead, that the calculations have been made and that the total effect of selection, expressed in magnitudes, is represented by $K$. Then the photographic increment is

$$\Delta m_{pg} \ (E.E.) = 2.5 \log (1 + d\lambda/\lambda) + K$$

where

$$K = \Delta m_{pg} - \Delta m_b$$

and $K$ varies with the red-shifts.

The evaluation of $K$ depends to a certain extent on properties of the initial unshifted radiation which cannot be directly observed. The properties must be inferred, and the major uncertainties in the calculated effects of red-shifts arise from this necessity. The plausible assumption that nebulæ radiate like stars whose effective temperatures are about 6,000°, slightly higher than that of the sun, leads to values of $K$ which are probably of the right order. The values are considerably larger than the increment representing the number-effect alone. Therefore, the uncertainties, although relatively small fractions of $K$, may be rather large compared to the number-effect, which is the object of investigation.

The significant points in this rather long discussion may be stated with commendable brevity. Effects of red-shifts on photographic magnitudes are

$$\Delta m \ (\text{cal.}) = 5 \log (1 + d\lambda/\lambda) + K$$

or

$$2.5 \log (1 + d\lambda/\lambda) + K$$

according to whether the red-shifts do, or do not, represent motion. For effective nebular temperatures of 6,000°, the calculated values of $\Delta m$, are closely proportional to the red-shifts, $d\lambda/\lambda$, and are represented by the relations[6]

$$\begin{aligned}\Delta m \text{ (cal.)} &= 4\, d\lambda/\lambda \text{ (motion)}\\ &= 3\, d\lambda/\lambda \text{ (no motion).}\end{aligned}$$

### *Effects of Red-Shifts and Observed Departures from Uniformity.*

THESE simple relations may now be compared with the apparent departures from uniformity observed in the surveys. The departures increase with distance, and the relation is approximately linear. On the assumption that the relation is really linear, the apparent distribution is expressed by

$$\log \overline{N}_m = 0.6\ (m - \Delta m) + \text{constant}_1$$

where

$$\log \Delta m = 0.2\ (m - \Delta m) + \text{constant}_2$$

and the constants can be derived from the observed data by the usual methods of least squares. The solution is shown graphically in Fig. 16 by the smooth curve through the observed points. The fidelity with which the observations are represented by the curve supports the validity of the assumption that $\Delta m$ is a linear function of distance.

Now red-shifts, $d\lambda/\lambda$, are also linear functions of dis-

[6] Various methods of calculating the increments are discussed by Hubble and Tolman, "Mt. Wilson Contr.," No. 527; *Astrophysical Journal*, 82, 302, 1935, although the numerical results are not given. de Sitter (*Bulletin of the Astronomical Institutes of the Netherlands*, 261, 1934) gives detailed calculations, using one method, for the increment here termed the "energy-effect." The results are expressed in the form

$$\Delta m_{pg} = 2.9\, d\lambda/\lambda + (d\lambda/\lambda)^2$$

which closely approximates that given above, over the range of the surveys, namely, $d\lambda/\lambda < 0.25$.

The addition of the number-effect, 2.5 log $(1 + d\lambda/\lambda)$, leads to the series

$$\Delta m_{pg} = 3.99\, d\lambda/\lambda + 0.48\, (d\lambda/\lambda)^2 + \ldots$$

which also approximates the relation given above for red-shifts interpreted as velocity shifts.

tance[7] (Chap. V). Therefore $\Delta m$ is a linear function of $d\lambda/\lambda$. The relation,[8] as derived from observations is about

$$\Delta m \text{ (obs.)} = 2.7\, d\lambda/\lambda.$$

The observed coefficient is smaller here than that in the relation calculated on either interpretation of red-shifts, but is much closer to the coefficient representing no motion. Careful examination of possible sources of uncertainties suggests that the observations can probably be accounted for if red-shifts are not velocity-shifts. If red-shifts are velocity-shifts then some vital factors must have been neglected in the investigation.

A review of the problem discloses at least one neglected factor, namely, the difference in time required for light to reach the observer from the limits of the various surveys. As we look out into space, we look back into time. The surveys were made recently, but the light left the twenty-first magnitude nebulæ perhaps 120 million years before it passed the twentieth magnitude nebulæ, and 250 million years before it reached the 18.5 magnitude nebulæ. During these immense periods, the nebulæ, if red-shifts were velocity-shifts, would have receded to appreciably greater distances than those estimated from the apparent faintness. Thus the observed distribution would require corrections to reduce it to a "simultaneous" description.

Attempts to determine the corrections raise questions concerning the measurement of distances and their interpretation, and ultimately force the research into the field of relativistic cosmology.

[7] The function is observed to $d\lambda/\lambda = 0.14$, corresponding to $m = 19.5$ for isolated nebulæ, and to this limit it is sensibly linear. Beyond this limit the function must be extrapolated, but unless abrupt departures occur (an unlikely contingency) the assumption of linearity should approximate the general order of the red-shifts.

[8] Since log $d\lambda/\lambda$ and log $\Delta m$ are both represented by the expression $0.2\,(m - \Delta m)$ + constant (the constant, of course, differing in the two cases), $\Delta m$ is directly proportional to $d\lambda/\lambda$, and the factor is the anti-log of the difference between the constants.

## *Theories of Cosmology.*

CURRENT theories of cosmology employ a model known as the homogeneous, expanding universe of general relativity or, more briefly, as the expanding universe. It is derived from the cosmological equation which expresses a principle of general relativity—that the geometry of space is determined by the contents of space. The equation transcends the body of factual knowledge and can be interpreted and solved only with the aid of assumptions concerning the nature of the universe.

The first solutions, by Einstein and de Sitter (1917), employed the assumptions that the universe is homogeneous and isotropic and also that it is static, i.e., does not vary systematically with time. These solutions were special cases of the general problem and have since been abandoned—Einstein's, because it did not account for red-shifts; de Sitter's, because it neglected the existence of matter. The Einstein universe, it was said, contained matter and no motion, while the de Sitter universe contained motion and no matter. The general problem was first discussed by Friedmann (1922). Subsequently, Robertson (1929) derived the most general formulation (of the line element) from properties of symmetry alone.[9]

The solution involved the "cosmological constant" and the "radius of curvature of space" as undetermined quantities. By arbitrarily assigning different values to the parameters, various classes of possible universes were described, and among them, it was supposed, the type corresponding to the actual universe would be included. The problem for the observer was to determine

[9] For further information on this great field of theoretical investigation, the reader is referred to Robertson's authoritative review of the development of the subject up to the end of 1932; "Relativistic Cosmology," *Reviews of Modern Physics*, 5, 1, 1933. A complete bibliography of the more important contributions, with short descriptions of their contents, is included, and also a list of recommended, nontechnical discussions of the field. Among the latter is an exceptionally clear statement from the mathematical point of view, by Robertson himself ("The Expanding Universe," *Science*, 76, 221, 1932).

the actual values of the constants, or at least to narrow the range within which they must lie.

The general solution was nonstatic, and the radius of curvature of space varied with time. Therefore, the possible universes were contracting or expanding. The equations did not indicate which alternative to expect, but the observed red-shifts were generally accepted as direct evidence that the actual universe was expanding at the present time, and this interpretation was incorporated in the theory. Thus the model came to be known as the homogeneous, expanding model of general relativity.

The cosmological problem is of wide interest and discussions have not been confined entirely to the field of general relativity. Milne, in particular, has developed a "Kinematical" model which appears to possess unusually significant features.[10] For our present purpose, however, it does not require special consideration since it has been shown to correspond very closely to a particular case of the general relativity model—namely the hyperbolic model with negative curvature.

For any theory of the structure of the universe it is possible to make a "world-map," to use Milne's phrase, which indicates the actual distribution of nebulæ at a given epoch of time. The apparent distribution, which an observer should expect to record on his photographs (if the theory corresponds to the actual universe), has been called a "world-picture," another of Milne's phrases. World-pictures must differ from world-maps, if red-shifts are velocity-shifts, since the nebulæ continue to recede while light travels to the observer. The theories can be tested by comparing the observed distribution with the calculated world-pictures.

Certain features of the world-picture for the general relativity model have been calculated by Tolman.[11]

[10] *Relativity, Gravitation and World Structure,* 1935.

[11] *Relativity Thermodynamics and Cosmology* (1934), chap. x. The application to the immediate problem of nebular surveys is discussed in Hubble and Tolman, *Two Methods of Investigating the Nature of the Nebular Red-Shifts,* "Mt. Wilson Contr.," No. 527; *Astrophysical Journal,* 82, 302, 1935.

Among them is the formula expressing the relative numbers of nebulæ that should be observed at a given epoch, within various limits of apparent magnitude. From this relation the effects of red-shifts (interpreted as velocity-shifts) on the nebular counts, are readily derived. It then appears that the effects of red-shifts in the world-picture, are precisely those discussed in the previous section, with the addition (to the former) of a term in *R*, the radius of curvature of space.

Curvature was neglected in the previous discussion, and the discrepancies which were found when red-shifts were interpreted as velocity-shifts, might conceivably be explained by the neglected factor. It will be recalled that it was just possible to explain the nebular counts on the assumption that red-shifts were not velocity-shifts. If red-shifts were velocity-shifts, additional corrections, called "number-effects," were required, and these appeared as discrepancies. The question now arises whether sufficient curvature can be introduced to balance the number-effects and thus remove the apparent discrepancies.

Tolman's formula indicates that a positive value of *R* would reduce the discrepancies, while a negative value would increase them. Thus a negative curvature, implying an open universe, is ruled out, and the possible, expanding universes are restricted to those with positive curvature. If red-shifts are velocity-shifts, it follows that the universe is closed, having a finite volume and finite contents.

The curvature required to remove the discrepancies is very great, and hence the radius of curvature, *R*, is very small. Actually, it is comparable with the radius of the observable region as defined with existing telescopes. Thus, in order to save the velocity-shifts, we would be forced to conclude that the universe itself is so small that we are now observing a large fraction of the whole.

Some further information may be found in the fact that the radius, *R*, in a closed universe, bears a definite relation to the density of matter (and radiation) in space. A radius of the dimensions necessary to save the velocity-shifts represents a mean density appreciably higher than $10^{-26}$ gm/cc. This value is many times greater than even the maximum estimates of the smoothed-out density of material concentrated in nebulæ, and we find no evidence of any considerable amount of internebular material which might increase the density. A sufficient amount of the latter material might conceivably exist provided it were in forms that cannot be detected, but an upper limit can be set even to matter in these forms. The density along the borders of the galactic system, as previously mentioned, is probably not greater than $10^{-25}$ gm/cc., and the density in the surrounding space is presumably still less. Radiation would not change the general order of the density.

If the estimations of density were completely reliable, a radius of curvature of the necessary dimensions would be ruled out by the evidence. But so definite a solution is probably unwarranted. The crucial data are surrounded by uncertainties. By pressing the data to the limit of their tolerance, always in one direction, we might force the velocity-shifts into the framework of the surveys. The universe would then be small, and filled with matter to the very threshold of perception.

On the other hand, if the interpretation as velocity-shifts is abandoned, we find in the red-shifts a hitherto unrecognized principle whose implications are unknown.[12] The expanding universe of general relativity would still persist in theory, but the rate of expansion would not be indicated by the observations.

THUS the explorations of space end on a note of uncertainty. And necessarily so. We are, by definition, in the very center of the observable region. We know our imme-

[12] A new method of approach has already been developed by Zwicky, in a paper, the main purpose of which "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," "Remarks on the Redshift from Nebulæ," *Physical Review*, 48, 802, 1935.

diate neighborhood rather intimately. With increasing distance, our knowledge fades, and fades rapidly. Eventually, we reach the dim boundary—the utmost limits of our telescopes. There, we measure shadows, and we search among ghostly errors of measurement for landmarks that are scarcely more substantial.

The search will continue. Not until the empirical resources are exhausted, need we pass on to the dreamy realms of speculation.

# INDEX

## SILLIMAN MEMORIAL LECTURES
## PUBLISHED BY YALE UNIVERSITY PRESS

LIAM E. FORD, R. B. SOSMAN, HORACE L. WELLS, HARRY W. FOOTE, LEIGH PAGE, WESLEY R. COE, and GEORGE L. GOODALE.

A TREATISE ON THE TRANSFORMATION OF THE INTESTINAL FLORA WITH SPECIAL REFERENCE TO THE IMPLANTATION OF BACILLUS ACIDOPHILUS. By LEO F. RETTGER, Professor of Bacteriology, Yale University, and HARRY A. CHEPLIN, Seessel Fellow in Bacteriology, Yale University. (Out of print.)

THE EVOLUTION OF MODERN MEDICINE. By SIR WILLIAM OSLER, BART., M.D., F.R.S. (Fourth printing.)

RESPIRATION. By J. S. HALDANE, M.A., M.D., F.R.S., HON. LL.D. BIRM. AND EDIN., Fellow of New College, Oxford; Honorary Professor, Birmingham University. (Second printing.)

AFTER LIFE IN ROMAN PAGANISM. By FRANZ CUMONT. (Second edition.)

THE ANATOMY AND PHYSIOLOGY OF CAPILLARIES. By AUGUST KROGH, PH.D., LL.D., Professor of Zoö-physiology, Copenhagen University. (Second printing, Second edition.)

LECTURES ON CAUCHY'S PROBLEM IN LINEAR PARTIAL DIFFERENTIAL EQUATIONS. By JACQUES HADAMARD, LL.D., Member of the French Academy of Sciences; Foreign Honorary Member of the American Academy of Arts and Sciences.

THE THEORY OF THE GENE. By THOMAS HUNT MORGAN, LL.D., SC.D., PH.D., Professor of Experimental Zoölogy in Columbia University. (Third printing, Second edition.)

THE ANATOMY OF SCIENCE. By GILBERT N. LEWIS, PH.D., SC.D., Professor of Chemistry and Dean of the College of Chemistry, University of California. (Second printing.)

BLOOD. A Study in General Physiology. By LAWRENCE J. HENDERSON, A.B., M.D., Professor of Biological Chemistry in Harvard University.

ON THE MECHANISM OF OXIDATION. By HEINRICH WIELAND, Professor of Organic Chemistry, University of Munich. (Second printing.)

MOLECULAR HYDROGEN AND ITS SPECTRUM. By OWEN WILLANS RICHARDSON, Yarrow Research Professor of the Royal Society, King's College, London.

THE CHANGING WORLD OF THE ICE AGE. By REGINALD ALDWORTH DALY, Sturgis Hooper Professor of Geology, Harvard University. (Second printing.)

THE REALM OF THE NEBULÆ. By EDWIN HUBBLE.