

UNIVERSAL
LIBRARY

OU_152338

UNIVERSAL
LIBRARY

OSMANIA UNIVERSITY LIBRARY

Call No. 523.1 S62K

Accession No. 15783

Author Siller, W. D.

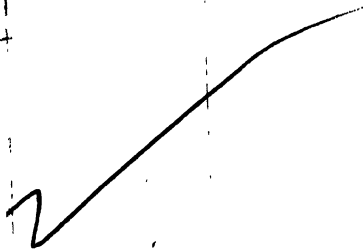
Title KOSMOS

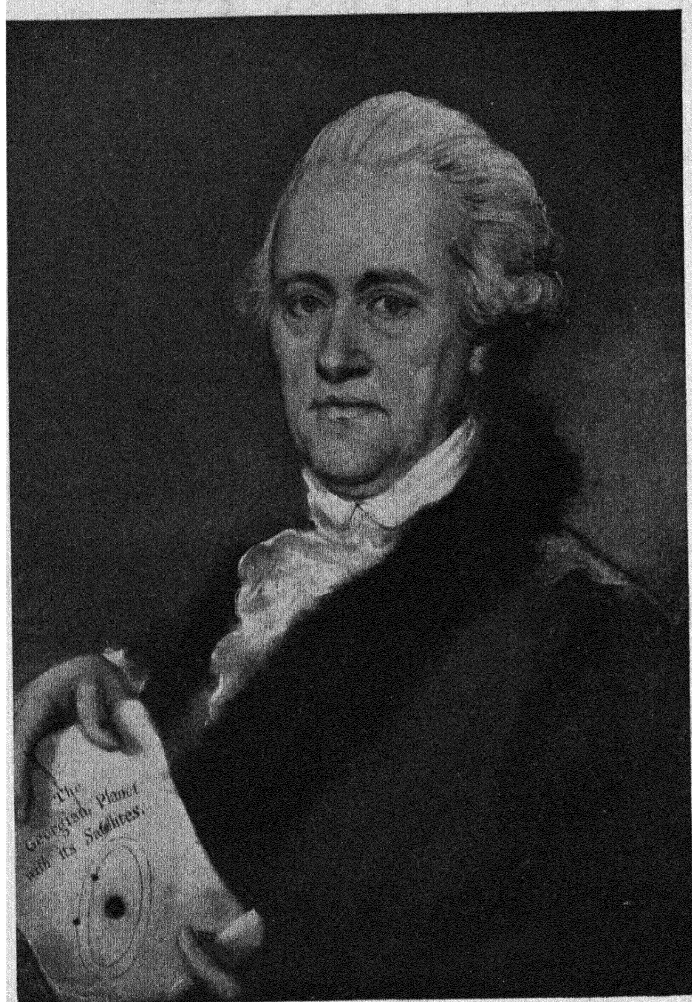
This book should be returned on or before the date last marked below.

11 NOV 1955

23 NOV 1962

M





SIR WILLIAM HERSCHEL

KOSMOS

*A Course of Six Lectures on the
Development of our Insight into the Structure of the Universe,
delivered for the Lowell Institute in Boston, in November 1931*

BY

W. DE SITTER

*Director of the Observatory and Professor of Astronomy
at the University of Leiden*



Cambridge, Massachusetts

HARVARD UNIVERSITY PRESS

1932

COPYRIGHT, 1932¹
BY THE PRESIDENT AND FELLOWS OF HARVARD COLLEGE

PRINTED AT THE HARVARD UNIVERSITY PRESS
CAMBRIDGE, MASS., U. S. A.

To

MY WIFE

PREFACE

THE Lowell lectures, held on the invitation of the Lowell Institute in Boston in the autumn of 1931, had as their subject the evolution of the ideas of mankind regarding the structure of the universe. They were not, and could not possibly be, a complete history of the subject. Some salient points were selected, and treated more or less fully, whilst long periods were entirely skipped. My aim was not so much historical accuracy in details — though of course I have endeavoured to be accurate — as the presentation of the leading principles and the evolution of the general outlook. And especially I hope to have made clear the fundamental elements common to the attitudes at different epochs, which are often apparently so different, and to have shown how at all times the efforts of great men were directed towards the discovery of the great principle of unity, which I cannot express better than by the Greek word chosen as the title of the book.

Originally the course was to have consisted of eight lectures, which had, by the force of circumstances, to be reduced to six. Some traces of the effects of this reduction may still be visible at the end of the second and the beginning of the fourth lecture. The substance of the lectures is essentially as they were delivered, though the actual wording will probably be very different, as I did not refer to a written text while delivering them. The last two lectures have been brought up to date. Especially the last needed considerable alterations and additions. The conclusions of these two

chapters may now be considered to represent the position of science, or rather my personal appreciation of that position, in March 1932.

In the preparation of the lectures I owe much to my collaborators at the Leiden Observatory, especially to Dr. Oort, and in the revision of the last two chapters I have been materially assisted by the great benefit I have derived from conversations with many astronomers and physicists during my delightful stay in the United States and Canada, especially at Cambridge and Pasadena.

It is a pleasure, as well as a duty, to thank the Harvard University Press for the great care devoted to the publication, and the Mount Wilson Observatory for permission to reproduce several of its photographs.

LEIDEN, April, 1932 .

CONTENTS

I. INTRODUCTION: AIMS AND METHODS OF SCIENTIFIC RE- SEARCH • THE ANCIENTS; PTOLEMY AND COPERNICUS . .	3
II. BIRTH OF MODERN ASTRONOMY • KEPLER, GALILEO, NEW- TON • THE EIGHTEENTH CENTURY	19
III. THE BEGINNING OF SIDEREAL ASTRONOMY • WILLIAM HER- SCHEL AND HIS SUCCESSORS	34
IV. FURTHER OBSERVATIONAL ADVANCES • THE SURVEY OF THE “LOCAL SYSTEM” • LIFE-WORK OF J. C. KAPTEYN . . .	52
V. PHOTOGRAPHY AND SPECTROSCOPY • THE GREAT TELESCOPES • THE GALACTIC AND THE EXTRA-GALACTIC SYSTEMS . .	78
VI. RELATIVITY AND MODERN THEORIES OF THE UNIVERSE • CONCLUSION	103

ILLUSTRATIONS

PORTRAITS

I. SIR WILLIAM HERSCHEL	<i>Frontispiece</i>
II. NICOLAUS COPERNICUS	18
III. TYCHO BRAHE	20
IV. JOHANNES KEPLER	22
V. GALILEO GALILEI	26
VI. SIR ISAAC NEWTON	28
VII. PROFESSOR J. C. KAPTEYN	52

FULL-PAGE PLATES

VIII. THE SPIRAL NEBULA MESSIER 101	88
IX. SOUTH-PRECEDING PORTION OF THE GREAT SPIRAL NEBULA IN ANDROMEDA	90
X. STAR CLOUDS IN SAGITTARIUS	90
XI. THE MILKY WAY IN OPHIUCHUS	92
XII. THE SPIRAL NEBULA N. G. C. 4565	94

FIGURES IN THE TEXT

1. THE SYSTEM OF EXCENTRIC CIRCLES AND THE SYSTEM OF EPICYCLES	16
2. THE PTOLEMAIC AND THE COPERNICAN SYSTEMS	17
3. SECTION OF THE GALACTIC SYSTEM ACCORDING TO SIR WILLIAM HERSCHEL'S FIRST THEORY	40
4. WILLIAM HERSCHEL'S DIAGRAM ILLUSTRATING HIS SECOND THEORY (1817)	45
5. THE KAPTEYN SYSTEM (1922)	73

6. THE GLOBULAR CLUSTERS AND THE MILKY WAY ACCORDING TO H. SHAPLEY (1918)	84
7. DR. EASTON'S DIAGRAM ILLUSTRATING THE SPIRAL STRUCTURE OF THE GALACTIC SYSTEM	88
8. DIAGRAM ILLUSTRATING THE ROTATION OF THE GALACTIC SYSTEM	92
9. DETAIL OF FIGURE 8 ON ENLARGED SCALE	93
10. DR. OORT'S DIAGRAM OF THE GALACTIC SYSTEM	94
11. THE THREE FAMILIES OF NON-STATIC UNIVERSES	125

K O S M O S

I

INTRODUCTION : AIMS AND METHODS OF SCIENTIFIC RESEARCH · THE ANCIENTS; PTOLEMY AND COPERNICUS

WE ARE living in a time of most remarkable development of scientific ideas. Never in all the history of science has there been a period when new theories and hypotheses arose, flourished, and were abandoned in so quick succession as in the last fifteen or twenty years.

Some of you will remember the publication in 1914 of Eddington's book: "On Stellar Movements and the Structure of the Universe." This book contained the most recent results in a domain on which during the last ten years the interest of astronomers had been strongly concentrated. It was a review of the position of science after the remarkably fertile period of development which had followed after Kapteyn's discovery of the star streams in 1904, and was considered by all astronomers, including its author, as marking a resting place on the road, where the data gathered in different ways were collected, discussed, and arranged, in order to make new plans for the continuation of the search in the same direction. Looking back on it now, it impresses us as the close of a former period. It is true that investigations of the same nature have been continued; the plans that were drawn up have been executed and have yielded the important results that were hoped for, but they have not, as

was expected, dominated the evolution of the science, nor indicated the chief direction of development.

On the contrary, it looks as if since that time we could not speak any more of a chief direction, of one line in the progress of science. New branches of astronomy have sprung up, apparently independent of each other, and often leading to contradictory results.

In the same year, 1914, Einstein's "Entwurf einer verallgemeinerten Relativitätstheorie und Theorie der Gravitation" was published. I need not call back to your mind how many controversies, not always conducted in a dignified manner, this theory has elicited. In November 1915 the theory was completed and began its triumphal march through the world, gaining its final victory as early as 1919 by the results of the English eclipse expeditions. It seemed at first as if here at least a safe harbour had been reached, where we could rest and quietly sort out the treasures that the venture-some voyage into new regions had brought us. But presently the further development went in unexpected directions, and brought most strange and paradoxical results.

In the domain of atomic physics the confusion has been still worse. About the same time, now nearly twenty years ago, Bohr's theory was developed. The quantum theory assumed a new aspect and soon rose to high esteem. But this glory did not last long. The theory of fifteen years ago is now an historical curiosity, known under the name of the "classical quantum theory." It has been successively replaced by the "new quantum theory," the theory of matrices, wave mechanics, each stranger and more paradoxical than its predecessor. What used to be the most fundamental con-

cepts of physical science, determinism and causality, are called into doubt, the foundations of science appear to be shaking, and it seems as if the whole building were tottering.

Is then the unity of science lost? Has the tremendous exertion of the last twenty years only resulted in abolishing our formerly so well-ordered kosmos, giving us chaos instead?

It would be a mistake to think so. To begin with, without a solid faith in the existence of order and law no science is possible. And modern science, however revolutionary and undisciplined it may appear to be, is certainly science. It is trying to do what science has always done, what is its most important, if not its sole legitimate, function, viz. to correlate different phenomena. The ultimate aim of science is to discover as much as is discoverable about the reality behind the phenomena. Of the fundamental assumption that there *is* a reality behind the phenomena, existing independent of the fact whether and how it is observed, physical science has as little doubt as simple common sense. Strictly speaking every assertion about what has not been observed is outside physics and belongs to metaphysics. And although there is nothing an orthodox physicist abhors more than metaphysics, still he has no hesitation at all in believing that his laws are universal and that the phenomena continue to happen according to them just the same when nobody is looking. And, since it would be impossible to prove that they do not, he is fully entitled to his belief. The observed phenomena are the interaction between this reality, of which we assume the existence, and our senses or our instruments, which are only extended and refined sense-organs. Thus they certainly contain something of that reality, and that part of it which they

contain is all we ever can learn about it. But they also contain much that is due to ourselves, and the difficulty lies in separating those different elements. The laws of nature are the expression of the regularities that are found in the phenomena and the correlations between the different phenomena, and it is these regularities and correlations that we ascribe to reality and not to the observer. The laws of nature, after all that is relative, due to the observer, has been eliminated, *are* for us the reality.

In different ages the appreciation of the meaning and the aims of science may have been different, and in some epochs more emphasis may have been laid on aspects which were at other times considered to be of subordinate interest. But this difference in attitude is more apparent than real; the ultimate criteria by which scientific theories are judged are invariably the same.

Every physical theory must begin and end in observation. Its origin is the attempt to account for observed phenomena on a basis of reason, and consequently the final test necessarily is comparison with observations; no theory can survive which is not able successfully to stand this test. This truth has sometimes been lost sight of. Thus, during that period of the middle ages when there was any science at all, the works of Aristotle enjoyed an undisputed authority, and in order to decide whether a theory was true or not, it was compared not with observations but with Aristotle. The Greek philosophers themselves, however, certainly were well aware of the fact that science must be based on observation. Aristotle himself, Archimedes, Hipparchus, Ptolemy, must,

without any doubt, be counted as experimental physicists and astronomers. Even in the much abused middle ages the critical attitude of mind is not entirely absent. The derogatory remarks of King Alfonso X regarding the Ptolemaic system of the world are well known, but it is only in the sixteenth and seventeenth centuries that the necessity of making observations is again felt very strongly. Perhaps never in all history has the requirement of testing every theory by observations been so strongly emphasised as by the two great men whose work inaugurated the new epoch of science, Kepler and Galileo.

During the whole development of science there is thus a continual interaction between theory and observations. New observations give birth to new theories and hypotheses, and, on the other hand, the development and generalisation of existing theories point out where new observations are required.

There are two methods by which science proceeds in the building up of her theories. The one is by generalisation and induction, the other by hypothesis. Both are legitimate tools of science, and both have been used at all times. In some periods, however, the one method has been in higher esteem and in others the other.

The best examples of pure generalisation or abstraction are found in the development of the science of mechanics, and especially in the work of Sir Isaac Newton, whose outspoken abhorrence of hypotheses is well known. By his influence and the great authority of his name the purely inductive method came to be considered as the only correct one amongst his followers, and Laplace even finds it neces-

sary to make a special apology when he indulges in an hypothesis.

On the other hand the atomic theory and theories of radiation are typical examples of successful application of the method of hypotheses from the beginning, and especially in the last quarter of a century the hypotheses have been growing very luxuriantly in this corner of the garden of science.

But whether the crop raised are generalisations or hypotheses, the seed is always derived from observation, and the fruit must be confirmation by and correlation of more observations.

The direction in which the development of science proceeds in any particular epoch depends, of course, greatly on the general tendency of prevalent philosophical ideas. An example of contrasting points of view which can be followed throughout the centuries is the antithesis between a statical and a dynamical view of the universe. At the very beginning of Greek philosophy, about the year 500 B.C., we meet with the conflicting systems of Heraclitus and the Eleatic philosophers, the first emphasising the ever-changing aspect of nature, while the others considered existence the fundamental fact and change to be more or less apparent. The Greek mind was essentially static; Heraclitus is rather exceptional. The whole development of Greek mathematics, culminating in Euclid and Apollonius, is in the spirit of the Eleatic school. It considers only discrete conceptions; the concept of motion and continuity, of the infinitely small and the infinitely large, is entirely absent until the time of Archimedes, who must be considered as the first modern mathematician, as he also was the first theoretical physicist. Archimedes, however, was an

exceptional genius, whose ideas were a long way ahead of his time, and have had little influence on the development of Greek science in his own time and in the centuries immediately following. It is only when Galileo takes them up again that they bear fruit and are further developed.

Also in the later development of science, and in our own time, we find sometimes more stress laid on the evolution in nature and at other times more on the permanency of the laws. About the middle of the last century the evolutionary idea received a great impetus from the work of Charles Darwin. It is not entirely accidental that our present views on the evolution of the planetary system and the universe are intimately connected with the name of Sir George Darwin. There are many characteristics in the work of the son which remind one very forcibly of that of the father.

During long periods in history the development of science appears to be slow, whilst at some epochs a sudden revival springs up. These sudden developments of new ideas always come after a great improvement in the technical means of observation. In the history of astronomy there are three of these epochs of accelerated progress.

The first, which may be characterised by the name of Hipparchus, is so long ago that the reconstruction of the circumstances cannot but be largely hypothetical. But it is certainly connected with the general technical progress during the century of expansion of Greek civilisation after the conquests of Alexander the Great.

The second, in the seventeenth century, comes after the invention of the telescope and reached its climax in Newton's work.

The third is inaugurated in the last quarter of the nineteenth century by the discovery of photography and spectroscopy. We are still in the midst of it, and it does not yet show any signs of abating its activity, or of having reached the stage where the collecting of new results is succeeded by the ordering of them into an all-embracing system.

A general feature of the development of science from its first almost unconscious beginnings to the very elaborate structure of today is the more and more artificial aspect that the theories present to the superficial critic. The generalisations become wider and wider; general principles, like that of conservation of energy, and others of equally wide range, arise beside, and partly replace, the special laws. New concepts, utterly different from those based on direct sense impressions, are introduced by hypotheses. The picture that science draws of the reality behind the phenomena is very unlike the directly observed phenomena themselves.

This is unavoidable. The essential unity of nature is hidden deep below the surface, which is all that direct observation can show us. The belief in the existence of this unity, of course, lies, and has always lain, at the base of all science, the conviction that there is order and law in nature, that the whole outer world, the universe, is a well-ordered system, a unit. This is not a scientific thesis, or dogma, it is pre-scientific, being rooted much deeper in our consciousness than science is, it is what makes science possible. When Kosmos was born out of Chaos, the soil was prepared on which science could grow up.

The origin of Greek science is lost in the hazy distance of prehistoric times. The prediction of the solar eclipse of 585 B.C. by Thales is certainly only legendary. The beginning of Greek science may perhaps be dated from the school of Pythagoras (570 to 490 B.C.), though it is difficult to separate the truth from the fables surrounding the mythical person of its founder. It seems certain, however, that Pythagoras had a great influence on the whole spirit of Greek science. And particularly the form of early planetary theories has been largely determined by the strong preference given in Pythagorean philosophy to the sphere and the circle as the most perfect body and curve. The development of the planetary motion in series of uniform motions in circles, which is still used every day in its algebraical form of trigonometric series, may be traced to this origin. It seems probable that Pythagoras is the first who had the courage to think of the earth as freely floating in space in the centre of the universe, without needing any material means of support.

Pythagoras is also credited by some, though on entirely insufficient grounds, with what is certainly one of the most remarkable discoveries in astronomy, viz. the identity of the morning and the evening star. When this discovery was made, is, however, entirely uncertain. In the Greek cosmology of the time of Homer, as well as in early Babylonian inscriptions, they are still two distinct stars. It seems probable that the identity was only realised by Greek science about, or shortly after, the time of Pythagoras. This discovery may be compared with the similar discovery of the identity of a comet before and after perihelium passage. This last, of course, was a direct consequence of Newton's theory of gravi-

tation, according to which the comet's orbit must be a conic section, and especially a parabola, instead of a straight line as had been formerly assumed. Similarly the identification of the two appearances of the inner planets as morning and evening star may perhaps be connected with the theory of circular motion of the planets.

At about the same time, round 500 B.C., the first attempts at general theories of all physical phenomena appear, although to our eyes in very crude form. The atomic theory of Democritus dates from this same epoch, as well as the system of Heraclitus, who explained everything by motion.

These theories were, however, still mostly speculative; the origin of empirical, observational science comes later.

The first example of the mutual influence of theory and observation is the system of homocentric spheres, devised by Eudoxus (408 to 355 B.C.) and modified thirty years later by his pupil Kalippus, a contemporary of Aristotle, so as to make it agree with the observed facts within the limits of accuracy attainable by the observations of that time.

Soon afterwards the system was abandoned altogether, and replaced by the theory of excentric circles, chiefly to account for the great changes of brightness of the planets in the different parts of their orbits.

Aristotle (384 to 322 B.C.), who for the later middle ages was the great authority, lived in a time when Greek science was still in its beginnings. About a century later we find a stage of much farther development. Of the great names of that time we shall mention only three: Eratosthenes, Aristarchus, and Archimedes.

Eratosthenes, about 200 B.C., measured the diameter of

the earth by determining the difference of latitude between Syene in Egypt and Alexandria. His result is only about 1% in error.

Aristarchus was the first to determine the distance of the sun from the earth by actual measurement. He also determined the apparent diameter of the sun, which he makes equal to $1/720$ of the circumference of the ecliptic. Aristarchus then finds the distance of the outermost planet of the solar system, Saturn, by the proportion $b:a = a:R$, b being the radius of the earth, a the distance of the earth from the sun, and R the radius of the sphere containing the whole solar system. The remarkable convincing power of analogies like this, which to our mind appear to be absolutely baseless, has survived during a great many centuries; the same reasoning is used, not only by Archimedes, but even by Kepler.

Aristarchus also was the author of the heliocentric system of the planetary motions. As Dreyer has shown, this system was probably developed from that of the excentric circles, by bringing all centres to co-incidence in the sun, by which it was practically transformed to Tycho's system. Having gone thus far, Aristarchus at once took the further step — from Tycho's system to that of Copernicus — by making the earth revolve round the sun instead of the sun with all the planets round the earth.

Aristarchus realised that his system involved enormous distances for the fixed stars, and he says that the sphere of the earth is to that of the stars, as the centre of a sphere is to the surface. Archimedes, as the conscientious mathematician that he was, objects to this, since the centre, having no dimensions, has no ratio to the surface. This mathematical

quibble is, however, not used as an argument against the system. It does not appear whether Archimedes adopted or rejected the heliocentric system.

Archimedes (289 to 212 B.C.), the greatest of all, is the founder of theoretical physics, the first who made experiments with the special purpose of subjecting them to mathematical reasoning. He laid the foundations of the science of mechanics by discovering the laws of the centre of gravity, of the lever, and of floating bodies. His mathematical methods contain the germ of the infinitesimal calculus; he is the first to define a curve by the *motion* of a point instead of by statical properties, and to introduce the concepts of continuity, and of the infinitely large and of the infinitely small. In his remarkable little treatise, called *ψαμμίτης*, or in the Latin translation "Arenarius," he proves that the number of grains of sand which would be sufficient to fill up the whole universe, though extremely large, is not mathematically infinite.

About a century again after the epoch of Aristarchus and Archimedes comes Hipparchus (190 to 125 B.C.), who has been rightly called the father of astronomy. The problems first tackled by him are the great problems of fundamental astronomy, which still continue to occupy us: the determination of the equinox, the determination of the length of the year, the elements of the orbits of the sun and the moon. He also made the first catalogue of stars, and is the discoverer of precession. The value of the precessional constant generally quoted as determined by Hipparchus, viz. 36" a year, is not the one actually derived by him. He determined the difference in longitude between the star Spica and the moon during

a lunar eclipse, and found a difference of two degrees in 150 years, i. e. $48''$ per annum, only 4% different from the true value.

The discoveries of Hipparchus were made possible by the improvement of the instruments, especially by the introduction of diopters and divided circles.

The knowledge of astronomy of the ancients has been crystallized in the work of Ptolemy, about A.D. 130, the "Almagest" or "Great Array." This work contains the well-known theory of the planetary system; the motions of the planets are analysed in uniform motions along circles, the centres of which in their turn move uniformly on other circles, etc. The earth is the centre of the universe. The accuracy of the observations had been so much increased that one circle for each planet, as in the system of excentrics and in the heliocentric system of Aristarchus, was no longer sufficient, but several epicycles were required for each planet.

The relation between the two systems may be seen from Fig. 1, where the full lines represent the Ptolemaic system of epicycles, and the dotted lines the system of excentric circles. The earth is at A , the planet at P . In the older system the planet moves on the "excentric circle," and the centre C of this circle describes the "concentric circle." In the later system the planet moves on the "epicycle," of which the centre D is carried on the "deferent." The sun, in both systems, moves in a circle round the earth. Its position at the moment is represented in the diagram by the dot with a small circle round it. In the system of excentrics the centres of all the excentric circles are always on the line joining the sun to the earth; in the Ptolemaic system the radius DP is always paral-

lel to the direction from the earth to the sun. It will be seen that the two systems give the same representation of the observed relative positions.

In the system of excentric circles the simplification of putting the sun (of which we know only that it is on the line AC , but not where it is on that line) in C , and having only *one* concentric circle (viz. the orbit of the sun) for all planets,

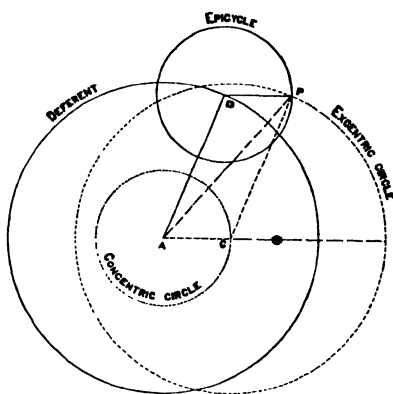


FIGURE 1

The system of excentric circles and
the system of epicycles

is, of course, rather obvious. This leads to the Tychonic system. The further step to the Copernican system is not wide. It consists only in keeping C fixed as the centre of the world, and making A revolve round C , instead of C round A .

Figure 2 shows directly the equivalence of the Ptolemaic system, represented in the left-hand half of it, and the Copernican system, represented in the right-hand half. A is the earth, Z the sun, and P the planet. The figure shows very clearly the superior simplicity of the Copernican system.

This gain in simplicity, however, would not be so striking if we had not reduced the system to its simplest form, by leaving out all the secondary epicycles that Hipparchus and Ptolemy had added to the system, in order to make it represent the observations.

The difference between the system of Ptolemy and that of Copernicus is a purely formal one, a difference of interpreta-

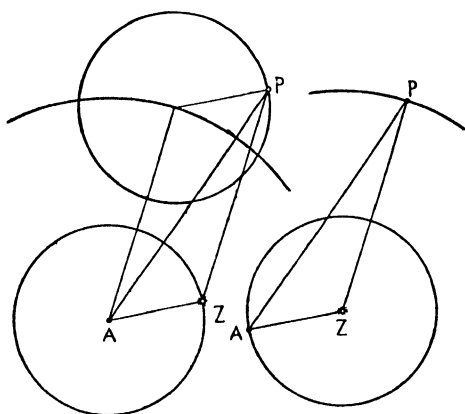


FIGURE 2

The Ptolemaic and the Copernican systems

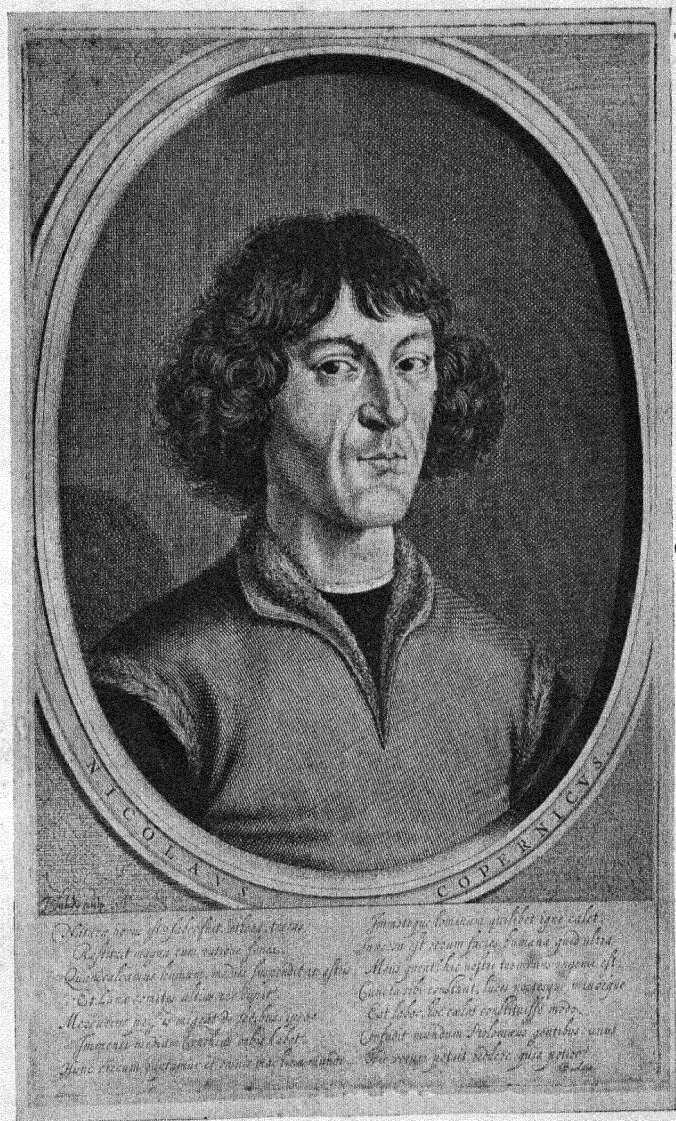
tion only, not of the mathematical contents of the theory. That the heliocentric system was never mentioned in Greek science after Aristarchus, must no doubt be explained by the greater complexity of the later planetary theories. After the introduction of so many epicycles, the simplification realised by placing the sun in the centre was so much less striking than in the case of the simple theory of Aristarchus' days, that it appeared hardly worth while. The equivalence of the geocentric and heliocentric systems was, however, perfectly

realised. But the Greeks, at least since the first century B.C., left the examination of "true causes" to physics, i. e. to what we would call metaphysics. Astronomy had for its only task to describe the planetary motions in a formally correct way, to account for the observed phenomena: *σώζειν τὰ φαινόμενα* — which we should be careful not to translate too literally as "saving the appearances."

In the later middle ages, the time of the scholastics, the writings of Aristotle were the great authority, and were supposed to contain all the wisdom of the ancients. There is, however, a distance of nearly five centuries between Aristotle and Ptolemy, and it is, of course, historically entirely incorrect to represent the work of these two men as belonging to practically the same stage in the development of science, a mistake that even in our own time is often made.

Ptolemy is the last Greek astronomer. Nothing was added to science after his time, and Greek civilisation was finally destroyed by the burning of the library of Alexandria by the Christian mob in A.D. 389.

Then come the dark ages of orthodox ignorance. The Greek writers were only rediscovered in the twelfth and thirteenth centuries in Arabian translations, but the teaching of Aristotle was still forbidden at the University of Paris in 1215. Only very gradually the way was cleared for Copernicus, who took the same step that Aristarchus had taken eighteen centuries before him.



NICOLAUS COPERNICUS

II

BIRTH OF MODERN ASTRONOMY · KEPLER, GALILEO, NEWTON · THE EIGHTEENTH CENTURY

THE greatest observer of the pretelescopic time is TYCHO BRAHE. A Swedish nobleman by birth, he travelled in his youth and was recommended to King Frederic II of Denmark by the astronomer-duke William IV of Hessen. The King gave him the isle of Hveen, where he built his famous observatories Uraniborg and Stjerneborg. He was a born observer, and for him the making of observations was the primary and necessary thing. Tycho has often been blamed for rejecting the system of Copernicus and replacing it by his own having the earth in the centre. He does, however, deserve no blame for this, but rather praise. He adopted Copernicus' system so far as it could be reconciled with his observations, i. e. he made all the planets revolve round the sun. But in the further step of placing the sun in the centre he could not follow Copernicus, since he had, as he thought, observational evidence against it. He had tried to determine the parallaxes of the fixed stars and found that they were smaller than could be observed with his instruments. The enormous distance which this implied for the fixed stars was of course one of the current arguments of the time against the Copernican system. But this was not Tycho's objection; he was convinced that he had observed the diameters of the



TYCHO BRAHE

fixed stars, and his argument was that, as he found no parallax, the real diameter must be much larger than the orbit of the earth round the sun, and this was to him unacceptable. Galileo has shown where the flaw in the argument lies, viz. that the diameters Tycho thought he had measured were not true angular diameters but were produced by irradiation in the eye. As Galileo says, the eye provides the star with an envelope of hair that is shorn off by the telescope. Kepler had previously found the same explanation.

After the death of King Frederic difficulties arose for Tycho in Denmark and he went to Prague, where Kepler came to him to work as his assistant. It is very interesting to read Kepler's appreciation of Tycho, as he expressed it in a letter to Maestlin. He writes:

My judgement about Tycho is this: he is immensely rich, but, as most rich people, does not know how to use his riches. One must therefore try (and I for one have done so with all convenient discretion) to wrest his treasures from him, to persuade him to publish his observations, all of them, without any reserve.

And again in another letter to the same lifelong correspondent, written shortly after Tycho's death:

That which Tycho has achieved, he has achieved before the year '97; from that time onwards his condition turned to the worse, he was besieged by tremendous cares and began to become puerile. The inconsiderate departure from his native country oppressed him. The court here [i. e. in Prague] has absolutely ruined him. He was not the kind of man, who might live with everybody without serious difficulties, not to speak of the intercourse with men in such high positions, the supercilious counsellors of kings and princes. Tycho's principal achievement are his observations, many voluminous books as numerous as the years during which he has

conducted this work. But also his *Progymnasmata* (in which he treats of the fixed stars and the motion of sun and moon for our time) really have a fragrance of ambrosia. I hope to publish them at the next annual fair; I am making a supplement to it. With regard to the moon: in the last years the work has chiefly been done by a certain Christian Severini Longomontanus of Denmark under Tycho's direction. This work does not show the superiority that is found in the theory of the sun. About the comets Tycho intended to write another book; about the planets he has performed quite learned, industrious investigations, but rather in the way of Ptolemy, *mutatis mutandis*, just as Copernicus also has done. You can see from that how God distributes his gifts; no one of us can do everything. Tycho has done the same as Hipparchus; that is with reference to the foundation of the building; he has in this completed an enormous work. But one man cannot do all. Hipparchus requires his Ptolemy to build the theories of the other five planets on his foundation. Already during Tycho's lifetime I have done this.

No apology is needed for these lengthy quotations. They are interesting not only because they show that Kepler formed an entirely correct estimate of the importance of Tycho's work, and the relation between it and his own, but also because they give an insight into the character of that very remarkable man, JOHANNES KEPLER. He was born in 1571, studied divinity in Tübingen, and in his twenty-third year was appointed State Mathematician and Professor in Graz, where he married and where he wrote his first work: "*Prodromus dissertationum cosmographicarum seu mysterium cosmographicum.*"

He had a rich phantasy and a mystical disposition, he was still entirely under the influence of the scholastics and readily accessible to their ways of reasoning; thus, e. g., he adopts with absolute conviction the argument of Aristarchus about

the proportionality between the radius of the earth, the distance of the earth from the sun, and the radius of the sphere containing the whole solar system. He even adds two more members to this sequence. As another example of this influence may be considered his method of finding the total mass of the universe. He divides the total mass into three equal parts, viz. the mass of the sun, the mass of the planets, together with that of the aether within the sphere of the fixed stars, and the mass of the fixed stars and the material which forms that sphere. By a similar reasoning about proportionalities he computes the radius and the volume of the sphere of the fixed stars and thus finally finds the density of the aether.

On the other hand he often expresses opinions far in advance of his time. Thus, in letters written about the same time he suggests the possibility that the sun is a fixed star and that the fixed stars are suns surrounded by planets.

Evidently he feels closely related to the Pythagorean school; the name Pythagoras often occurs in his letters, whilst Aristotle is never mentioned.

In the *Prodromus* he exposes the system of the five regular bodies (he calls them the Pythagorean bodies), which he uses to explain the distances of the planets. He supposes Saturn to be attached to a sphere; within that sphere is described a cube; to the sphere inscribed in that cube Jupiter is attached. Then comes a tetrahedron and within that the sphere of Mars, etc. The large distance between Jupiter and Mars is explained by the fact that it is the tetrahedron which is between their spheres. The thickness of the spheres is used to account for the excentricities of the planetary orbits.

However phantastic these first attempts may be, they show the independence of his mind. He wants to find out the truth for himself, unfettered by any tradition and recognising no authority beyond the evidence of the ascertained facts. In Prague, when he had access to Tycho's observations, he saw the possibility of making new planetary theories, free from all hypotheses, and based on observation alone, and he forthwith made this his life-work. He is the first who has finally done away with the circle and the uniform motion. From his point of view there was not much difference between the systems of Ptolemy and Copernicus. Kepler went to work without any preconceived ideas about what the true form of the planetary orbit was going to be. He tried one curve after another, before he finally found the ellipse. But he was never discouraged by his failures: "Who never doubts, can never be sure of anything," he says, and goes on to make another attempt.

Especially the finding of the second law took him a long time; he tried several relations before he found the true law expressing how the velocity depends upon the distance. First he supposed the velocity to be proportional to the diameter of the sun as seen from the planet; afterwards he made it proportional to the quantity of light the planet receives from the sun, which is the correct law. Concurrently with these investigations he indulges in the wildest speculations. But these never were more than "working hypotheses" to him, and he was as ready to give them up as he was fertile in inventing them.

Gilbert's book on the magnet, which appeared about this time, made a deep impression on him and confirmed him in

his conviction that the force which the planets experience originates in the sun.

As an evidence of his enthusiasm, optimism, and force of mind may be given another quotation from the same letter to Maestlin, written after Tycho's death. He has been given the care of Tycho's instruments and unpublished observations, but he knows not yet whether he will also succeed Tycho as Imperial Astronomer, and by the salary attached to that position (which salary was, however, only rarely paid) will be enabled to publish Tycho's works and to continue his researches. He writes:

How the answer will be time will show. I myself am full of hope. For, when God has provided me with that which is most important, viz. a subject on which I can use my talents, He will also provide the necessary money. If, however, God takes an interest in astronomy, which to believe piety requires, I hope that I shall achieve something in this subject, as I see how God has united me with Tycho by constant destiny and has not let us be separated, even through the worst of disagreements.

Modern physics is usually taken to originate with GALILEO (1564 to 1642). He was born in a time when new ideas were already rife, and doubt of the authority of Aristotle began to arise. His education was of the best that the times could afford. At the university of Pisa, where he stayed for three years, he heard lectures on astronomy in his third year. In these lectures the system of Copernicus was mentioned, as a heresy, of course, like those of Pythagoras and Aristarchus, but also the new determination of the length of the year by Copernicus (which is at the base of the Gregorian calendar) was discussed. After his three years at Pisa he went to Flor-

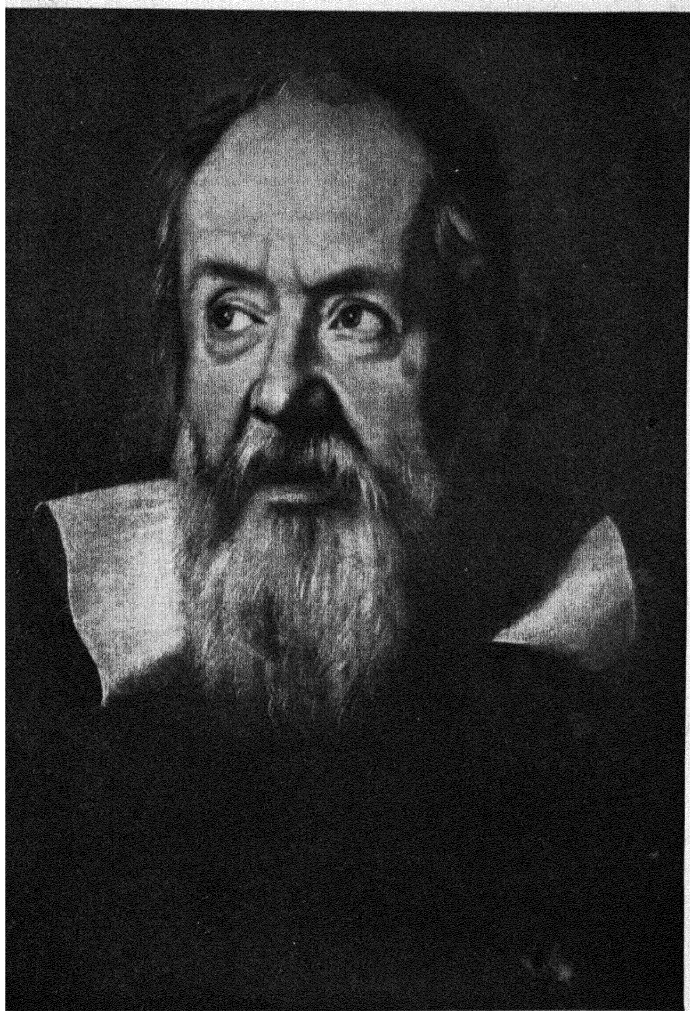
ence, and there he continued his work independently. It is in this time that he studied the works of Archimedes, which made a great impression on him. He may be said to be the direct successor of Archimedes as a theoretical physicist. Thus his first researches are about the centre of gravity, hydrostatic pressure, and the determination of specific weights by means of the hydrostatic balance.

The important event which decided his destiny was the invention of the telescope. It seems certain that Galileo found out for himself how to construct a telescope, and actually made one, after hearing rumours about those made in Holland. The many important discoveries made with his telescope have been of the greatest influence on the development of astronomy. He was the first to see the satellites of Jupiter, the phases of Venus, the strange appendices of Saturn, afterwards interpreted as the rings by Huygens. He discovered the spots on the sun and from them derived the rotation of the sun. He observed that the milky way consists of a great number of stars; this was the first time in history that the fixed stars were made a subject of research for their own sake, and on this observation is based the first theory of the milky way, that of Kepler, who assumed it to be a ring of stars.

All these discoveries, especially that of Jupiter's satellites and the phases of Venus, have been of great weight in the decision between the system of Copernicus and that of Ptolemy. It is well known that Galileo's championship of the Copernican system is what led to his martyrdom, which to a large extent is responsible for the prominent position he occupies in history. It is a rather remarkable fact that

this formal difference between the current and the new planetary systems should have been singled out by the Church, Roman Catholic as well as Protestant, and by public opinion, as the one outstanding modernisation and heresy. As an explanation of the observed phenomena they are entirely equivalent, as was already known to the ancients. In astronomy the work of Kepler is much more revolutionary and of much greater fundamental importance. And the same is true in still higher degree of Galileo's own work on the subject of mechanics. Public opinion is not a good judge in picking out the essential points of the profound changes of outlook that are going on at any time. In the first decades of the sixteenth century the real conflict was between new and old, between science based on observations and science based on authority. The Copernican system was, erroneously, selected as the scapegoat. Similarly in our own times the theory of relativity is popularly held up as upsetting everything and contrary to tradition and common sense, whilst in reality the work of Rutherford, Bohr, and their successors is much more revolutionary.

The really important part of Galileo's work is not his advocacy of the Copernican system, but, apart from his observations, his investigations in the domain of mechanics, which are a direct continuation of the work of Archimedes. Until Galileo's time the doctrines of Aristotle were generally adhered to, in which the position of a body is considered as the primary thing: a change of position, i. e. a *velocity*, requires a force to cause and maintain it. For Galileo the velocity is the primary thing, that does not require an explanation: it is *acceleration* for which a cause must be found. His laws about



GALILEO GALILEI

free-falling bodies, the pendulum, the law of inertia, date from his early years, about 1604. He enounced for free-falling bodies the law of the increase of the velocity proportionally with the time, i. e. the constancy of the acceleration. He has been on the verge of the definition of mass, inertia, constancy of momentum, all those fundamental concepts of the science of mechanics which have only been made into a consistent system by Newton, three-quarters of a century later.

In his subsequent life Galileo was drawn away from these theoretical investigations by his observational work with the telescope, and his interests were more in a practical direction. Amongst other problems he turned his attention to the finding of longitude at sea, and pointed out the possibility of using the eclipses of Jupiter's satellites to that end.

In his last years, when he was old and blind, Galileo made projects for the application of the pendulum to the measurement of time; so he was, more or less, on the way towards the invention of the pendulum clock.

By the invention and improvement of the telescope a period of great progress is inaugurated. It does not come suddenly but ripens only gradually. The telescope at first, even after it had been optically much improved, was only an instrument of discovery. The great need of theoretical astronomy, after Kepler had derived from the observations of Tycho all that they could give, was for increased accuracy in measurement for the determination of positions. Before this could be achieved, the mechanical parts of the instruments, the mountings, the divided circles, had to be improved so as to catch up with the sudden enormous optical improvement.

In this general improvement of mechanical devices one of the leading personalities is PICARD (1620 to 1682, the founder of the *Connaissance des Temps*). He may be called the first modern observer, the Bessel of the seventeenth century. He was the first to use a telescope on an instrument for measuring angles, and also the first to give a theory of instrumental errors.

So the advantages of the new observational results became only gradually available, and not before Newton's time was the corresponding progress in theory made. The same unavoidable delay of theoretical after observational advance is notable in our times: photography and spectroscopy were discovered about 1865, whilst the great revolution in theory did not take place until about 1915.

The dominating personality of the time of theoretical development is SIR ISAAC NEWTON (1643 to 1727).

His principal achievements are the discovery of the law of gravitation and the creation of the system of mechanics based on the law of inertia.

His great book, the "Principia," is by far the most beautiful example of reasoning purely inductive, without one single hypothesis.

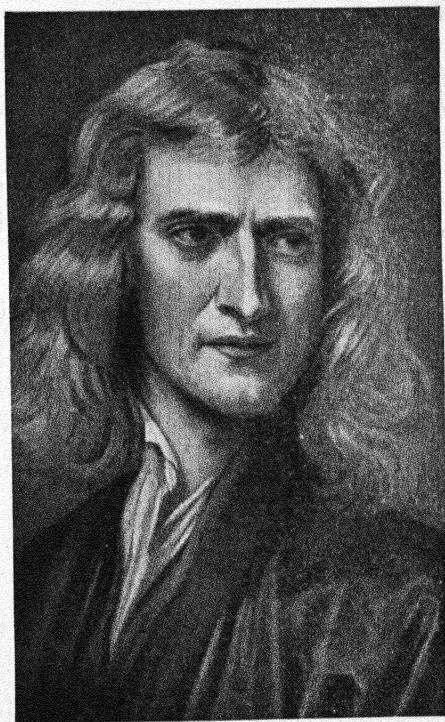
He starts from his definitions:

Definition I. The quantity of matter is the measure of the same, arising from its density and bulk conjunctly.

Definition II. The quantity of motion is the measure of the same, arising from the velocity and quantity of matter conjunctly.

As regards I, we are accustomed nowadays to define density as the quotient of mass and volume, but to Newton's mind density evidently is the fundamental conception.

PLATE VI



SIR ISAAC NEWTON

Newton adds that if there is a medium, filling up the space between the parts and able to move freely between them, that should not be taken into account.

Newton's definitions are not axioms but simply explanations of the meanings of the technical terms he is going to use.

The chapter of the definitions is closed by the scholium, in which Newton explains that it is not necessary to define space and time, as being sufficiently familiar, but he says expressly that there exist an absolute space and an absolute time, independent of the material contents and well to be distinguished from the relative space and time which are used to measure absolute space and time by material means.

Then follow the three laws of motion, the first of which is the law of inertia.

Law I. Every body perseveres in its state of rest, or uniform motion in a right line, unless it is compelled to change that state by forces impressed thereon.

Law II. The alteration of motion is ever proportional to the force impressed; and is made in the direction of the right line in which that force is impressed.

Thus to Newton's mind *mass* (derived from density and volume) is fundamental, then momentum is defined as the product of mass and velocity, and change of momentum is proportional to the force. In the mechanics of the beginning of the twentieth century it is natural to consider the second law as the definition of (inertial) mass, thus starting from acceleration and force, force being a fundamental concept. Einstein returns to Newton's conception, considering not the force but the density as the fundamental thing, and the force, or the potential, of which it is the gradient, is derived from the material density.

Law III. To every action there is always opposed an equal reaction; or the mutual actions of two bodies upon each other are always equal and directed to contrary parts.

This law more or less loses its importance if mechanics is also applied to forces that are not central. In the way Newton expresses it, it is only valid for forces in the direction of the line connecting the two bodies considered, and can be derived from the law of conservation of momentum. As being a combination of I and II, it is less fundamental than these.

After the introductory chapters of Definitions and Laws of Motion, the first two books are devoted to the working out in detail of the system of kinematics and dynamics based on them. The third book brings *the general law of gravitation*, which is deduced step by step from the observations:

1st phenomenon: the orbits of the satellites of Jupiter around the planet are in accordance with the laws of Kepler,
2nd phenomenon: the planets describe orbits around the sun,
3rd phenomenon: in the orbits of the planets Kepler's second law is obeyed,

and so on. Thus gradually by careful deduction the general law of gravitation is arrived at.

The "Principia" ends with the "Scolium generale," which gives somewhat the impression of a profession of faith. It begins with the words, "The hypothesis of vortices is pressed with many difficulties," and ends:

But hitherto I have not been able to discover the cause of those properties of gravity from phaenomena and I frame no hypotheses; for whatever is not deduced from the phaenomena is to be called an *hypothesis*, and hypotheses, whether metaphysical or physical, whether of occult qualities, or mechanical, have no place in *Experimental Philosophy*. In this philosophy particular propositions are

inferred from the phaenomena and afterwards rendered general by induction. Thus it was that the impenetrability, the mobility and the impulsive force of bodies, and the laws of motion and of gravitation were discovered. And to us it is enough that gravity does really exist, and acts according to the laws which we have explained and abundantly serves to account for all the motions of the celestial bodies and of our sea.

Newton's great renown and authority have contributed much to the high esteem enjoyed by purely inductive reasoning. The period of pure logical induction began with Kepler, reached its height in Newton's work, and went on during the whole of the eighteenth and nineteenth centuries. It seems to be closed now however: in our times theory is again bristling with hypotheses.

A point of interest is the question why Newton waited such a long time before publishing the "Principia." It is almost certain that he discovered the law of gravitation in 1665, whilst the book only appeared in 1687, and he had kept it so secret that only some very intimate friends knew about it. It has often been assumed that the reason of this delay was the imperfect knowledge of the radius of the earth, which spoilt the comparison of the theory with observations. A more plausible explanation, however, is the one given in 1887 by Adams and Glaisher. Newton had found it very difficult to prove that the resultant force, with which a homogeneous sphere acts on a point outside it, is the same as it would be if the whole mass of the sphere were concentrated in the centre. So long as this apparently comparatively unimportant but really fundamental point had not been cleared up to his entire satisfaction, he considered the theory as not finished and shrinked from publishing it.

Newton's theory was far from being generally accepted at once. Huygens and Leibniz felt an aversion of its formal character, an antipathy against action at a distance. Huygens saw very well that Descartes' theory of vortices was in many respects untenable, but it explained at least *how* the force is transmitted, and did not use the mysterious action at a distance. Many of the contemporaries of Huygens had the same objection, and only in the time of rationalism, the age of Voltaire, did philosophers and men of science generally surrender to Newton's way of thinking.

Gradually the spirit of Newton has conquered the whole world. The law of gravitation came to be considered as the ideal physical law, the model after which all laws had to be shaped. Thus we find attractions and repulsions proportional to some inverse power of the distance all over physics for a considerable time. The elaboration of the law of gravitation itself and its application to the planetary perturbations was the work of the great mathematicians of the eighteenth century: Euler (1707 to 1783), Clairaut (1713 to 1765), Lagrange (1736 to 1813), culminating in Laplace's great work, the "*Mécanique céleste*."

Besides this mathematical development there is also a progress in the development of observational astronomy. Important events are the foundation of the observatory in Paris in 1669, and that of Greenwich in 1675; the first Astronomer Royal, Flamsteed (1646 to 1719) may be mentioned here. His great successor, Halley (1656 to 1742), has enriched astronomy by many important discoveries. He was the first to discover the proper motions of the fixed stars

and the secular term in the longitude of the moon, by a comparison of observations made in his own time with those of the Greek astronomers and with Arabian observations made in the tenth and eleventh centuries. He made the first catalogue of southern stars, having undertaken to that end an expedition to St. Helena. He was not only an observer, but a theoretical astronomer as well; he applied the theory of gravitation to the comets, and his tables of the motions of the planets were for a long time the best in existence.

It cannot be our aim to give a history of astronomy during the eighteenth century, or even to mention the chief events in that history. I will, however, mention two more names, before we make the jump over a century from Newton to Herschel.

Öle Roemer (1644 to 1710) discovered from the eclipses of Jupiter's satellites the finite velocity of light, he built the first meridian circle, and probably also the first equatorially mounted telescope.

James Bradley (1693 to 1762), "*vir incomparabilis*," discovered nutation and aberration and laid the foundations of modern fundamental astronomy.

All these discoveries are the natural outcome of the development of the art of observation as a consequence of the evolution in the technique of instrument making. Astronomical observation for a time reached its height in Bradley, and has only been surpassed much later.

III

THE BEGINNING OF SIDEREAL ASTRONOMY · WILLIAM HERSCHEL AND HIS SUCCESSORS

THROUGHOUT the whole of astronomy up to the end of the eighteenth century hardly any attention had been given to the fixed stars. The conception of the ancients that the fixed stars are attached to a sphere and thus all have the same distance from us, has only gradually been overcome. Kepler still speaks of the fixed stars as being confined to a shell between two spheres and says not much more about them.

The first actual determination of the distance of a star is that of Sirius by Huygens. There are two simple properties of light on which the determination of distances can be made to depend: the rectilinear propagation, which is the foundation of the trigonometric methods of determining parallaxes, and the diminution of the intensity inversely as the square of the distance. The trigonometric method is the more fundamental one; we only resort to the second alternative for distances which are too large for the first. So it was in all times, only up to nearly the middle of the nineteenth century *all* distances were too large for the direct trigonometric method. It had been tried by Tycho; Kepler had suggested it to Galileo, indicating the exact observations to be made and the dates on which they should be taken. The principal incentive for making these observations, however, was not to find the distances of the stars, but to prove the Copernican

system by the parallactic motion. The distance as determined from the intensity of the light, of course, would be practically the same whichever planetary system were the true one.

Huygens compared the light of Sirius with that of the reflection from a small sphere of glass of a small portion of the sun's disc cut out by a diaphragm. Assuming the true luminosity of Sirius to be the same as that of the sun, he found the distance 28000 times the distance of the sun, and concluded that the parallax must be too small to be measured. The same conclusion he also drew from the fact that the stars Mizar and Alcor show no relative motion with an annual period. He thus anticipates by more than a century Herschel's suggestion to determine relative parallaxes from the relative displacements of bright and faint stars, which are close together in the heavens, assuming, as Herschel did, that the intrinsic luminosity is the same. It is well known that Herschel, by actually trying this method in several cases, did not find a parallax, but made the discovery of binary stars. Had Huygens known, as we do now, that the light of Sirius is about twenty-five times that of the sun, his distance would have come out only four times too large, instead of twenty times.

WILLIAM HERSCHEL is the first to have the idea that the fixed stars form a system having a certain structure and to study this structure for its own sake, basing the investigation of it on actual observations.

Kepler and Huygens had suggested that the sun was a star, and the stars were suns, attended by planetary systems.

Kant and Lambert — starting from very different principles and reasoning on different lines — had both built up systems of the world which are entirely speculative. The only certain fact concerning the structure of the stellar system known before Herschel began his work was that the milky way consists of many stars. This had been discovered by Galileo.

Special mention should however be made of Herschel's elder contemporary John Michell who published (in 1767) some remarkably correct views regarding clusters of stars, and on proper motion, which he considered must be partially due to the motion of the sun, pointing out that the secular parallax may in time give a means for the determination of the distance. The direct measurement of parallaxes, he says, will have to await the construction of better telescopes, able to show a twentieth of a second of arc, which he does not consider hopeless in future. But he says nothing on the milky way, nor on the general structure of the stellar system.

Herschel points out that we know from direct observation only two coördinates of the stars; the third, the "profundity" as he calls it, is unknown.

The stars are very different in brightness — is this due only to their different distances, or are they also different in intrinsic brightness, "luminosity" as we call it now, by a name introduced by Kapteyn?

The distribution of the stars over the sky also is very inhomogeneous. Is their real distribution in three-dimensional space regular and homogeneous or not? We have thus really *two* unknowns: the luminosity and the density. It was only Kapteyn who tackled the problem in its completeness. Herschel only began. He makes the assumption that the

luminosity of all stars is the same, *and* that the density is constant. Evidently these two assumptions are contradictory, i. e. not a priori: a universe *might* exist in which both are true, but it would look very different from ours.

In Herschel's first communication "On the parallax of the fixed stars" (1781) he enounces explicitly the two following

GENERAL POSTULATA

1. Let the stars be supposed, one with another, to be about the size of the sun.
2. Let the difference of their apparent magnitudes be owing to their different distances, so that a star of the second, third, or fourth magnitude is two, three, or four times as far off as one of the first.

With respect to the first I remark that it appears in several places that Herschel considers the surface luminosity of the sun and the stars to be the same; "size" thus refers to the actual dimensions.

With respect to the second Herschel himself says in a footnote that this amounts to a definition of magnitude. The scale of magnitudes, by which the apparent brightness of the stars is measured, is derived by tradition from the Greek astronomers. The connection between the conventional scale and the actual ratio of intensities had to be found by photometric methods. Herschel does not attempt to do this, but postulates his definition. Since he accepts, and uses, the law of the inverse square, this defines the magnitude as being inversely proportional to the square root of the intensity. Sir John Herschel, by a photometric investigation at the Cape, confirmed this, finding the formula $l = 1/(m + 0.414)^2$. The logarithmic scale which we use now was, of course, only

adopted much later. Herschel evidently considers this law, $m = \Delta$, as almost self-evident, and he has stuck to it consistently throughout his life, also inverting it, and, e. g., speaking of a star of the magnitude 1342, meaning one at a distance 1342 times the average distance of a star of the first magnitude. In the footnote quoted he calls attention to the possibility of absorption in space, suggested to him, as he says, by the fact that he "generally found the very small telescopic stars red, or inclining to red." If there were absorption, of course the distance would be smaller than the magnitude number. But this footnote is the only place in his works where absorption is mentioned. He takes no account of it in his further deductions.

Herschel's great work are his star gauges, started as soon as the 20-foot telescope (focal length 20 feet, aperture 18.7 inches) was completed, i. e. 1784. As the first result of his scrutiny of the heavens he comes to the conclusion that most of the so-called nebulae are resolved in stars, and he surmises that all of them will be so resolvable with larger telescopes. The milky way — the galactic system, we should say — thus is an immense "sidereal stratum." His conception is explained by a figure, which clearly shows the tacitly assumed even distribution, i. e. constant density, on which his first theory depends.

In the second paper he begins by stating his ideas concerning different stages of nebular development — all nebulae still are assumed to consist of stars. Then comes his conclusion: "We inhabit the planet of a star belonging to a Compound Nebula of the third form." A "compound nebula of the third form" in his nomenclature is a system containing

globular clusters, irregular clusters, and stars arranged in "links, hooks," etc. Its detailed structure is thus not one of uniform density. Still Herschel continues to use the hypothesis of strictly uniform density consistently, but it would seem that even at that time he considered this to be only a large-scale approximation. However this may be, he goes on: "I shall now proceed to shew that the stupendous sidereal system we inhabit, this extensive stratum and its secondary branch, consisting of many millions of stars, is, in all probability, a *detached Nebula*," i. e., in Herschel's vocabulary, an isolated system, surrounded everywhere by empty space.

If the density is constant, then the distance to which the system extends in a certain direction is directly proportional to the cube root of the number of stars seen in that direction, provided we can see to the limit of it. Herschel puts it thus:

Problem: The stars being supposed to be nearly equally scattered and their numbers, in a field of view of a known angular diameter, being given, to determine the length of the visual ray.

This does, of course, depend on the arrangement, whether the stars are supposed to be at the vertices of a cube, or of a tetrahedron, or any other possible regular distribution. He takes as a plausible distribution that in the vertices of cubes. Then the solution of the problem is simple. The "Table of Star-Gages" contains the result of 683 gauges. The numbers of stars vary from 0.5 to 588. Each number is a mean of several fields in the neighbourhood, often ten. It is remarkable that many of the very poor fields are in the milky way; so, e. g., the poorest of all, 0.5 stars (mean of 10 fields), is in Scorpio. Herschel calls it "an opening in the heavens."

From these numbers he then derives the distances by the rule resulting from his solution of the problem mentioned before. Thus he is able to construct a model of the system. In his paper he only gives a section by a plane perpendicular to the central plane of the milky way. This section, depending on 129 gauges, lying approximately in a great circle through the constellations Aquila, Coma, Canis Major, Cetus (galactic longitudes 191° and 11°), is reproduced here (Fig. 3). It will

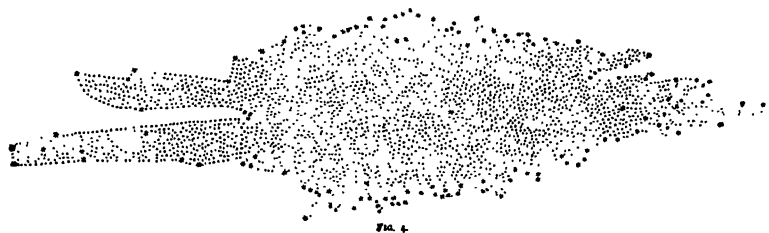


FIGURE 3

Section of the galactic system according to Sir William Herschel's first theory (1785)

be seen that the hypothesis of uniform density gives a rather irregular outline. The shortest distances from the sun to the boundary are 75 units towards Coma (top of figure) and 83 units in the opposite direction, towards Cetus. The longest radii towards the left, in Aquila, are 420 and 497 units respectively, against 352 units in the direction of Canis Major. The unit is the distance of Sirius, or two parsecs. Herschel then goes on to state his grounds for believing that he has really reached the confines of the system. For

it is true, that it would not be consistent confidently to affirm that we were on an island unless we had actually found ourselves every where bounded by the ocean.

But, says Herschel, if there were still more stars beyond, too faint to be seen separately in his telescope, the gauges would show a milky unresolvable nebulosity as a background, and this they do nowhere:

so that from these considerations it appears again highly probable that my present telescope, not shewing such a nebulosity in the milky way, goes already far beyond its extent.

Our system thus is, he concludes, "*A very extensive, branching, compound Congeries of many millions of stars.*" The number of stars it would contain according to Herschel's data is from 75 to 100 millions, but Herschel himself makes no more accurate statement than "many millions."

This then is the system of 1785.

At the end of this paper he enumerates many other nebulae which are, in his opinion, also galactic systems. The nearest and largest of these is the Andromeda nebula. Its distance is estimated as two thousand times the distance of Sirius, i. e. 4000 parsecs or 13000 lightyears. Combining this estimated distance with Herschel's estimate of the apparent diameter, "above a degree and a half," we find for the true diameter 50 to 60 of Herschel's units, or 100 to 120 parsecs, whilst the diameter of our own system, according to Herschel's figures, is about 850 units, or 1700 parsecs.

The foundation on which Herschel's first theory (1785) rests is his belief that all nebulae are resolvable in stars. His continued observations eventually gave him the conviction that this is not so. In a paper of 1791 he distinguishes between resolvable nebulae, or clusters, and "milky nebulosity," which he considers to consist of "a shining fluid of a nature totally unknown to us." He then enlarges on several

hypotheses as to the nature of this "shining fluid": "Can we compare it to the coruscations of the electric fluid in the aurora borealis?" This he dismisses, as well as the hypothesis that it should be of the same nature as the zodiacal light, since that is too faint to be visible from any distance, and he finally ventures to suggest that it consists of particles of light, deflected from their straight course through space by various influences, among which is the gravitational attraction of stars and stellar systems. "However," he winds up,

we shall lay no stress upon a surmise of this kind, as the means of verifying it are wanting: nor is it of any immediate consequence to us to know the origin of the luminous matter. Let it suffice, that its existence is rendered evident, by means of nebulous stars.

The great 40-foot telescope (aperture 4 feet) was completed by the end of 1789, and although its use was cumbersome, and only very few nights were good enough for it, it certainly was of great help to Herschel in clarifying his ideas. Even one good night with it would have sufficed to bring about the evolution in his ideas that finally resulted in his second theory of the galactic system. Besides the existence of unresolvable nebulosity, the fact that with the great telescope he saw more stars in the same field than with the 20-foot led him to abandon the view that he had reached the confines of the universe: "The 20-feet telescope could not fathom the Profundity of the milky way." No more can the 40-foot; there remains always a background.

Hence we may conclude, that when our gages will no longer resolve the milky way into stars, it is not because its nature is ambiguous, but because it is fathomless.

"Ambiguous," in Herschel's language, means composed of stars and nebulous matter mixed.

Moreover he has gradually come to see that the uniform density is not even an approximation: he has discovered in the milky way the "clustering power"; the stars are inclined to unite into groups, and he believes that the milky way will ultimately be entirely broken up into separate groups and clusters. The galactic system is, as a consequence of the mutual attraction of the stars, "subject to clustering power . . . [and] must be finally broken up, and cease to be a stratum of scattered stars."

He thus gives up the uniform density as a guiding principle, and uses in its stead the uniform luminosity:

Stars, at least those which seem to be promiscuously scattered, are probably one with another of a certain magnitude.

Therefore the fainter the stars we see in a certain direction, the further we are looking into space. Then, taking the magnitude to be inversely proportional to the square root of the intensity, according to his definition of magnitude, the magnitude of the faintest stars is derived theoretically from the light-grasp of the telescope. Or, rather, he gives up the use of magnitudes entirely, and works with the distances only. To determine these he computes what he calls the "penetrating power" of his telescopes. The loss of light by absorption and reflection in the eyepiece is carefully determined and duly taken into account. The aperture of the pupil of the eye is supposed to be a fifth of an inch. The penetrating power of the naked eye, which is in this paper, contrary to his usual practice, supposed to go to the seventh

magnitude, being taken as unity, that of the 20-foot used as a Newtonian is thus found to be 61.2, as a front-view it is 75, whilst of the 25-foot it is 95.85 and of the 40-foot 191.69. This latter thus shows stars of the magnitude 1342 ($= 7 \times 191.7$). Later (1817) he determined the unit in terms of the average distance of the stars of the first magnitude. By comparing the numbers of stars of different magnitudes up to the seventh, he found the law that the magnitude is equal to the distance to be inexact. By actual photometric comparison (by equalising the light by means of diaphragms) he found for different stars of the sixth magnitude, which is now again taken as the limiting magnitude for the naked eye, between $(1/10)^2$ and $(1/15)^2$ th of the intensity of a star of the first magnitude. He then adopts as an average for the "visual ray" of the naked eye the number 12, the unit being, as already stated, the average distance of a star of the first magnitude. The visual ray of the 20-foot thus becomes $75 \times 12 = 900$, and of the 40-foot 2300. Even at that distance the limit of the system was not reached. The "Concluding Remarks" of his paper, dated Slough, near Windsor, May 10, 1817, may be taken to represent his latest views on the "construction of the heavens," to the investigation of which he has given all his best efforts. They are as follows:

X. CONCLUDING REMARKS

What has been said of the extent and the condition of the milky way in several of my papers on the construction of the heavens, with the addition of the observations contained in this attempt to give a more correct idea of its profundity in space, will nearly contain all the general knowledge we can ever have of this magnificent

collection of stars. To enter upon the subject of the contents of the heavens in the two comparatively vacant spaces on each side adjoining the milky way, the situation of globular clusters, of planetary nebulae, and of the far extended nebulosities, would greatly exceed the compass of this Paper.

Apart from the inclusion of planetary nebulae and far extended nebulosities (which are the irregular nebulae, like the one in Orion) amongst the extragalactic objects, this could



FIGURE 4

William Herschel's diagram illustrating his second theory (1817)

almost have been written fifteen instead of a hundred and fifteen years ago. A small diagram is added, which is reproduced in Fig. 4. The similarity with Shapley's diagram of 1918 (see fifth lecture, Fig. 6) is very striking.

The diameter of the circle, which is a section of a sphere containing all naked eye stars, is 24 units, or, taking one of Herschel's units to be two parsecs, — the parallax of a star of the first magnitude has been assumed to be $0''.5$ by Herschel throughout his life, — 48 parsecs. The thickness of the milky way stratum is 2×39 units = 156 parsecs. These dimensions, which appeared colossal to Herschel and his contemporaries, look very moderate to us.

Herschel's first theory led to a completed system; his later observations and deductions invalidated these conclusions, and he began — at the age of about sixty — to reconsider his old views, and to build his system of the world on entirely new lines. In the great paper of 1785 he says already that he considers his present results as provisional, and liable to be changed by further observations.

I shall therefore wait till the observations in which I am at present engaged shall furnish me with proper materials for the disquisition of so new a subject. And though my single endeavours should not succeed in a work that seems to require the joint effort of every astronomer, yet so much we may venture to hope, that, by applying ourselves with all our powers to the improvement of telescopes, which I look upon as yet in their infant state, and turning them with assiduity to the study of the heavens, we shall in time obtain some faint knowledge of, and perhaps be able partly to delineate, *the Interior Construction of the Universe*.

And he characterises his method of working in the following words:

But first let me mention that, if we would hope to make any progress in an investigation of this delicate nature, we ought to avoid two opposite extremes, of which I can hardly say which is the most dangerous. If we indulge in a fanciful imagination and build worlds of our own, we must not wonder at our going wide from the path of truth and nature; but these will vanish like the Cartesian vortices, that soon gave way when better theories were offered. On the other hand, if we add observation to observation, without attempting to draw not only certain conclusions, but also conjectural views from them, we offend against the very end for which only observations ought to be made. I will endeavour to keep a proper medium; but if I should deviate from that, I could wish not to fall into the latter error.

The paper of 1811, dated Slough, May 26, 1811, begins:

A Knowledge of the construction of the heavens has always been the ultimate object of my observations, and having been many years engaged in applying my forty, twenty, and large ten feet telescopes, on account of their great space-penetrating power to review the most interesting objects discovered in my sweeps, as well as those which had before been communicated to the public in the *Connoissance des Temps*, for 1784 [i. e. Messier's catalogue], I find that by arranging these objects in a certain successive regular order, they may be viewed in a new light, and, if I am not mistaken, an examination of them will lead to consequences which cannot be indifferent to an inquiring mind.

If it should be remarked that in this new arrangement I am not entirely consistent with what I have already in former papers said on the nature of some objects that have come under my observation, I must freely confess that by continuing my sweeps of the heavens my opinion of the arrangement of the stars and their magnitudes, and of some other particulars, has undergone a gradual change; and indeed when the novelty of the subject is considered, we cannot be surprised that many things formerly taken for granted, should on examination prove to be different from what they were generally, but incautiously, supposed to be.

In a former paper he once said:

It is sometimes of great use in natural philosophy to doubt of things that are commonly taken for granted.

He had the force of mind to doubt not only generally accepted views, but his own former conclusions.

It was pointed out in the beginning of this chapter that two unknowns are required to determine the structure of the stellar system: the luminosity law and the density law. Herschel saw no means of determining these two unknowns. He based his investigation on *assumptions* regarding them, and naturally chose the simplest assumptions, viz. that the

luminosity and the density were constants. The conclusions drawn from his observations by the help of these two assumptions were contradictory, and he was compelled to confess that he had not solved the problem. Nevertheless no single astronomer, either before or after him, has done so much for the final solution as William Herschel.

His son, Sir John Herschel, continued his father's observations and especially extended them to the southern hemisphere. He did not, however, make any further progress with the solution of the great problem.

The next to take up the general problem was F. G. W. STRUVE, the founder of the Pulkowa Observatory. He also had to make an assumption regarding the density law; he takes it to be a function of the perpendicular distance z from the plane of the milky way only, i. e. he assumes firstly complete symmetry with regard to the galactic plane, and in all directions, and further, like Herschel in his last theory, he takes the extent of the system parallel to this plane to be infinite. The only difference from Herschel is that he makes the density decrease with z instead of constant. From this assumption and the observational data supplied by W. Herschel's gauges (the southern observations of John Herschel were not yet published) he then derives a density law, i. e. he derives empirically the coefficients of a certain algebraical formula, which is supposed to represent the density ρ as a function of z . The formula is rather cumbrous and does not have a convincing look about it. The special feature of Struve's work is that he deliberately abstracts from all details, which had so much fascinated Herschel. All he wants

to find is the great lines, i. e. he seeks a "typical" image (according to Seeliger's nomenclature), leaving the filling in of details for a future time. The discussion of the material is very careful and a model of sound scientific criticism.

Having found the density law from Herschel's gauges, he proceeds to determine the same law from the stars contained in Argelander's *Uranometria* and in Bessel's *Zones*. The latter cannot be assumed to be complete. He therefore derives plenitude-factors by comparison with the *Uranometria* and with Piazz's *Zones*. The density law is thus derived from Bessel's stars of the seventh magnitude and from those of the eighth magnitude. Next, to compare the units of distance in the three cases, these are taken inversely proportional to the cube root of the numbers in the galactic latitude zero. Then, as a control, he computes the density for two linear distances from the galactic plane from the different laws, and finds satisfactory agreement.

He then proceeds to compute, from his density law, the radius of the sphere containing all stars of the separate magnitudes. The radius of the sphere containing all Argelander's stars of the sixth magnitude is taken as unit. Then the radius of the sphere containing Argelander's stars of the first magnitude is 0.142, for the second 0.241, etc. For Bessel's stars of the ninth magnitude the radius is 4.25, and for Herschel's stars, the faintest visible with the 20-foot telescope, it is 25.67. These still are all expressed in an arbitrary unit. But in Struve's time there were some determinations of parallax available. These had been discussed by Peters, and Struve takes from this discussion the mean parallax of a star of the second magnitude to be $p = +0''.116 \pm 0''.014$, correspond-

ing to a distance of 28 lightyears. This number is used to convert the relative distances into absolute ones. Thus the distance of the faintest stars seen by Herschel is found to be 3541 lightyears.

Struve now further discusses these numbers. Instead of assuming the aperture of the pupil of the eye to be 5 mm., as Herschel had done, Struve constructed a small telescope, which he called his "Herschelian modulus," with that aperture (0.211 inch), and magnifying three times. With this modulus he saw 183 stars on the same surface on which Argelander's *Uranometria*, which was intended to contain all stars visible to the naked eye, had 100. The penetrating power of Herschel's telescope was 61.2, but this is referred to an aperture of 0.2 inch — i. e. to the "modulus" — as a unit. Since the modulus shows 1.83 times the number of stars of the sixth magnitude in Argelander's catalogue, this number must be multiplied by the cube root of 1.83, making it 74.8, or practically three times the distance given in the table and derived from the density law.

Struve thus finds a contradiction. This is not to be wondered at, for in computing this distance he tacitly assumes, as Herschel did, that the luminosity is constant — and also that the density is constant along the galactic plane. Struve, however, does not explain the contradiction by these causes, but by absorption in space. He thus finds an absorption of about 1 % of the light over the unit distance, i. e. the distance of a star of the first magnitude.

It is evident that the apparent contradiction, which Struve explains by extinction, can just as well be explained in other ways. John Herschel had suggested that the stars in the

milky way — which he, like Kepler, considered as a ring — were intrinsically fainter than those in our neighbourhood. Evidently the contradiction can also be explained if we assume a decrease of density with distance, as Kapteyn did later. The complete treatment of the problem, considering both the density and the luminosity laws as unknowns to be determined from observations, has only been solved by Kapteyn. This solution, to which we shall return in the next lecture, was not possible until more reliable data on the distances of the stars, as well as on their numbers and magnitudes, were available.

IV

FURTHER OBSERVATIONAL ADVANCES · THE SURVEY OF THE “LOCAL SYSTEM” ·

LIFE-WORK OF J. C. KAPTEYN

THE work of KAPTEYN, about one century after that of Herschel, became possible by the observational progress made during that century. In the preceding lecture I quoted a sentence from William Herschel in which he expressed his hope and conviction that our knowledge of the structure of the universe would increase with the improvement of the telescope, which he considered to be still in its infancy. Not only have telescopes been improved enormously, but, what Herschel could not possibly dream of, the invention of photography and spectroscopy has entirely changed the aspect of the science.

As to the problem of the structure of the galactic system, the first thing required was the determination of parallaxes. Without a knowledge of the distances of the stars no further progress was possible. Herschel always assumed for the parallax of a star of the first magnitude $0''.5$; he does not say how this value is derived — probably it is from the photometric determination made in 1760 by Lambert, who had compared the luminosity of the sun with that of stars of the first magnitude, using Saturn as an intermediate step. Michell did the same some years later, but with less accuracy, and was only able to state that the parallax of stars of the



PROFESSOR J. C. KAPTEYN

first magnitude must be smaller than one second of arc, which was already known to Bradley. Some time afterwards Olbers determined, with Mars and Saturn as intermediate steps, the distances of Aldebaran and Procyon.

The first measured parallaxes are those of 61 Cygni by Bessel, of Vega by F. G. W. Struve, both in 1838, and that of α Centauri by Henderson at the Cape, published in January 1839. Thus when Struve did his work on the stellar system, which was mentioned in the last lecture, he could dispose of some parallaxes, but the number was still too small to base reliable conclusions on.

The measurement of the distances of the stars must, like everything else that we know about them, be based on the study of the light we receive from them. Two simple properties of light are used for the purpose, viz. the rectilinear propagation, and the decrease of the intensity in inverse proportion to the square of the distance. The second method had, as we have seen, been used already by Huygens, Michell, Lambert, and Olbers. But these were hardly to be called determinations, since they all implied the arbitrary assumption that the intrinsic brightness of all stars is the same as that of the sun. The first method, which is independent of this or any other assumption, could only be used after the invention of the screw micrometer (and the heliometer) by Fraunhofer about 1820, and the improvement of optics, also chiefly by Fraunhofer.

The principle is quite simple: of a triangle one side (the base-line) and the two angles at both ends are measured, and the third angle (the angle at the star, called the parallax) and the other sides (the distances from the two ends of the

base-line) can then be found. The base-line must be known. In astronomy the unit is the astronomical unit, i. e. the mean distance of the sun from the earth, and in the ordinary, so called trigonometric, determinations of distances the base-line is the diameter of the orbit of the earth round the sun, or two astronomical units. To express this in our ordinary unit of distance, say the centimeter, we should have to *measure* it, i. e. to apply a tape-line of known length and count how many times it must be applied to reach the end. This, of course, is impossible. The largest distance that can be measured directly with measuring rods or tapes is one of a few miles. From that we have to climb up by means of triangulation first to longer distances along the surface of the earth, then, with the aid of astronomical observations, to the radius of the earth, and, finally, by means of observations of planets, to their distance from the earth. This latter distance is known in terms of the astronomical unit from the planetary theories. The whole sequence of successive steps is thus rather long, but for most astronomical purposes we need not bother about the longest part of it, the linking up of the astronomical and the terrestrial units. All we need is the distances of the stars in astronomical units. These are given by the last step, the trigonometric determination of the parallax.

In this method the base-line is, as already explained, the diameter of the earth's orbit, or to be more exact, the distance between two positions of the earth on its orbit, chosen as far apart as circumstances will permit. The earth's orbit being known, the length of this base-line is easily expressed in the astronomical unit. For the unit of stellar distances

we use the "parsec," an ugly word, but one which has unfortunately come to stay. It is the distance of a star of which the parallax is one second of arc, and it is a little over 200,000 astronomical units, or about three and a quarter lightyears. (A lightyear is very nearly ten million million kilometers.) By this method of ordinary trigonometric parallaxes, with our largest modern telescopes, we can reach distances of about five hundred lightyears. The parallax of a star at a distance of 500 lightyears corresponds on the photographic plates taken with our largest modern telescopes to an interval of about a third of a mikron. We can arrange our observations so as to measure nearly twice that distance, say one forty-thousandth of an inch. This is a good example of one of the characteristic features of practical astronomy: the most important results depend on the measurement of extremely minute quantities. As a consequence of the improvements in the construction of the telescope, and the application of photography, we can now by the method of trigonometric parallaxes measure distances about fifty times as large as those determined by Bessel and Struve nearly a century ago. It seems doubtful whether the accuracy attainable by this method is capable of being still further increased.

For larger distances we must look out for a triangle with a larger base-line. For this we can avail ourselves of the motion of the sun through space, carrying the earth with it. This base-line, i. e. the distance the sun has travelled from a certain initial time, increases proportionally to the time, and can thus be made arbitrarily large by waiting long enough. A century gives a base-line of over 400 astronomical units. As has been said already, the motion of the sun was de-

terminated for the first time by William Herschel, afterwards by Bessel, Airy, and others, and again by Kapteyn. The apparent displacement of the stars due to the motion of the sun is called their "secular" parallax. It is, of course, always in the same direction, whilst the "annual" parallax, due to the motion of the earth round the sun, is periodic. The method of secular parallaxes can, consequently, give no *individual* parallaxes, because of the proper motion of the stars. The stars do not stand still whilst their measure is being taken, but have a motion of their own. In the case of the annual parallax this does not matter: the proper motion is eliminated, because the earth returns after the lapse of a year to its initial position. But in the case of the secular parallax it is impossible to separate the parallactic motion, due to the motion of the sun, from the star's own motion. This method thus can only give *mean* parallaxes of a number of stars, the proper motion being as often one way as the other, and being cancelled out in the mean. But for the knowledge of the structure of the universe these means are just what we want, and more important than individual parallaxes.

With the present data, i. e. accurate observations since the time of Bradley, during about one and three-quarters of a century, we can penetrate into the universe to a few thousands of lightyears; thus by successive application of the method of triangulation we can climb up from a distance of a few miles, through the radius of the earth, the astronomical unit, the sun's motion through space, to one that is more than ten thousand million million times larger. For still larger distances the method of triangulation fails us and we

must have recourse to the second principle, that of the inverse square law. But we will defer the discussion of methods based on that principle until the time when we shall need to consider these larger distances.

When Kapteyn began, about 1890, his investigations concerning the structure of the stellar system, there were known about 130 directly measured parallaxes, fifteen of which he had determined himself. These were mostly of bright stars, the data for faint stars being then still extremely scanty. About forty years ago a star of the ninth magnitude was considered very faint. Today, of course, it is called a bright star. It is, in fact, due in great part to Kapteyn's influence that astronomers have since then attached so much importance to getting data for faint stars. In addition to these individual parallaxes his results are based in great part on secular parallaxes. In order to be able to determine these, an accurate knowledge of the sun's motion is required, and a great part of Kapteyn's energy and time have been given to the determination of the amount and the direction of that motion. This research has involved the development of several new methods, and criticism of previous results; it is the most exhaustive investigation of the subject that has ever been made. All the great problems of fundamental astronomy had to be reconsidered; Kapteyn's work included a new determination of the precessional constant, of systematic errors in the fundamental system of declinations, etc., and, as a by-product, it led to the discovery of the two star streams, about which more will have to be said presently.

All these investigations, and many others, of widely dif-

ferent nature, show Kapteyn's clear insight and command over the most heterogeneous parts of astronomical science. In every new domain that he enters what he achieves is new, carefully investigated in all details, and worked out with sound criticism. However widely various these researches may appear to be, in Kapteyn's mind they do not stand apart from each other, but are all parts of the great whole, the problem of the structure of the system of the fixed stars. However deeply he may be absorbed in the details of a special investigation, the great problem is always present to his mind and is never for a moment lost sight of. That is the secret of his success. In him one sees the rare combination into a harmonic whole of two intellectual faculties that by their very nature seem to be opposed to each other — and that are in reality indeed often conflicting, and seldom combined in the same person, — viz. the wide outlook and command of great problems on the one hand, on the other hand the painstaking devotion to technical details of observation and computation. Never was any labour too much for him if the purpose was to clear up a doubtful point, perhaps uninteresting in itself, if that was necessary for the whole, but also he has never let himself be tempted to follow a side path, however fascinating it might be, farther, or to work out a secondary result with more refinement or elegance, than was necessary for the solution of the main problem.

Every scientific problem has two sides, viz. the compilation of the data and the derivation from them of general results, "the grinding of huge masses of facts into law," as Darwin calls it. These two parts of the work require an entirely different attitude of mind and entirely different talents, and

accordingly in the history of science we see them generally performed by different individuals. The characteristic of Kapteyn's method of working is the parallel growth of these two aspects of the problem and its solution at the same time in incessant interaction. Again and again we see him applying a new method to the material available at the time, even though the data are perhaps not yet so complete as would be necessary, and the method has not yet been worked out in all details, with the double purpose, on the one hand of testing the efficiency of the method, on the other hand of getting indications of the direction in which the data require addition and completion.

In Kapteyn's published work several utterances are found that express this method of working very clearly. It is especially emphasised in the beautiful lecture held on the occasion of the opening of the Astronomical Laboratory at Groningen (1897), and again in one of his last publications (1920):

The progress of such a solution being necessarily slow, the present provisional treatment was undertaken mainly for the purpose of finding the most promising lines for conducting the further treatment of the sidereal problem.

Like William Herschel, Kapteyn has devoted his life to the investigation of the structure of the stellar system. The problem in its most general form is to find for any class of stars, e. g. all stars of a certain spectral class, or even for all stars together, the unknown function $D(x, y, z, i)$, which determines the density (i. e. the number of stars per unit of volume), for a certain absolute magnitude, at different places in space.

William Herschel devoted much time to the detailed structure, but still he kept always present to his mind the idea of constructing a *mean* image of the universe, showing the great lines of the architecture, all details being smoothed out.

Kapteyn deliberately dismisses the discussion of the detailed structure; what he sets out to find is what Seeliger calls "the typical image." He presupposes the symmetry with regard to the milky way, so that only two of the three coördinates are left: the distance r and the galactic latitude b . The strongly marked symmetry with regard to the plane of the milky way is the chief feature of the structure of the stellar system. Regarding it Kapteyn quotes Sir John Herschel, who knew the milky way as few others. In the volume of his observations at the Cape the latter writes:

It will of course be understood, and a very superficial inspection of the synopsis of gauges above given, suffices to show that great local departures from the general law of distribution indicated in this table occur, and *that* nowhere more remarkably than in the Milky Way itself, whose inequalities of breadth and structure are most conspicuous and singular. On this head I shall presently have more to say. But, with exception of these portions of the galaxy itself, I find nothing in the smallest degree meriting to be regarded as *systematic* in any part of the heavens which I have examined, as respects these deviations from perfect regularity. On the contrary, so purely local are they that on a careful revision of the whole chart, I find it difficult to specify any considerable areas over which an average density of stars prevails materially differing from what might be expected from the law of density (regarded as a function of the Galactic polar distance) indicated in the last article.

When he embarked on his work Kapteyn had not enough data to be able to solve even the problem of the "typical

image," and therefore restricted himself to the construction of the "schematic image," as Seeliger calls it, in which only one coördinate, the distance, is considered. Of course this "schematic image," in which the milky way is, so to say, smeared out equally over the whole sky, is no approximation to the truth, but it is a very nice problem to begin with, to try the working of the methods on. In this lecture I will also restrict myself to this simplified problem, on which all the essential points of the larger problem can be demonstrated quite satisfactorily.

Kapteyn's methods are entirely statistical; he deals with averages. Suppose for a moment that we knew the distances of all the millions of stars which constitute the galactic system. We should then be able to make an accurate model of the system. If we had this model, what would we do with it? Since we want to find out the great lines of the plan according to which it is built, its general architecture, we should not fix our attention on the individual stars, but should have to take averages; we would count the stars, in order to find the number per unit of volume, i. e. the density, at different points of space. It is thus not the individuals that are important, but the averages. Accordingly Kapteyn sets out to find the averages without going through the long and tedious (and also impossible) process of determining the individual values first. His methods are based on the laws of the theory of probability, the law of great numbers, or, as we might paradoxically say, the law of accidents. Paradoxically — since an accident is by definition a thing that obeys no law. But even because they obey no law, accidents are in the long run liable to cancel one another, and the law of great numbers is,

in fact, one of the most reliable laws of nature — and not only of inanimate nature, but of human society and economics as well. Take the example of an insurance company. It has many insured lives of men of a certain age. One of these may live to be a hundred, another may die tomorrow. It does not interest the company a bit who these individuals are; what it wants to know is the *average* expectancy of life of a man of a certain age. Its whole business is built on the law of averages, the law of great numbers. And the fact that insurance companies pay dividends proves that the law of great numbers is a reliable law.

Originally Kapteyn had developed a method for the determination of the function D , which looked theoretically perfect, but required an extensive mathematical apparatus. This apparatus was all prepared beforehand, but when it came to the application it proved to be entirely impossible to use it. The reason why it would not work was afterwards found to be the existence of the two star streams. Probably this experience has contributed largely to Kapteyn's preference for numerical over formal mathematical methods, and to his habit of going forward step by step, continually testing his methods as he developed them.

In his second method, Kapteyn assumed, in order to make the problem more tractable, that it is possible to separate the density law and the luminosity law, in other words that the distribution of the stars among the different magnitudes is independent of the distance, i. e. everywhere the same. Mathematically expressed, this means that the function $D(r, M)$ is the product of two functions, $\Delta(r)$, the density law, and $\Phi(M)$, the luminosity law. This hypothesis is made

only with a view of making the problem easier to tackle; it can afterwards be verified, and if it proves to be inexact, the necessary corrections can be applied.

William Herschel had in his first theory practically supposed that the density is constant; in the second theory he assumes the absolute magnitude M to be the same for all stars. Neither of these hypotheses, of course, is true.

Kapteyn's determination of the two functions Δ and Φ is free from any hypotheses, except the fundamental one that the law of great numbers is applicable. His method is essentially very simple, and can be explained almost without using any technical language. It depends on the previous determination of three other functions, viz:

- (1) the numbers $N(m, \mu)$, i. e. the total numbers of stars in the sky having the apparent magnitude m and the proper motion μ ,
- (2) the mean parallax $\bar{\pi}(m, \mu)$ for the stars of apparent magnitude m and proper motion μ ,
- (3) the function $\psi(\pi/\bar{\pi})$, the frequency function of the ratios of the true and the mean parallaxes.

We will say a few words of explanation about each of these.

(1) The numbers $N(m, \mu)$ are, of course, simply the result of *counting*. But this counting is not such a very simple matter. The first thing required, before we can begin to count, is a correct scale of magnitudes. In the scale generally adopted since about the middle of the nineteenth century the magnitude is proportional to the logarithm of the intensity. Thus a step of five magnitudes means a decrease of the intensity in the proportion of a hundred to one: a star of the

first magnitude has one hundred times the light of one of the sixth, and so on. Now the square root of a hundred is ten, so if the two stars had the same intrinsic brightness, the one that is apparently the fainter, the star of the sixth magnitude, would be at a distance ten times that of the star of the first magnitude. If the difference of apparent brightness were only one magnitude, the ratio of the intensities would be the fifth root of a hundred, or about 2.51, and the ratio of the distances would be the square root of this again, or 1.58. Thus, making the same computations the other way round, we can, if the distance and the apparent magnitude are known, calculate the magnitude that the star would have if it were placed at any other distance, and especially at the unit of distance. For this unit Kapteyn has always used a distance of ten parsecs. We will call this the Kapteyn unit. Its value is a little over two million astronomical units, or 32.6 lightyears. The apparent magnitude that any particular star would have if placed at this unit of distance is called its "absolute magnitude," and it is a measure of the star's intrinsic brightness or "luminosity," in the same way as the apparent magnitude is a measure of its apparent brightness. The words "luminosity," "absolute magnitude," like so many others that are now household words in astronomy, were first introduced by Kapteyn, when he felt the need of replacing the more or less vague terminology that was in use when he began his work by accurate definitions.

Kapteyn has devoted much care to reducing the empirical scales of magnitude of all the different catalogues of stars to one standard scale, and to determining the limiting magnitudes of the telescopes of William and John Herschel; of the

gauges of John Herschel he has made an extensive use, in preference to those of the elder Herschel, because they were distributed equally over the sky according to a definite scheme, without giving preference to specially interesting regions, as those of William. It is, of course, impossible to count *all* stars; there are too many of them. We must take samples, but for the application of the law of great numbers it is necessary that the samples be selected entirely at random, without any possible bias for either rich, or poor, or otherwise exceptional regions.

Besides the apparent magnitudes, we must have the proper motions of the stars. These are measured in seconds of arc per century. Like the magnitudes, the proper motions are *apparent* quantities, depending on the distance. If one star is at twice the distance of another, the components of their real velocities perpendicular to the line of sight being the same, then the proper motion of the more distant star will be only one-half of that of the other. Thus, if all stars had the same real velocities, and if these had no preference for any special direction, so that in the average of a large number of stars the component perpendicular to the line of sight would also be the same, then the proper motion would be a measure of the distance; it would, in fact, be proportional to the parallax. Since the stars have not all the same real velocity, any more than they have all the same absolute magnitude, neither the proper motion nor the apparent magnitude can be used as a direct measure of the distance. But it is clear that statistically the distribution of the apparent magnitudes and the apparent proper motions must depend on the distribution of the distances, i. e. on the density law. That is

why Kapteyn's investigations start with the counting of the numbers $N(m, \mu)$.

For the proper motions Kapteyn's work was at first based entirely on the results derived by Auwers from the observations made by Bradley (mean epoch 1755), and from the Greenwich Catalogues of 1860 and 1864 (mean epoch 1863). The mean interval is therefore 108 years. The number of proper motions in Bradley's catalogue, as edited by Auwers, is about 3600. For the fainter stars, of which no proper motions are known, of course the numbers $N(m, \mu)$ cannot be determined, and we have to be content with using the numbers $N(m)$ of stars of apparent magnitude m , irrespective of their proper motion.

(2) The second datum on which Kapteyn's researches are based is the mean parallax $\bar{\pi}(m, \mu)$ of a star of apparent magnitude m and proper motion μ . When Kapteyn began his work there were, as has already been mentioned, only about 130 measured parallaxes available. The determination of the function $\bar{\pi}$ depends on these 130 individual stars, and on the secular parallaxes, derived from mean proper motions with the aid of the solar motion, which, as related before, was determined anew by Kapteyn. It is wonderful to reflect how scanty the data were on which Kapteyn based his theory. Only a man of great courage and self-confidence could even think of building such a large superstructure on so feeble foundations. But *fortes fortuna juvat*, and it is most remarkable to see how well even now Kapteyn's first mean parallaxes represent the enormously richer and more accurate observational data of the present time. In later years, as the observational data accumulated, the determination was of course

repeated and improved. Kapteyn gives a definite formula for the mean parallax $\bar{\pi}(m, \mu)$ as a function of m and μ , from which for every magnitude and proper motion the mean parallax can be computed.

(3) A star of given parallax and proper motion has the *mean* parallax given by the formula for $\bar{\pi}$; thus, e. g., if we take a large number of stars, each of magnitude 6.1, and centennial proper motion $4''.5$, then the mean of their parallaxes will be $0''.0102$, corresponding to a distance of 9.8 Kap-

TABLE I
DISTRIBUTION OF TRUE DISTANCES

Limits of ratio true/ average	Number out of 100
Smaller than $\frac{1}{4}$	1
$\frac{1}{4}$ to $\frac{1}{2}$	$7\frac{1}{2}$
$\frac{1}{2}$ to 1	29
1 to 2	40
2 to 4	19
Exceeding 4	$3\frac{1}{2}$

teyn units. But the *individual* distances will all be very different from this average, because the stars have not all the same absolute magnitude, nor the same real velocity. Some of the distances will be larger than the average and some will be smaller. The function ψ gives the frequency of the ratio of the real to the average distances. The function adopted by Kapteyn is derived from a comparison of the known individual parallaxes with the formula for the mean parallax $\bar{\pi}(m, \mu)$. It is again given by a special mathematical formula, from which we have computed the small Table I. Thus, out of a hundred stars, one is at a distance less than a quarter

of the average; of forty the distance is between the average and twice the average, and so on.

We now have collected all the data, and can proceed to explain the use Kapteyn makes of them. His method is very simple. He begins by making a table of the $N(m, \mu)$. Next to each of these numbers can be written the mean parallax $\bar{\pi}(m, \mu)$ or the corresponding average distance, computed by the formula. A sample of this table is given below as Table II.

TABLE II
NUMBERS OF STARS AND AVERAGE DISTANCES

Centennial proper motion	Apparent magnitude			
	3.5 — 4.5	4.5 — 5.5	5.5 — 6.5	6.5 — 7.5
1".0–2".0	52 (17.5)	194 (19.6)	638 (21.7)	1896 (23.8)
2".0–3".0	41 (12.0)	177 (13.5)	595 (14.9)	1910 (16.4)
3".0–4".0	27 (9.5)	188 (10.5)	542 (11.6)	1910 (13.0)
4".0–5".0	27 (8.0)	93 (8.8)	461 (9.8)	1490 (10.8)
5".0–6".0	22 (6.9)	86 (7.7)	357 (8.6)	1249 (9.4)
6".0–7".0	25 (6.1)	77 (6.8)	252 (7.4)	963 (8.3)
7".0–8".0	18 (5.6)	71 (6.1)	247 (6.8)	752 (7.5)

The table is one of double entry, the arguments being magnitude and proper motion. The numbers given in ordinary figures are the $N(m, \mu)$, those in italics are the corresponding average distances. Thus, e. g., there are in the sky 461 stars of which the magnitude is between 5.5 and 6.5 and the proper motion between 4".0 and 5".0, and the average distance of these stars is 9.8 Kapteyn units.

Now Kapteyn divides the whole of space into different shells, limited by spheres at such distances that the ratio of the distances of the inner and outer limits of a shell corre-

sponds to a difference of one magnitude. Thus, as has been explained above, the ratio between the inner and outer boundary of each shell is 1.58, and if we move out five shells the distance is multiplied by ten. The first shell (or rather the core) is taken from the centre to the sphere of radius, 0.63 Kapteyn units, the second shell thus ranges from 0.63 to one unit, and so on, the seventh extends from 6.3 to 10 units, and the twelfth from 63 to 100.

TABLE III
DISTRIBUTION OVER SHELLS

Shell	Radii of boundaries		5461 star $m = 6.1$ $\mu = 4''.5$ (aver. dist. 9.8)	Volume	Absolute magnitude
I	0	0.63	} 0.4	1.0	7.45
II	0.63	1.00		3.1	6.45
III	1.00	1.58		12.5	5.45
IV	1.58	2.51		49.7	4.45
V	2.51	3.98	16.5	198	3.45
VI	3.98	6.31	54	788	2.45
VII	6.31	10.00	104	3137	1.45
VIII	10.00	15.85	127	12490	0.45
IX	15.85	25.12	96	49730	-0.55
X	25.12	39.81	45	198000	-1.55
XI	39.81	63.10	13	787600	-2.55
XII	63.10	100.00	2	3137000	-3.55

If the distance of each star were equal to its average distance, it would be easy to assign to each star its place: the first 52 stars of our Table II (with $\mu = 1''.5$, $m = 4.0$) could be put at once into shell IX, the next 41 into shell VIII, the 27 with $\mu = 3''.5$ and the 27 with $\mu = 4''.5$ into shell VII, and so on.

This would already give us an idea of the distribution, but

that idea would be very inaccurate. It is essential to take account of the fact that the distances given in Table II are only averages, and that the true distances are widely spread; moreover, they are not even distributed symmetrically with respect to the average. We must thus use the formula for the frequency of the parallaxes, of which Table I is an abridged representation.

As an example consider the 461 stars with $\mu = 4''.5$ and $m = 6.1$, of which the average distance is 9.8 Kapteyn units. When the formula for ψ is applied it is found that they are distributed as shown in Table III.

In this way all the numbers in Table II are treated, and after that the total number of stars of a given apparent magnitude contained in each shell is found by a simple addition. We now know the number of stars with a given apparent magnitude in each shell, but the individuality of the stars has been lost.

As a consequence of the way in which the limits of the shells have been chosen, the apparent magnitudes can at once be converted into absolute ones. Thus we know the absolute magnitude of each star, and the foregoing table can be made into one that gives the number of stars of each absolute magnitude per shell. At the same time we can divide by the volume of the shell, and thus find the number of stars of given absolute magnitude per unit of volume at every distance.

In this way Table IV is derived. The part of the table inclosed between the heavy lines is supposed to be fairly reliable. The fringe outside is considered by Kapteyn to be of less accuracy.

We can take the density of any one shell, say the fifth, as our unit. Then dividing the number of stars in any other shell by the number in shell V, we find the density of that particular shell. Each vertical column of the table thus gives a separate determination of the density law. The mean of

TABLE IV
NUMBERS OF STARS PER 1000 UNITS OF VOLUME

Shell	Absolute magnitudes							Density
	-2.55	-1.55	-0.55	0.45	1.45	2.45	3.45	
V	...	10	33	89	299	719	1860	1.000
VI	...	9	28	94	265	780	1970	1.031
VII	2	7.5	24	79	265	766	1680	0.985
VIII	1.6	5.2	20.4	72	232	590	1680	0.869
IX	0.98	4.25	15.5	53	160	520		0.656
X	0.71	2.68	9.8	35	130	...		0.418
XI	0.39	1.54	6.9	28	...	...		0.234
XII	0.21	1.2	5.2	...	...	...		0.122

the densities thus derived is given in the last column of the table. Similarly every horizontal line gives a determination of the luminosity function $\Phi(M)$. Since all these different determinations are found to give practically the same result, the hypothesis, originally made for simplicity's sake, that the function $D(r, M)$ can be split up into Φ and Δ is proved to be sufficiently accurate.

Kapteyn has applied this method over and over again on new data; at first he used all stars together; afterwards he was able to do it with more refinement by considering separately

TABLE V
KAPTEYN'S DENSITY LAW (1920)

Distance		Number of stars per cubic lightcentury	
Kapteyn units	Lightyears	In milky way	Towards pole
0	0	1300	1300
10	326	1300	716
25	818	1300	520
63	2060	750	160
158	5160	270	21
398	13000	65	1½
1000	32600	10	...

TABLE VI
KAPTEYN'S LUMINOSITY LAW (1920)

(The unit of luminosity is a star of magnitude 0 at the distance of one lightyear)

Out of 1000 stars				
1	has	a luminosity exceeding	1000	
18	have	"	between 100 and 1000	
122	"	"	"	10 " 100
328	"	"	"	1 " 10
352	"	"	"	0.1 " 1
151	"	"	"	0.01 " 0.1
27	"	"	smaller than 0.01	

the different galactic latitudes, and later, when more data became available, even the different spectral types.

The results as given in Groningen Publication 11 (1901) are: the density is nearly constant as far as the sixth shell;

after that it decreases, and is about two-thirds of its initial value in the ninth shell, one-quarter in the eleventh, and about one-eighth in the twelfth. In later determinations, of course, the density law is different for different galactic latitudes.

Tables I to IV are taken from Kapteyn's first treatment of the problem, published in 1901. They have been given

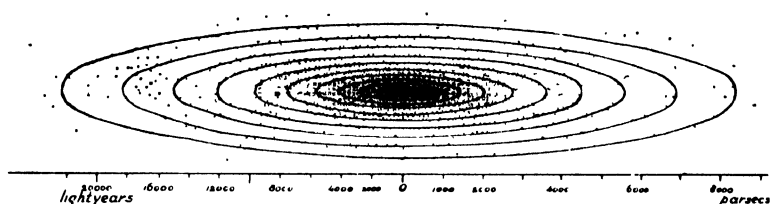


FIGURE 5

The Kapteyn system (1922)

here only as an illustration of the method. The results derived in 1920 are given in Tables V and VI.

It will be seen that the density decreases rapidly in the direction towards the poles of the milky way, and much more slowly in the plane of the milky way itself. The luminosity curve has a maximum in the neighbourhood of the luminosity that is here chosen as unit, which is about one-third of the luminosity of the sun. Later investigations have thrown doubt on the existence of this maximum. Faint stars are probably more numerous than Kapteyn's work implies.

The image of the structure of the galactic system that Kapteyn had formed towards the end of his life is given in Figure 5. The determination of the density is continued up to one-hundredth of the initial value; this is reached at a

distance of 8465 parsecs in the galactic plane and of 1660 parsecs in the direction perpendicular to it. The figure is a section through the system perpendicular to the galactic plane, which is a plane of symmetry. The ellipses are sections of the surfaces of equal density, which are ellipsoids of revolution. The number of black points denotes the density: it is 100 in the centre as against 1 on the outer ellipsoid. The sun is no longer supposed to be in the centre, but at a distance of about 65 Kapteyn units away from it, and 3.8 units to the north of the plane of symmetry. The centre of the milky way, as seen from the sun, is assumed to be at galactic longitude 77° .

In the last publication written by Kapteyn, he gives dynamical considerations about the galactic system, ascribing the flattening to rotation and computing, from a comparison of the decrease in density in the galactic plane with that in a direction perpendicular to it, the gravitational field and the velocity of rotation necessary to cause the observed flattening. The linear velocity at a distance of about 700 parsecs from the centre, where the sun is, would come out at about 20 km/sec. Kapteyn thus gives the following representation: the galactic system consists of two separate systems that penetrate each other and rotate in opposite directions, and in this way he explains the observed relative velocity of 40 km/sec of the two star streams with regard to each other. The direction of the streams must therefore be at right angles to the direction towards the centre of the system, and from this has been determined the longitude of 77° , mentioned above, for that centre.

Of course Kapteyn, as has already been pointed out several times, has in the course of his thorough investigations obtained many additional results. Of these the most important is certainly the discovery of the two star streams, which was first communicated by him to the International Congress of Science at St. Louis in 1904, but only became generally known when he had laid it before the British Association in Cape Town in 1905. It was a particularly momentous discovery, which has had an enormous effect on the development of astronomy. Before 1905 Kapteyn worked nearly alone in the field of stellar astronomy; after that date the number of investigators in that field was suddenly enormously increased. The work done, and the conclusions arrived at, up to 1914 are admirably set forth in Eddington's excellent book, "Stellar Movements and the Structure of the Universe." During the last fifteen years the study of the star streams has not occupied that prominent position in the work of astronomers that was still assigned it in 1914. They may not be such a dominating feature of the system as was generally supposed immediately after their discovery. Still the discovery of the star streams is of momentous importance, not only on account of the great impetus it has given to the study of stellar astronomy, but also as the first example of a regularity in the motion of the fixed stars.

During the last years new observational data have come to light, that have thrown doubt on some of Kapteyn's conclusions, and have induced us to assign to Kapteyn's system no longer that central position in the universe, or rather in the galactic system, that was, in the time when it was constructed, generally ascribed to it.

Even though the image of the galactic system, such as Kapteyn finally worked it out as a result of his investigations, is to be considered nowadays already obsolete, still Kapteyn is one of the most important personalities in the whole history of astronomy. His principal achievement is the creation of statistical astronomy. Kapteyn's whole life has been devoted to the problem of the structure of the stellar system, and by his wonderful mastership in the discussion of large amounts of observational data he has been able to make enormous progress, and to inaugurate an entirely new epoch in the study of stellar astronomy.

The use of statistical methods, introduced by Kapteyn, combined with the enormous increase of the observational output of our instruments as a consequence of photography and the big telescopes, makes it more necessary than formerly to concentrate our efforts. It is essential to let the choice of the observations, both the objects to be observed and the special observations to be made on them, be guided by theory. Kapteyn's "Plan of Selected Areas" is a grandly conceived scheme of coördinated observations that must be brought together by international coöperation and have, indeed, already been brought together for a great part.

Kapteyn's work has put its mark upon the whole of modern astronomy; many of the most useful concepts, that we use every day, and have become so accustomed to that we are apt to think they have always been there, were actually introduced for the first time by Kapteyn; thus, e. g., the colour index, the absolute magnitude, and others.

And yet nothing has been said about the huge observational work, the compilation of observational data, that

Kapteyn has performed, the thousands of parallaxes and proper motions determined in the Groningen Laboratory, and especially the Cape Photographic Durchmusterung, that in itself would be enough to be the life-work of an astronomer and to secure for him an honourable position amongst the greatest of his day.

V

PHOTOGRAPHY AND SPECTROSCOPY · THE GREAT TELESCOPES · THE GALACTIC AND THE EXTRA-GALACTIC SYSTEMS

THE essential difference between the work of Kapteyn and that of Herschel more than a century earlier is due to the enormous increase of the observational data in the meantime.

Also Kapteyn's own work and the great influence of his methods on the development of science have been largely instrumental in emphasising one of the characteristics of modern astronomy, the collecting of data referring to ever fainter stars, and especially the development of methods of determining the distances of faint and far-away objects. We shall have to come back to this point later on.

In the last third of the nineteenth century two new methods of observation were developed which have made possible this large increase of data, viz. photography and spectroscopy.

Photography has not only multiplied the observer's productivity by a large factor, but it has also made the observations more impersonal. The result is fixed on the photographic plate, and it is always possible to return to it and verify or amend the conclusions.

It is not necessary to enter into the details, but it is a fact that nowadays we can hardly imagine astronomy without

photography. I need not recall to your mind all the important discoveries made with its aid. And not only have these discoveries widened the field of astronomy, but also the old fundamental astronomy has profited largely by the remarkable properties of the photographic film. Photographic methods are specially suited for the measurement of proper motions and parallaxes; greater precision is obtained in less time than with the old methods. Also for the determination of positions photography begins nowadays to give results.

The second new method is spectroscopy; this has opened up an entirely new world, and has made things possible that half a century or two-thirds of a century ago nobody could dream of. At first, after the discovery of the origin of the spectral lines by Kirchhoff, the chief point of interest was the investigation of the chemical constitution of the atmospheres of the stars, the "chemistry of the stars." Presently it was found that it is possible to arrange the stars according to their spectra in one continuous series of types, and even so accurately that eventually decimals of a spectral type had to be used. Of course it was very soon assumed that this order was the order of evolution. Even apart from this, and even if this simplistic view has had to be abandoned or greatly modified, the fact that this arrangement in a continuous sequence is possible remains one of the most important assets in the study of the universe.

We have already spoken about the discoveries of Kapteyn, who, as a pioneer in this field, was the first to find and study statistical relations between spectral type and proper motion, absolute magnitude, distribution in space, etc.

From the spectroscopy of the stars an entirely new kind of physics has gradually arisen during the last fifteen years. There is a strong interaction between astronomy and physics; astronomy supplies the physicist with many kinds of extreme circumstances that he cannot realise in his laboratory: enormous distances, long times, high densities, low densities, high temperatures. Physics in return gives the astronomer much valuable help. The new science of astrophysics has become something entirely different from what used to be called physical astronomy half a century ago. It is outside the scope of these lectures to enter into the details of this branch of astronomy; still I will not omit to point out the important researches about the internal constitution of the stars by Eddington, Milne, and so many others. The modern theories regarding the constitution of the stars and the processes that keep their radiation going are singularly interwoven with the theories of atomic processes. Modern physics would be impossible without astrophysics. The intimate interconnection between them is typically exemplified in the title of Sir Arthur Eddington's delightful little book, "Stars and Atoms," of 1927, which is representative of the attitude of mind of the physicist and the astrophysicist of today.

With this development is closely connected the evolution in scientific thought, which leads us more and more away from Newton's rigorous standpoint allowing only pure generalisation and induction and "framing no hypotheses," towards a much more dominating position of hypothesis in modern science.

Spectroscopy has, however, done even more. Besides

creating the possibility of the new science of astrophysics, it has also opened up a new domain of astronomy of precision. By the application of Doppler's principle, first put into practice by Huggins in 1868 and Vogel in 1871, the radial components of the velocities of the stars, the rate at which they are receding from us or approaching us, can now be measured with great precision, whatever the distance of the observed star may be (so long, of course, as it is not too faint for the spectrum to be observed).

The reaping of the harvest of all these new methods has even surpassed our expectations, in consequence of the development of optical and mechanical technique, and of the great telescopes. Herschel's forty-foot telescope (aperture 48 inches, completed in 1789) was, properly speaking, too big for the mechanical possibilities of that time. It was extremely cumbrous to handle, and even an enthusiastic observer like Herschel could seldom work with it. Thus, up to the middle of the nineteenth century, the size of Herschel's twenty-foot telescope (aperture 19 inches), with which practically the whole work of William and John Herschel has been done, remained the standard size.

In February 1845 the 72-inch reflector of Lord Rosse, at Parsonstown in Ireland, was completed, and as early as April of that year the first spiral nebula was discovered with it. It was mounted in the meridian, had only a small sideways possibility of motion, and was not convenient in practice. The enormous advance in mechanical engineering can hardly be better appreciated than by comparing the mounting of Lord Rosse's telescope with that of the modern reflector of the same aperture at the observatory at Victoria,

British Columbia, built in 1918. For more than seventy years Lord Rosse's telescope remained the largest existing. It was only superseded by the 100-inch Hooker telescope of the Mount Wilson Observatory, completed in 1918, and perhaps before long the 200-inch will be constructed.

The great advantage of the large mirrors, of course, is that, by the large amount of light collected, the mass of observational data is very much increased, much fainter objects can be observed, and investigations about the structure of the milky way, continued in the way of William Herschel, can be pushed much farther. Also by the aid of photography more details are brought out, and the structure can be studied much more exhaustively than by visual observations.

As has been pointed out already, the difference between Herschel and Kapteyn lies in the fact that the latter disposed of good determinations of distances, whilst Herschel did not know with certainty the distance of one single star. Since the completion of Kapteyn's system new methods of determining distances have been found, viz. that of spectroscopic parallaxes and the determination of distance from cepheids.

The method of spectroscopic parallaxes was discovered in 1914 by Adams and Kohlschütter. They found that from certain characteristics of the spectrum, by the comparison of the intensities of certain lines, the intrinsic brightness of the star could be derived. Thus, the apparent brightness being also known, the distance can be determined.

The measuring of distances by means of cepheids is based on the same simple property of light, viz. that the intensity is inversely proportional to the square of the distance. Cepheids are a special kind of variable stars, very regular and

having a characteristic light curve, which makes their identification easy and unmistakable. In 1912 Miss Leavitt, of the Harvard Observatory, found an empirical relation between the periods and the luminosities of these stars, so that, if the period is known, the medium absolute brightness can be read from the curve. This method of deriving distances was used for the first time by Hertzsprung in 1914 for the determination of the parallax of the small Magellanic Cloud. Bailey had in 1900 discovered many variables, called cluster-variables, in some globular clusters. These are probably not true cepheids, but they fit into the period-luminosity curve. Consequently the same method could be extended to the measurement of the distances of globular clusters. The weak point is the determination of the zero of the scale. That must be found from the known distances of cepheids, but as most cepheids have a very small parallax, the direct determination of their distance is very uncertain, and it must be found from proper motions. This, however, brings in all the difficulties of fundamental astronomy, such as systematic errors in declination, etc.

Having thus, very briefly and superficially, described the new tools which modern research has at its disposal, we proceed to review the principal results that have been obtained.

The most notable progress in our ideas concerning the size and the structure of the stellar system, immediately after Kapteyn's investigations, is due to Shapley's researches, of which the results are described in a series of papers called "Studies based on colours and magnitudes in stellar clusters," published from 1915 to 1920. This work was done entirely with the 60-inch telescope of the Mount Wilson Observa-

tory. The chief results are the determination of the distances and the distribution in space of the globular clusters, and the indication of the connection of the system of the globular clusters with the galactic system. Shapley's classical series of papers in the Mount Wilson Contributions was completed in 1920, but the principal conclusions are con-

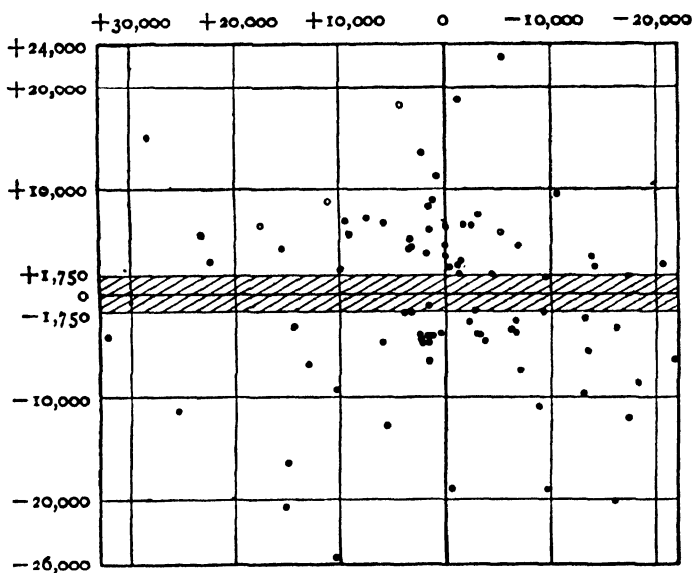


FIGURE 6

The globular clusters and the milky way, according to H. Shapley, 1918

tained in number XII, of April 1918, called "Remarks on the arrangement of the sidereal universe," and in paper XIV, "Further remarks on the structure of the galactic system," published in November 1918. The diagram, Figure 6, is taken from the latter paper. It represents the system of the globular clusters as it results from Shapley's work. Subsequent investigations have made some changes in the details,

but the great lines have not been altered. The globular clusters, of which there are about one hundred, are arranged symmetrically with regard to the plane of the milky way. Roughly speaking, they fill a sphere and the distances on both sides of the plane of symmetry are about 15000 parsecs. (We now quote the dimensions as at present most generally adopted, not Shapley's original ones of 1918.) In that sphere with radius 15000 ps the distribution of the globular clusters is approximately uniform, with a slight condensation towards the centre. In a space between two parallel planes on both sides of the plane of symmetry, which is shaded in the diagram, there are no, or very few, globular clusters, at least if there are we do not see them. We will return to this point later on. In that same central strip, of course, are situated the clouds of the milky way.

It is most remarkable to compare Shapley's diagram of 1918 (Fig. 6) with that representing William Herschel's last ideas, published in 1817 (Fig. 4). In both cases we have the central disc, containing the milky way clouds, and extending indefinitely in the direction of the central plane, and on both sides of this central disc is the "region of the globular clusters." And not less remarkable is the comparison of Herschel's first theory of 1785 (Fig. 3) with the Kapteyn system (Fig. 5). Kapteyn's results were essentially complete about 1900. Thus we see history repeating itself after a century. About the beginning of the nineteenth century we have the two theories of William Herschel, published at an interval of about thirty years, and in the first years of the present century we have two very similar theories, at a somewhat shorter interval, but this time by two different investigators.

Of course the modern editions depend on much more complete observational data, and more rigorous discussion, while Herschel's results were based mostly on assumption and reasoning from analogy. But this only serves to increase our admiration for the wonderful intuition of this great man, and it is certainly most remarkable to find the same evolution of ideas that took place in his mind repeated after a century in what we might call the collective mind of the astronomical world. About the year 1800 the astronomical world, at least so far as sidereal astronomy is considered, consisted of William Herschel alone.

Within the central disc of Shapley's diagram are situated, as already mentioned, the milky way clouds, of which the Kapteyn system may be one. The centre of the Kapteyn system lies at a distance of about 10000 ps from the centre of the system of globular clusters, which is also the centre of the whole galactic system, while according to the most generally adopted views the density decreases when we recede from the centre of the Kapteyn system, not only in the direction perpendicular to the galactic plane but also in that plane. It may then increase again as we come to other clouds. Such is about the image that we formed of the galactic system after Shapley's work, say about ten years ago.

On April 26, 1920, took place the homeric fight on the scale of the universe, the discussion between Shapley and Curtis, in which Shapley defended the scale of his new model of the galactic system, depending chiefly on the observed distances of the globular clusters, against the contrary view of Curtis, who favoured a scale about ten times smaller. In this discussion reference is also made to what is known under

the name of the "theory of island universes," though it played only a very subordinate part in the argument. The idea of this theory, viz. that our galactic system is a nebula and that some of the nebulae that we see in the sky are systems of the same size and nature as our own, is, of course, as old as William Herschel. The spiral structure of the galactic system was, as far as I know, first suggested by Easton in 1900, and its analogy with the spiral nebulae has been of great influence on our ideas regarding the structure of the galactic system.

Easton's diagram, which was constructed by him as an illustration and explanation of his theory, is reproduced in Fig. 7. In the border is an accurate photometric map of the milky way as we see it: we must think of the milky way as projected on the inside of a cylindrical surface, and then the cylinder is bent out flat. Of course the picture is a negative; the darkest shades represent the brightest parts in the sky. In the middle is a schematic picture of a spiral. It is not intended to represent any particular nebula, but it looks very much like, say, M. 33 or M. 101 (see Plate VIII). The sun is supposed to be at S, and lines of sight are drawn to show that what we should see from this position agrees in general character with the actual milky way as we see it in the sky. Easton explains the two branches of the milky way by supposing the two arms of the spiral to be not exactly in the same plane. This is now known to be inexact. Also the direction of the centre in Easton's diagram is in the constellation of Cygnus, about at right angles to where we now know it to be. So not much remains of Easton's theory beyond the general idea that the galactic system is of spiral structure.

PLATE VIII



THE SPIRAL NEBULA MESSIER 101

In the ten years elapsed since the completion of Shapley's work on the globular clusters three new points of view have come to the front, which have been of great importance for the development of our conceptions of the galactic system.

The distribution of the stars in the milky way is very irregular. There are many darker spaces between the clouds, and there are dark lanes and patches without stars across and right in the middle of them. Some of the most noted examples are in the constellation of Ophiuchus (Plate XI). William Herschel used to call these "holes in the sky," and it has long been thought that they were actually places where there are no stars. The conviction that they are, on the contrary, places where there is something preventing us from seeing the stars has only slowly dawned upon us. It began by Barnard's discovery of "dark nebulae." Then, by the investigations of Pannekoek and others it has become certain that in the galactic plane there are very extensive clouds of obscuring matter, either gaseous, non-luminous, nebulae, or clouds of small particles or dust.

The analogy with many spiral nebulae seen edgewise, and showing a dark band running along the whole length of their equatorial planes, such as, e. g., N. G. C. 4565 (Plate XII,) tends to strengthen both our belief in the similarity of our own system to the spirals and our belief in the existence of this zone of obscuring matter all around the galactic plane. This same absorbing layer explains why we do not see any globular clusters in the milky way, nor any extra-galactic (spiral) nebulae. Another very important theoretical conclusion to be drawn from the existence of this obscuring band is that the decrease of density with increasing distance, as

derived by Kapteyn's theory, may be only apparent and due in great part, or perhaps entirely, to absorption. An absorption of one magnitude per thousand parsecs would be sufficient to explain the apparent decrease of density derived by Kapteyn from the observed numbers of stars for different magnitudes, the actual density along the galactic plane being constant. In the direction perpendicular to the galactic plane it is almost certain that there is no appreciable absorption. In that direction therefore the decrease of density as determined by Kapteyn is probably real. Taking the thickness of the central segment to be 2000 ps, the density at its limits would then be about one-fortieth of its value in our neighbourhood. In the direction of the galactic plane nothing can be said with certainty; we do not know whether the density decreases, remains constant, or perhaps increases in some directions, and the only certain thing we can say is the same as the conclusion William Herschel arrived at at the end of his life: that the milky way is unfathomable.

This need not, however, always remain so, and even now this pessimistic view is too drastic in its generality. As star counts are being extended to fainter and fainter stars it is not impossible to separate the effects of a real decrease in density from those of absorption, and, by investigations recently published by Mr. Bok of the Harvard Observatory, it has already been possible not only to delineate approximately the limits of our own "local system," the particular star cloud that we are in, but also to determine where the next star cloud begins, after an intervening region of small density. It is particularly gratifying to note that it is essentially by Kapteyn's own methods that these results are reached.



SOUTH-PRECEDING PORTION OF THE GREAT SPIRAL NEBULA
IN ANDROMEDA

PLATE X



STAR CLOUDS IN SAGITTARIUS

Similar researches, with similar results, have recently been made by Dr. Seares at the Mount Wilson Observatory. This is, of course, not the place for a critical discussion of these recent researches. We cannot do more than mention them, and draw the conclusion that, notwithstanding the difficulties resulting from the absorption, Kapteyn's statistical methods will be found applicable, and will give very valuable results, far beyond the limits of the Kapteyn system.

Still the existence of the absorbing matter sets a natural limit to the application of these methods, though the limit may be farther away than was thought when it was first realised that Kapteyn's conclusions, based on the hypothesis of no absorption, could not be accepted unaltered. When we come to distances where the absorption is so strong that we can no longer see the stars, there is an end to the counting, and we do not know how much there is beyond. Still the galactic system must be finite. Its total extent, however, must be found by other means.

The second important new fact of the last ten years is the rotation of the galaxy, discovered by Lindblad and Oort.

The relative velocity of the centre of gravity of the stars near the sun with reference to the system of globular clusters is about 300 km/sec in a direction perpendicular to the direction towards the centre, and in the plane of the milky way. Now if this is to be interpreted as a rotation there must be an effect both on the proper motions and on the radial velocities. The diagram (Fig. 8) has been constructed to explain this. The sun and its immediate neighbours are situated at the point S. At the points A, B, . . . H are other groups of stars, all at the same distance from the sun. Lines

are drawn from each of these points towards the centre of rotation, which is outside the diagram, at the bottom. The velocity of rotation at each of these points is indicated by a heavy line, with an arrow point to show the direction, such as Aa in Fig. 9, which is a part of Fig. 8, drawn on a larger

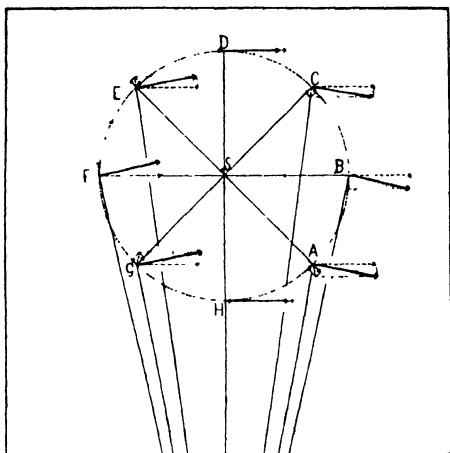


FIGURE 8

Diagram illustrating the rotation of the
galactic system

scale. It is, of course, perpendicular to the direction towards the centre. Then at each of the eight points a broken arrow-pointed line like Ab is drawn equal and parallel to the velocity of the sun. The line Ac, completing the parallelogram Acab, is the relative motion of A with reference to S. This is what we observe. The relative velocity Ac is then at each of the eight points resolved in its two components Ad perpendicular to, and Ae along the line of sight. The first is observed as a systematic angular proper motion, the other



THE MILKY WAY IN OPHIUCHUS

as a radial velocity. As is evident from the diagram, there is an effect of the rotation both in the transverse proper motions and in the radial velocities. If the whole system rotated as a solid body there would evidently be no change of the distances and no systematic radial velocities due to the rotation, but if, as in the figure, the velocity of rotation changes with the distance, the relative radial velocity will show a double period: it is zero at H, outwards at A, zero again at

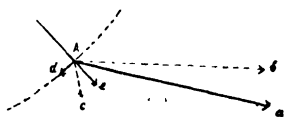


FIGURE 9

Detail of Figure 8 on enlarged scale

B, inwards at C, and so on. There is a similar effect in the proper motions, but that can only be observed in its angular projection, and it is very small on account of the great distance. The radial velocity is observed as such spectrographically, the observation is independent of the distance, and the effect itself increases with the distance. It is thus much more favourable for the detection of the rotation than the transverse component. The observed rotational effect in the radial velocities (and the proper motions) gives us the data for the determination of the ratio V/R between the velocity and the distance to the centre, and the rate of change dV/dR , V being the rotational velocity at the distance R from the centre. From these data and the observed velocity V we can derive the distance to the centre R as well as the centripetal force and the way in which this force depends



THE SPIRAL NEBULA N. G. C. 4565

is not represented in the diagram, but it is there, and prevents us from seeing the middle part of the nucleus (and all that is beyond it) and thus causes the appearance of the two branches of the milky way. It will be noticed that Oort's diagram is very much like a spiral seen edgewise, such as, e. g., N. G. C. 4565 (Plate XII).

This theory of the rotation of the galactic system provides the explanation for several phenomena, which had been found from observations, but could not be explained before. In the first place the stars of large velocity. It was discovered by Oort in 1922 that there is a sharp distinction between stars having a velocity smaller than 63 km/sec and those with a velocity larger than this, relatively to the centre of gravity of the stars in the neighbourhood of the sun. The small velocities are distributed entirely at random; the large velocities are all directed to one hemisphere. The centre of this hemisphere towards which the large velocities are directed is situated in the plane of the milky way in the direction ninety degrees from the centre of the system of the globular clusters. Thus, if the velocities are referred to the centre of the globular clusters, which on the theory of rotation is the centre of rotation, the stars with large velocity are the stars of which the velocity of rotation is smaller than the average, in other words those that are left behind by the sun and its neighbours of the Kapteyn system as they speed on in their rotation round the centre. There are no stars of which the velocity of rotation exceeds the average of 300 km/sec by more than 60 or 63 km/sec. The explanation which suggests itself, viz. that this velocity of about 360 km/sec is the velocity of escape from the galactic system, is in numerical

agreement with the distance to the centre and the mass as derived above. A star which has a velocity relatively to the centre of gravity of the stars near the sun larger than 63 km/sec in the direction of the rotation would escape from the system.

Another phenomenon of which the theory of the rotation gives a dynamical interpretation is that of the two star streams, discovered by Kapteyn in 1904. Notwithstanding the immense interest aroused by this discovery, and the large amount of study devoted to the subject, no acceptable explanation for this remarkable phenomenon had been found for more than twenty years. If the linear velocity of rotation increases as we approach the centre of the system, which it does according to the observed evidence of the radial velocities, or even if it decreases less rapidly than it would do in a rotating solid body, then the random relative velocities of the stars, i. e. the deviations of the true velocities from the uniform motion in a circle, must necessarily have preference for the radial direction. The agreement of the observed direction of the star streams with the direction towards the centre is not so close as we might wish. Whether the discrepancy is so large as to be really disquieting is difficult to say.

In some spiral nebulae, the Andromeda nebula, M. 33, N. G. C. 4594, and some others, a rotation has been observed from radial velocities by Slipher at Flagstaff and by the observers at Mount Wilson. This strengthens again the analogy which has already several times been mentioned.

The third of the recent discoveries mentioned above as having been of great importance in shaping our ideas about the galactic system is the following.

It has been known for some time that in the spectra of some of the brightest and hottest stars the calcium lines H and K show a different radial velocity from the other lines of the spectrum. The explanation of these "stationary lines," as they used to be called, suggested by Plaskett and theoretically worked out by Eddington, is that they are due to absorption in the extended and rarefied clouds, consisting of ionised calcium, which fill up the interstellar space between the star and the earth.

Recently it has been proved that this calcium cloud shows the same rotational effects in radial velocity as the stars themselves and consequently shares in the rotation of the whole system.

Another important fact is the distribution of planetary nebulae. Planetary nebulae are rather strongly condensed towards the plane of the milky way and, in that plane, towards the longitude of the centre of the globular clusters. The number of planetary nebulae per unit of volume is greater in the direction towards the centre than in the direction away from it. Incidentally we may remark that here again is a statistical result extending beyond the limits of the Kapteyn system. Moreover, the planetary nebulae decidedly show the Oort effect in their radial velocities and thus partake in the rotation of the whole system.

Reviewing the position now, we come to the conclusion that the model that we had of the galactic system after the completion of Shapley's series of papers of the period from 1915 till 1920 needs only little emendation. The stars, galactic nebulae, dark nebulae, planetary nebulae, open

clusters, galactic clouds, etc., are distributed over a narrow segment that extends not farther than 1000 or 1500 ps to both sides of the galactic plane. They are strongly condensed towards the middle of this strip. In the neighbourhood of the sun one-half of the mass is contained between the limits of about 270 ps on both sides of the central plane. Near the centre the width may be (and probably is) greater. This whole system is rotating round a centre which co-incides with the centre of the system of globular clusters. The distance of the solar system from the centre is of the order of 10000 ps and the velocity of rotation at that distance is in the neighbourhood of 300 km/sec. The period of revolution is thus of the order of two hundred million years. It appears probable that the stars, clusters, nebulae, etc., in this flat system show great local varieties of density and are collected in what are commonly described as the galactic star-clouds. The Kapteyn system may be one of these star-clouds. It is, of course, not impossible that the change of density in the direction of the galaxy, found by Kapteyn's method, is only apparent and due to the absorption by dark matter. Whether there exists a "local system" or not, cannot yet with absolute certainty be deduced from our observational data, but the most recent researches appear to go a long way towards establishing its reality.

The density in the direction perpendicular to the galactic plane probably decreases more or less in accordance with the results found by Kapteyn, which would make the density about one-tenth at a distance of 500 parsecs.

The centre of the great galactic system lies in the direction of galactic longitude 325° , roughly speaking on the border of

the constellations Sagittarius and Scorpio, between the two branches of the milky way, the central condensation of mass being hidden from our view by the obscuring matter.

The analogy of our milky way system with the spiral nebulae is very strong and it is very tempting to consider our system as a spiral nebula. If, however, the galactic system is a spiral nebula, our position must be in the outer and less dense parts of the nebula.

The surface luminosity of that part we are in, viewed from the outside, cannot differ much from four times the luminosity of the sky as seen from our position in the direction of the pole. This is very faint, and it appears doubtful whether the whole Kapteyn system would be bright enough to be visible on a photograph of our system taken from a distance, say, of a million lightyears in a direction perpendicular to the galactic plane. Some of the brighter clusters, such as the Pleiades, or Praesepe, would probably be visible, but they would appear as somewhat hazy star-like points from that distance. On the other hand, from the same distance *in* the plane of the milky way it seems probable that the Kapteyn system would be visible as a brighter portion near the end of the nebula, such as is present in the lower left-hand corner of the photograph of N. G. C. 4565, reproduced in Plate XII.

This idea of the larger milky way system will certainly still be changed in so far as the details are concerned, but one thing seems certain and is very strongly impressed on us, viz. that the whole system is one unit. The theory of rotation explains so many observed facts, the strong flattening, the behaviour of stars of large velocity, the star streams, that it can no longer be doubted. Further the fact that many

different kinds of stars, the planetary nebulae, the interstellar calcium clouds, all give the same rotation, shows convincingly that the whole system is one unit. Within the system there are, very probably, rather large differences of density, local conglomerations of stars, but these are not, as has been suggested, separate systems forming a more or less haphazard conglomeration. They all have the same plane of symmetry, and are evidently part of one whole, one organism so to say. As a consequence of the rotation itself, these local conglomerations must necessarily be rather short-lived and cannot survive more than a limited number of revolutions, unless they are very concentrated.

On the other hand we see in other galactic systems, the spiral nebulae, similar conglomerations of matter, and, as these nebulae also must have a rotation, which is evident from their flattening and has been observed in some cases, we have here exactly the same problem. The dimensions of the spiral nebulae are rather large as compared with their relative distances; the ratio is of the order of $1/20$ to $1/50$, whilst the ratio of the diameters of the stars to their mutual distances is of the order of several tens of millions. Near approaches between spiral nebulae may be considerably more frequent than between stars, and the effect of such a near approach would probably tend to destroy the homogeneity if it existed.

There is, however, one strong argument against the identification of our galactic system with these spiral nebulae, and that is the size. We have seen that the equatorial dimensions of the galactic system are of the order of 30000 ps, and true diameters of spiral nebulae, as derived from the best

available apparent diameters and distances, are very much smaller. The mean of about thirty gives something like 2700 parsecs. Our own system would thus be very exceptional. Now at the present time we do not like to be exceptional. There have been times in the history of mankind when any argument that seemed to prove that we occupied a privileged place in the universe would be hailed with approbation, but these times are past. Modern scientific conscience objects to the idea of exceptions being made in our favour, and accordingly we wonder whether there is not a flaw in the argument, and whether perhaps our system is not smaller than we think it is, or the spiral nebulae may be larger than we think they are. As to the latter point, on the one hand the observed apparent diameters of the spirals may be too small; it is not impossible that those portions of the spiral nebulae which are so far out of the centre as we are away from the centre of our system are too faint to appear on the photographs. On the other hand, the adopted distances may be too small. With only two exceptions they depend practically alone on the adopted magnitudes of the brightest stars in them, for which Hubble assumes the absolute magnitude -6.3 . Supposing the objects of which the apparent magnitude has been determined in the spiral nebula and to which this absolute magnitude is assigned to be not stars, but globular (or other) clusters, the true absolute magnitudes would be four or five magnitudes brighter, and the distance, and therefore also the true dimensions, would have to be multiplied by a factor nine or ten, which would make the spirals of the same size as our galactic system. The problem of the determination of the apparent diameters and mag-

nitudes and of the distances of the extra-galactic nebulae is one of the most important and one of the most difficult of the present day.

We have arrived in this chapter at a definite picture of the structure of the universe, which, of course, is not final, but most probably correct in its essential features. We have all these spiral systems, swimming as islands in a great sea of emptiness. Whether our own system is of the nature of a continent, or just like all the other islands, is still uncertain. We know, of course, only a very small part of the universe, which we may call our "neighbourhood," but we like to think that this neighbourhood is just an ordinary portion of the universe, and that those parts which we cannot yet reach with our observations are of the same nature.

In the next lecture we will speak of the expansion of the universe. That is a very remarkable thing, but more remarkable is the recent expansion of our knowledge of the universe. Our "neighbourhood" has expanded about a thousandfold in the last fifteen years, say from a radius of a hundred thousand to a hundred million lightyears, as the result of the observations with the large telescopes, especially the 100-inch reflector at Mount Wilson.

VI

RELATIVITY AND MODERN THEORIES OF THE UNIVERSE · CONCLUSION

THE last fifteen years have seen the most remarkable development of physical and astronomical theories, and of our ideas about the universe. It is outside the scope of the present series of lectures to consider the new theories of atomic physics and the corresponding new development of astrophysics, but we must say some words about the theory of relativity, which is now about sixteen years old.

The theory of relativity may be considered as the logical completion of Newton's theory of gravitation, the direct continuation of the line of thought which dominates the development of the science of mechanics, from Archimedes—who may be considered as the first relativist—through Galileo to Newton. Newton's theory had celebrated its greatest triumphs in the eighteenth and nineteenth centuries; one after another all the irregularities in the motions of the planets and the moon had been explained by the mutual gravitational action of these bodies. In the beginning of the nineteenth century Laplace's monumental work completed the application of the theory on the motions of the planets.

The final triumph came in 1846 by the discovery of Neptune, verifying the prediction by Adams and Leverrier, based on the theory of gravitation.

Gradually Newton's law of gravitation had become a

model on which physical laws were framed, and all physical phenomena were reduced to laws which were formulated as attractions or repulsions inversely proportional to some power of the distance, such as, e. g., Laplace's theory of capillarity, which was even published as a chapter in his "*Mécanique céleste*." Gradually, however, during the second half of the nineteenth century, the uncomfortable feeling of dislike of the action at a distance, which had been so strong in Huygens and other contemporaries of Newton, but had subsided during the eighteenth century, began to emerge again, and gained strength rapidly.

This was favoured by the purely mathematical transformation (which can be compared in a sense with that from the Ptolemaic to the Copernican system), replacing Newton's finite equations by the differential equations, the potential becoming the primary concept, instead of the force, which is only the gradient of the potential. These ideas, of course, arose first in the theory of electricity and magnetism — or perhaps one should say in the brain of Faraday. In electromagnetism also the law of the inverse square had been supreme, but, as a consequence of the work of Faraday and Maxwell, it was superseded by the field. And the same change took place in the theory of gravitation. By and by the material particles, electrically charged bodies, and magnets — which are the things that we actually observe — come to be looked upon only as "singularities" in the field. So far this transformation from the force to the potential, from the action at a distance to the field, is only a purely mathematical operation. Whether we talk of a "particle of matter" or of a "singularity in the gravitational field" is

only a question of a name. But this giving of names is not so innocent as it looks. It has opened the gate for the entrance of hypotheses. Very soon the field is materialised, and is called aether. From the mathematical point of view, of course, "aether" is still just another word for "field," or, perhaps better, for "space" — the absolute space of Newton — in which there may or may not be a field. From the point of view of physical theory (and it is especially in the theory of electromagnetism that this evolution took place), however, the "aether of space," as it used to be called about forty years ago, is not simply space, it is something substantial, it is the carrier of the field, and mechanical models, consisting of racks and pinions and cogwheels, are devised to explain how it does the carrying. These mechanical models have, of course, been given up long ago: they were too crude. But hypotheses have kept cropping up on all sides: electrons, atomic nuclei, protons, wave-packets, etc. At first the imagining of mechanical models went on. Fifteen, or even ten, years ago, although an atom was no longer, as the name implies, just a piece of matter that could not be cut into smaller pieces, atoms, electrons, and protons were still thought of as mechanical structures, models of the atom were imagined, having the mechanical properties of ordinary matter. The inconsistency of first explaining matter by atoms and then explaining atoms by matter was only slowly realised, and it is only comparatively recently that we have come to see that there is nothing paradoxical in the fact that an atom or an electron, which are not matter, may have properties different from those of matter, and must be allowed to do things that a material particle could not do.

However, whilst in all other domains of physics hypotheses have been found successful in accounting for the observed facts, and replacing the formal laws, the case of gravitation stands apart. Gravitation has been insusceptible to this general infection. By using this word I do not mean to suggest that the luxuriant growth of hypotheses in physics is a contagious disease, — it is not a disease, but a natural development, — but it is certainly contagious. Gravitation, however, seems to be immune to it. In the course of history a great number of hypotheses have been proposed in order to “explain” gravitation, but not one of these has ever had the least chance, they have all been failures. Why is that? How does it come about that we have been able to find satisfactory hypotheses to explain electricity and magnetism, light and heat, in short all other physical phenomena, but have been unsuccessful in the case of gravitation? The explanation must be sought in the peculiar position that gravitation occupies amongst the laws of nature. In the case of other physical phenomena there is something to get hold of, there are circumstances on which the action depends. Gravitation is entirely independent of everything that influences other natural phenomena. It is not subject to absorption or refraction, no velocity of propagation has been observed. You can do whatever you please with a body, you can electrify or magnetise it, you can heat it, melt or evaporate it, decompose it chemically, its behaviour with respect to gravitation is not affected. Gravitation acts on all bodies in the same way, everywhere and always we find it in the same rigorous and simple form, which frustrates all our attempts to penetrate into its internal mechanism. Gravita-

tion is, in its generality and rigour, entirely similar to inertia, which has never been considered to require a particular hypothesis for its explanation, as any ordinary special physical law or phenomenon. Inertia has from the beginning been admitted as one of the fundamental facts of nature, which have to be accepted without explanation, like the axioms of geometry.

But gravitation is not only similar to inertia in its generality, it is also measured by the same number, called the mass. The inertial mass is what Newton calls the "quantity of matter": it is a measure for the resistance offered by a body to a force trying to alter its state of motion. It might be called the "passive mass." The gravitational mass, on the other hand, is a measure of the force exerted by the body in attracting other bodies. We might call it the "active" mass. The equality of active and passive, or gravitational and inertial, mass was in Newton's system a most remarkable accidental co-incidence, something like a miracle. Newton himself decidedly felt it as such, and made experiments to verify it, by swinging a pendulum with a hollow bob which could be filled with different materials. The force acting on the pendulum is proportional to its gravitational mass, the inertia to its inertial mass: the period of its swing thus depends on the ratio between these two masses. The fact that the period is always the same therefore proves that the gravitational and inertial masses are equal. Gradually, during the eighteenth century, physicists and philosophers had become so accustomed to Newton's law of gravitation, and to the equality of gravitational and inertial mass, that the miraculousness of it was forgotten and only an acute mind

like Bessel's perceived the necessity of repeating those experiments. By the experiments of Bessel about 1830 and of Eötvös in 1909 the equality of gravitational and inertial mass has become one of the best ascertained empirical facts in physics.

In Einstein's general theory of relativity the identity of these two coefficients, the gravitational and the inertial mass, is no longer a miracle, but a necessity, because gravitation and inertia are identical.

There is another side to the theory of relativity. We have pointed out in the beginning how the development of science is in the direction to make it less subjective, to separate more and more in the observed facts that which belongs to the reality behind the phenomena, the absolute, from the subjective element, which is introduced by the observer, the relative. Einstein's theory is a great step in that direction. We can say that the theory of relativity is intended to remove entirely the relative and exhibit the pure absolute.

The physical world has three space dimensions and one time dimension; the position of a material particle at a certain time t is defined by three space coördinates, x , y , z . In Newton's system of mechanics this is unhesitatingly accepted as a property of the outside world: there is an absolute space and an absolute time. In Einstein's theory time and space are interwoven, and the way in which they are interwoven depends on the observer. Instead of three plus one we have four dimensions.

Is the fact that we observe the outside world as a four-dimensional continuum a property of this outside world, or is it a consequence of the particular nature of our consciousness,

does it belong to the absolute or to the relative? I do not think the answer to that question can yet be given. For the present we may accept it as an empirically ascertained fact.

The sequence of different positions of the same particle at different times forms a one-dimensional continuum in the four-dimensional space-time, which is called the *world-line* of the particle. All that physical experiments or observations can teach us refers to intersections of world-lines of different material particles, light-pulsations, etc., and how the course of the world-line is between these points of intersection is entirely irrelevant and outside the domain of physics. The system of intersecting world-lines can thus be twisted about at will, so long as no points of intersection are destroyed or created, and their order is not changed. It follows that the equations expressing the physical laws must be invariant for arbitrary transformations.

This is the mathematical formulation of the theory of relativity. The metric properties of the four-dimensional continuum are described, as is shown in treatises on differential geometry, by a certain number (ten, in fact) of quantities denoted by $g_{\alpha\beta}$, and commonly called "potentials." The physical status of matter and energy, on the other hand, is described by ten other quantities, denoted by $T_{\alpha\beta}$, the set of which is called the "material tensor." This special tensor has been selected because it has the property which is mathematically expressed by saying that its divergence vanishes, which means that it represents something permanent. The fundamental fact of mechanics is the law of inertia, which can be expressed in its most simple form by saying that it requires the fundamental laws of nature to be differential

equations of the second order. Thus the problem was to find a differential equation of the second order giving a relation between the metric tensor $g_{\alpha\beta}$ and the material tensor $T_{\alpha\beta}$. This is a purely mathematical problem, which can be solved without any reference to the physical meaning of the symbols. The simplest possible equation (or rather set of ten equations, because there are ten g 's) of that kind that can be found was adopted by Einstein as the fundamental equation of his theory. It defines the space-time continuum, or the "field." The world-lines of material particles and light quanta are the geodesics in the four-dimensional continuum defined by the solutions $g_{\alpha\beta}$ of these field-equations. The equations of the geodesic thus are equivalent to the equations of motion of mechanics. When we come to solve the field-equations and substitute the solutions in the equations of motion, we find that in the first approximation, i. e. for small material velocities (small as compared with the velocity of light), these equations of motion are the same as those resulting from Newton's theory of gravitation. The distinction between gravitation and inertia has disappeared; the gravitational action between two bodies follows from the same equations, and *is* the same thing, as the inertia of one body. A body, when not subjected to an extraneous force (i. e. a force other than gravitation), describes a geodesic in the continuum, just as it described a geodesic, or straight line, in the absolute space of Newton under the influence of inertia alone.

The field-equations and the equations of the geodesic together contain the whole science of mechanics, including gravitation.

In the first approximation, as has been said just now, the new theory gives the same results as Newton's theory of gravitation. The enormous wealth of experimental verification of Newton's law, which has been accumulated during about two and a half centuries, is therefore at the same time an equally strong verification of the new theory. In the second approximation there are small differences, which have been confirmed by observations, so far as they are large enough for such a confirmation to be possible. Thus especially the anomalous motion of the perihelion of Mercury, which had baffled all attempts at explanation for over half a century, is now entirely accounted for. Further the theory of relativity has predicted some new phenomena, such as the deflection of rays of light that pass near the sun, which has actually been observed on several occasions during eclipses; and the redshift of spectral lines originating in a strong gravitational field, which is also confirmed by observations, e. g. in the spectrum of the sun, and also in the spectrum of the companion of Sirius, which, being a so-called white dwarf, i. e. a small star with very high density and consequently a strong gravitational field, gives a considerable redshift. We cannot stop to explain these phenomena in detail. It must suffice just to mention them.

Two points should be specially emphasised in connection with the general theory of relativity.

First that it is a purely *physical* theory, invented to explain empirical physical facts, especially the identity of gravitational and inertial mass, and to coördinate and harmonise different chapters of physical theory, and simplify the enunciation of the fundamental laws. There is nothing meta-

physical about its origin. It has, of course, largely attracted the attention of philosophers, and has, on the whole, had a very wholesome influence on metaphysical theories. But that is not what it set out to do, that is only a by-product.

Second that it is a pure generalisation, or abstraction, like Newton's system of mechanics and law of gravitation. It contains *no hypothesis*, as contrasted with other modern physical theories, electron theory, quantum theory, etc., which are full of hypotheses. It is, as has already been said, to be considered as the logical sequence and completion of Newton's Principia.

A special feature of the development of physics in the nineteenth century has been the arising of general principles beside the special laws, such as the principles of conservation of mass and of energy, the principle of least action, and the like. These differ from the special laws, not only by being more general, but they aspire, so to say, to a higher status than the laws. Their claim is that they express fundamental facts of nature, general rules, to which all special laws have to conform. And they accordingly exclude a priori all attempts at "explanation" by hypotheses or mechanical models. It is characteristic of the theory of relativity that it enables us to include all these principles of conservation in one single equation.

We have a direct knowledge only of that part of the universe of which we can make observations. I have already called this "our neighbourhood." Even within the confines of this province our knowledge decreases very rapidly as we get away from our own particular position in space and time.

It is only within the solar system that our empirical knowledge of the quantities determining the state of the universe, the potentials $g_{\alpha\beta}$, extends to the second order of smallness (and that only for g_{44} , and not for the others), the first order corresponding to about one unit in the eighth decimal place. How the $g_{\alpha\beta}$ outside our neighbourhood are, we do not know, and how they are at infinity, of either space or time, we shall never know, otherwise it would not be infinity. That is what Archimedes meant when he said that the universe could not be infinite. The universe that we know cannot be infinite, because we ourselves are finite. Infinity is not a physical, but a mathematical concept, introduced to make our equations more symmetrical and elegant. From the physical point of view everything that is outside our neighbourhood is pure extrapolation, and we are entirely free to make this extrapolation as we please, to suit our philosophical or aesthetical predilections — or prejudices. It is true that some of these prejudices are so deeply rooted that we can hardly avoid believing them to be above any possible suspicion of doubt, but this conviction is not founded on any physical basis. One of these convictions, on which extrapolation is naturally based, is that the particular part of the universe in which we happen to be is in no way exceptional or privileged, in other words that the universe, when considered on a large enough scale, is isotropic and homogeneous. It should, however, be remembered that there have been epochs in the evolution of mankind when this was by no means thought self-evident, and the contrary conviction was rather generally held.

During the last years the limits of our “neighbourhood”

have been enormously extended by the observations of extragalactic nebulae, made chiefly at the Mount Wilson Observatory. These wonderful observations have enabled us to make fairly reliable estimates of the distances of these objects and to say something about their distribution in space. It appears that they are distributed approximately evenly over "our neighbourhood." They also are all of roughly the same size, so that we can make an estimate of the density of matter in space. Further the observations have disclosed the remarkable fact that in their spectra there is a displacement of the lines towards the red corresponding to a receding velocity increasing with the distance, and, so far as the determinations of the distances are reliable, proportional with it. If the velocity is proportional to the distance, then not only the distance of any nebula from us is increasing, but *all* mutual distances between any two of them are increasing at the same rate. Our own galactic system is only one of a great many, and observations made from any of the others would show exactly the same thing: all systems are receding, not from any particular centre, but *from each other*: the whole system of galactic systems is *expanding*.

It is perhaps somewhat difficult to imagine the expansion of three-dimensional space. A two-dimensional analogy may help to make it clear. Let the universe have only two dimensions, and let it be the surface of an indiarubber ball. It is only the *surface* that is the universe, not the ball itself. Observations can only be made, distances can only be measured, along the surface, and evidently no point of the surface is different from any other point. Let there be specks of dust fixed to the surface to represent the different galactic sys-

tems. If the ball is inflated, the universe expands, and these specks of dust will recede from each other, their mutual distances, measured along the surface, will increase in the same rate as the radius of the ball. An observer in any one of the specks will see all the others receding from himself, but it does not follow that he is the centre of the universe. The universe (which is the surface of the ball, not the ball itself) has no centre.

It is, of course, not essential that we have chosen for our illustration a rubber *ball*. We might just as well have taken any other surface, it is not even necessary that it should be a closed surface. Even a plane sheet of rubber might do just as well, if only the stretching to which it must be subjected to illustrate the expanding universe is the same in all directions.

These then are the two observational facts about our neighbourhood, which have to be accounted for by the theory: there is a finite density of matter, and there is expansion, i. e. the mutual distances are increasing, and therefore the density is decreasing. Of course we can only be certain of these facts so far as our observations reach, i. e. for our "neighbourhood," but, in agreement with our principle of extrapolation, we extend these statements to the whole of the universe.

We have thus to find a universe — i. e. a set of potentials $g_{\alpha\beta}$ satisfying the field-equations of the general theory of relativity — that has both a finite density of matter and an expansion. And, since we only consider the universe on a very large scale, and make abstraction of all details and local irregularities, our universe must be homogeneous and iso-

tropic. It follows at once from this condition of homogeneity and isotropy that the three-dimensional space of it must be what mathematicians call a space of constant curvature. Even so mathematics offers us a free choice between different kinds of space. The curvature may be positive, negative, or zero. It is not possible to picture, or imagine, the different kinds of three-dimensional space. We think that we have a mental picture of euclidian, or flat, space, i. e. space of which the curvature is everywhere zero, but I am not sure that this is not a self-deception, caused by the fact that the geometry of this special space has been taught in the schools for the last two thousand or more years. It is certain that for physical phenomena on the scale which our sense organs are able to perceive, i. e. neither too small nor too large, the euclidian space is a very close approximation to the true physical space, but for the electron, and for the universe, the approximation breaks down. To help us to understand three-dimensional spaces, two-dimensional analogies may be very useful (though also sometimes misleading). We can imagine different kinds of two-dimensional space, since we are able to place ourselves outside them. A two-dimensional space of zero curvature is a plane, say a sheet of paper. The two-dimensional space of positive curvature is a convex surface, such as the shell of an egg. It is bent away from the plane towards the same side in all directions. The curvature of the egg, however, is not constant: it is strongest at the small end. The surface of constant positive curvature is the sphere, say our indiarubber ball of a moment ago. The two-dimensional space of negative curvature is a surface that is convex in some directions and concave in others, such as the surface

of a saddle or the middle part of an hour glass. Of these two-dimensional surfaces we can form a mental picture because we can view them from outside, living, as we do, in three-dimensional space. But for a being, who would be unable to leave the surface on which he was living, that would be impossible. He could only decide of which kind his surface was by studying the properties of geometrical figures drawn on it. For the geometrical figures have different properties on the different surfaces. On the sheet of paper the sum of the three angles of a triangle is equal to two right angles, on the egg, or the sphere, it is larger, on the saddle it is smaller. On the flat paper — and on the saddle-shaped surface — we can proceed indefinitely in the same direction, on the egg or the sphere, if we continue to move in the same direction we ultimately come back to our starting point. The spaces of zero and negative curvature are infinite, that of positive curvature is finite. Thus the inhabitant of the two-dimensional surface could determine its curvature if he were able to study very large triangles or very long straight lines. If the curvature were so minute that the sum of the angles of the largest triangle that he could measure would still differ from two right angles by an amount too small to be appreciable with the means at his disposal, then he would be unable to determine the curvature, unless he had some means of communicating with somebody living in the third dimension. Now our case with reference to three-dimensional space is exactly similar. We have no intuitive knowledge of the kind of space we live in. So we must find out which kind it is by studying the triangles and other geometrical figures in it. As we are concerned with *physical* space, the triangles that

we must investigate are those formed by the tracks of material particles and rays of light, and naturally, in order to be able to distinguish different kinds of space, we must study very large triangles and rays of light coming from very great distances. Thus the decision must necessarily depend on astronomical observations.

Even the most refined astronomical observations, however, fail to show any trace of curvature. The triangles that we can measure are not large enough, and, I fear, never will be large enough, to detect the curvature. Fortunately, however, we are, in a way, able to communicate with the fourth dimension. The theory of relativity has given us an insight into the structure of the real universe: it does not consist of a three-dimensional space *and* a one-dimensional time, existing independently of each other, as in Newton's system of mechanics, but is a *four*-dimensional structure. The study of the way in which the three space-dimensions are interwoven with the time-dimension affords a kind of outside point of view of the three-dimensional space, and it is conceivable that from this outside point of view we might be able to perceive the curvature of the three-dimensional world.

At one time it was thought that this was so, and that we could actually prove that the curvature must be positive. However, recent mathematical investigation has shown that this was a mistake. We shall never be able to say anything about the curvature without introducing certain hypotheses. These rather vague statements will become clearer when we penetrate more deeply into the nature of the cross-connection between space and time in the four-dimensional universe. It is, of course, difficult to explain these without the use of

mathematical formulae, but, by following the historical line of development, I shall endeavour to lead up to an understanding of the present position without using too technical language.

Let us begin by considering the finite density of matter in the universe. The average density is very small. Matter is actually distributed very unevenly, it is conglomerated into stars and galactic systems. The average density is the density that we should get if all these great systems could be evaporated into atoms of hydrogen, or protons, and these distributed evenly over the whole of space. There would then probably not be more than three or four protons in every cubic foot. That is a very small density indeed: it is about a million million times less than that of the most perfect vacuum that we can produce in our physical laboratories. The universe thus consists mostly of emptiness, and it appears natural to consider a universe without any matter at all, an empty universe, as a good approximation to begin with for our grand scale model. The galactic systems are details which can be put in afterwards. But we may also take as our first approximation a universe containing the same amount of matter as the actual one, but equally distributed, i. e. having a finite average density of three or four protons per cubic foot. The local deviations from the average, caused by the conglomeration of matter into stars and stellar systems, are then disregarded in the grand scale model, and are only taken into account when we come to study details.

Now fifteen years ago, in the beginning of 1917, two solutions of the field-equations for a homogeneous isotropic uni-

verse had been found, which I shall provisionally call the solutions "A" and "B." It should be mentioned that at that time only *static* solutions were looked for. It was thought that the universe must be a stable structure, which would retain its large scale properties unchanged for all time, or at least change them so slowly that the change could be disregarded. In one of these solutions (B) the average density was zero, it was empty; the other one (A) had a finite density. Both, of course, were, as was well appreciated, only approximations to the actual universe. In B, to get the real universe, we should have to put in a few galactic systems, in A we should have to condense the evenly distributed matter into galactic systems. The universe A is really and essentially static, there can be no systematic motions in it. It has an average density, but no expansion. It is therefore called the *static universe*. B, on the other hand, is not really static; it expands, and it could only parade in the garb of a static universe because there is nothing in it to show the expansion. B is therefore called the *empty universe*. Thus we had two approximations: the static universe with matter and without expansion, and the empty one without matter and with expansion. The actual universe, as we have seen, has both matter and expansion, and can, therefore, be neither A nor B. In 1917 this dilemma had not yet become urgent, and was hardly realised. The actual value of the density was still entirely unknown, and the expansion had not yet been discovered.

Now in both the solutions A and B the curvature is positive, in both three-dimensional space is finite: the universe has a definite size, we can speak of its radius, and, in the

case A, of its total mass. In the case A, the static universe, there is a definite relation between the curvature and the density, in fact the density is proportional to the curvature, the factor of proportionality being a pure number ($1/4\pi$, if appropriate units are used). Thus, if we wish to have a finite density in a static universe, we must have a finite positive curvature.

At this point we must say a few words about the famous *lambda*. The field-equations, in their most general form, contain a term multiplied by a constant, which is denoted by the Greek letter λ (lambda), and which is sometimes called the "cosmical constant." This is a name without any meaning, which was only conferred upon it because it was thought appropriate that it should have a name, and because it appeared to have something to do with the constitution of the universe; but it must not be inferred that, since we have given it a name, we know what it means. We have, in fact, not the slightest inkling of what its real significance is. It is put in the equations in order to give them the greatest possible degree of mathematical generality, but, so far as its mathematical function is concerned, it is entirely undetermined: it may be positive or negative, it might also be zero. Purely mathematical symbols have no meaning by themselves; it is the privilege of pure mathematicians, to quote Bertrand Russell, not to know what they are talking about. They — the symbols — only get a meaning by the interpretation that is put on the equations when they are applied to the solution of physical problems. It is the physicist, and not the mathematician, who must know what he is talking about. At first, in Einstein's paper of November 1915, in which the

theory reached its final form, the term with λ was simply omitted, in other words λ was supposed to have the special value *zero*. That was the simplest way of avoiding the responsibility of attaching a label to it, and, of course, an entirely legitimate way: it is very bad physics to introduce more arbitrary constants than are needed for the representation of the phenomena. But fifteen months later, in February 1917, it was found that a static solution with a positive curvature — the solution A — was not possible without the λ . In fact the curvature is proportional to λ (in solution A, λ is equal to the curvature; in B, when treated as a static solution, it is three times the curvature). Thus, at the time when we had only the two static solutions A and B, and thought that these were the only possible ones, here was a plausible physical interpretation of the meaning of λ : it was the curvature of the world, and the square root of its reciprocal, the radius of curvature, could be conceived as providing a natural unit of length. It gave the electron something to measure itself by, so that it might know how large it ought to be, as Sir Arthur Eddington has expressed it. Recently Eddington has replaced this interpretation by another one of the same nature, involving the gravitational mass and the electric charge of the electron. But recent developments have made it very difficult to maintain these interpretations, as we shall presently see.

We must now take up the thread of the narrative where we left it a little while ago. We were in the position of having two possible solutions: the static universe with matter but without expansion, and the empty universe with expansion but without matter.

Now the observed rate of expansion is large: the universe doubles its size in about fifteen hundred million years, which is a short time, astronomically speaking. In the "static" universe expansion is impossible, the "empty" universe does expand. Therefore we may be tempted to consider the empty universe as the most likely approximation; and we can proceed to compute the radius of curvature of the universe, supposing it to be of the empty type, from the observed rate of expansion. It comes out as about two thousand million lightyears.

The universe, however, is not empty, but contains matter. The point is how much matter. Is the density anywhere near that corresponding to the static universe, or is it so small that we can consider the empty universe as a good approximation? We have seen that the universe is some million million times as empty as our most perfect vacuum. But this is not the correct way to measure the emptiness of the universe. We must use as a standard of comparison, not our terrestrial experience, but the theoretical density of the static universe. It is easy to compute the density of a static universe of a radius of two thousand million lightyears, and it comes out only very little larger than the observed density. The actual universe is thus very far from empty, it is, on the contrary, nearly full.

We thus come to the conclusion, which was already foreshadowed above, that the actual universe is neither the static nor the empty one. It differs so much from both of these that neither can be used as an appropriate grand scale model. We must thus look for other solutions of the general field-equations. On account of the expansion our solution must

necessarily be a non-static one, and it must have a finite density. There is only one possible static solution possessing a finite density, viz. our old friend A, but of non-static solutions with finite density there exists a great variety. I will now depart from the strictly historical narrative and enumerate these different possible solutions, not in the order in which they have been discovered, but in the sequence of a natural classification.

In the solutions A and B the curvature of three-dimensional space was necessarily positive, and the mysterious "cosmical constant" λ was also positive. In the non-static solutions this is not so. At first this was not realised. We had become so accustomed to think of λ as an essentially positive quantity, and of a finite world with positive curvature, that the idea of investigating the possibility of solutions with negative or zero values of λ and of the curvature simply did not arise. But when this oversight was corrected, it appeared at once that in the non-static case both λ and the curvature need not be positive, but can be negative or zero quite as well. I will therefore use the value of λ and the sign of the curvature as the principles of classification. The instantaneous state of the universe is characterised by a certain quantity occurring in the equations, which is denoted by the letter R , and which, if there is a curvature, can be interpreted as the radius of curvature, or the "radius" for short. The way in which the universe expands is determined by the variation of this R with the time. There are three types, or families, of non-static universes, which I will call the oscillating universes, and the expanding universes of the first and of the second kind. They are represented graphi-

cally in Figure 11. The horizontal coördinate is the time, the vertical coördinate is the "radius." Of each type only one example is given in the figure, but it should be realised that each of these is a representative of a family, comprising an infinite number of members differing in size and shape.

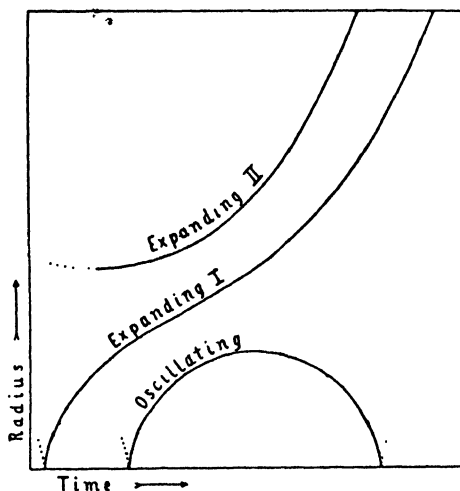


FIGURE 11

The three families of non-static universes

In the oscillating universes the "radius" R increases from zero to a certain maximum size, which is different for each member of the family, and then decreases again to zero. The period of oscillation has a certain finite (and rather short) value, different for each member of the family. In the expanding family of the first kind the radius is continually increasing from a certain initial time, when it was zero, to become infinitely large after an infinite time. In the expanding series of the second type the radius has at the initial time a certain minimum value, different for the different members

of the family, and increases to become infinite after an infinite time.

If λ is negative, only oscillating universes are possible, whatever the curvature may be, positive, negative or zero. The only choice is between the different members of the oscillating family, and to each special value of λ corresponds one member of the family.

If λ is zero, and the curvature positive, we have still an oscillating universe; if the curvature is zero or negative, the universe is one of the first family of expanding universes.

If λ is positive, then for positive curvatures the three families are possible; for negative and zero curvatures only the first family of expanding universes exists. The different possibilities are given in the following small table.

TABLE VII

λ	Curvature		
	Negative	Zero	Positive
Negative	Oscillating	Oscillating	Oscillating
Zero	Expanding I	Expanding I	Oscillating
Positive.....	Expanding I	Expanding I	Oscillating Expanding I Expanding II

We do not know to which of the three possible families our own universe belongs, and there is nothing in our observational data to guide us in making the choice. And even if we have decided on the family, we have still the freedom to select any particular member of it. This is not because the data are not accurate enough, but because they are deficient

in number. The observations give us *two* data, viz. the rate of expansion and the average density, and there are *three* unknowns: the value of λ , the sign of the curvature, and the scale of the figure, i. e. the units of R and of the time. The problem is indeterminate. If we make an hypothesis regarding either λ or the curvature, then we can find the other from the observed data, or rather we would be able to do so, if the data were sufficiently accurate. We might, e. g., decide a priori that the curvature should be zero, i. e. that the three-dimensional space should be euclidean. If we make that hypothesis, we have a sufficient number of data to enable us to determine the value of λ . We might, on the other hand, wish to get rid of the λ , on the ground that it ought never to have been there, being unfortunately introduced into the equations in a past stage of the development of the theory, when we were mistakenly trying to find a static solution. In other words we might make the hypothesis that the true value of λ is zero. In that case the data of observation will allow us, if they are sufficiently accurate, to determine the curvature.

As a matter of fact neither the average density nor the rate of expansion are at the present time known with sufficient accuracy to make an actual *determination* possible, even if an hypothesis of this kind is adopted. All we can say is that, if the curvature is small (as we know it must be, because it is imperceptible by ordinary geometric methods in our neighbourhood), then λ must be small, and *if* the curvature is *very* small, then λ must be very small. On the other hand, *if* λ is very small, or zero, then the curvature must be very small, and may even be zero, for aught we can say at present.

The interpretation of the expanding universe, the making of a mental picture, or a model, of it, which was, or appeared to be, easy when we knew that the universe was finite, is not such a simple matter now that we do not even know whether the curvature is positive, zero or negative, whether the universe is finite or infinite. It sounds rather strange to talk of an infinite universe still expanding. If we were certain that the curvature was negative, we might still, as in the case of positive curvature, replace the phrase "the universe expands" by the equivalent one "the curvature of the universe decreases." But if the curvature is zero, and remains zero throughout, what sort of meaning are we to attach to the "expansion"? The real meaning is, of course, that the mutual distances between the galactic systems, measured in so-called natural measure, increase proportionally to a certain quantity R appearing in the equations, and varying with the time. The interpretation of R as the "radius of curvature" of the universe, though still possible if the universe has a curvature, evidently does not go down to the fundamental meaning of it. The manner in which time and space are bound up with each other in the four-dimensional continuum is variable. It is difficult to express this variability of the cross-connections between space and time in simple language, and different interpretations of it are possible, corresponding to different mathematical transformations of the fundamental line-element, e. g. a different choice of the variable which we interpret as "time." Perhaps the best way we can express it is by saying that the solution of the field-equations of the theory of relativity shows that there is in the universe a tendency to change its scale, which at

the present time results in an expansion, but may perhaps at other times become, or have been, a shrinking. This is true of the grand scale model of the universe. If we put in the details, the singularities of the field, viz. the galactic systems and the stars, we find that there is also a tendency, called gravitation, to decrease the mutual distances of these "singularities." At short distances, within the confines of a galactic system, this second tendency is by far the strongest, and the galactic systems retain their size independent of the expansion or contraction of the universe; at large distances, such as those separating one galactic system from the next, the first preponderates.

The interpretation of λ as providing a natural unit of length also, of course, becomes difficult. The radius of curvature, or even the more vaguely defined quantity R , can no longer provide the electron with the means of knowing how large it ought to be, for R changes and the electron does not change its size. The electron could still measure itself in some mysterious way against the square root of the reciprocal of λ — or of *minus* λ , if λ happened to be negative, — but if λ were zero that also would not help it.

Also other interpretations, involving the total mass of, or the number of protons in, the universe, become untenable, if the universe may, for aught we know, be just as well infinite and have no total mass.

The theory of the expanding universe is at the present moment much less definite than we supposed it to be a few months ago, but that does not affect its real significance — it only brings out more clearly what this significance is, and what it is not.

Some consequences of the theory merit a special mention, however briefly.

A question which has long troubled astronomers and physicists is what becomes of the energy that is continually being poured out into space by the sun and the stars. To this question a complete answer is given by the new theory. It is used up, diluted, or degraded, by the expansion of the universe. Just as a man running to catch a bus or a tram-car gets out of breath and spends his energy, or a projectile thrown after a moving train hits it with less force than it would hit a stationary object, so the light travelling through the expanding universe and, so to say, trying to reach a particular star, or stellar system, which is continually receding with great velocity, is losing energy in trying to catch up with it. It is this degradation of the light, technically known as the red-shift of the spectral lines, by which we become aware of the receding velocities of the extra-galactic nebulae. It can be shown that the decrease of the total amount of radiant energy in the universe by this degradation exceeds the increase by the radiation of the stars. It would not be correct, however, to conclude that the expansion is caused *by* the energy thus lost by the radiation, any more than it would be correct to say that the tram-car is propelled by the energy expended by the man who runs after it.

There is one very serious difficulty presented by the theory of the expanding universe, which we shall have to face with careful deliberation.

In all solutions there is a certain minimum value of the "radius" R , either zero, or in the expanding family of the second kind a finite value, which the universe had at a defi-

nite time in the past. There appears to be a definite "beginning of time," a few thousand million years back in history, as there is a definite "absolute zero" of temperature, corresponding to *minus* 273 degrees on the ordinary scale. What is the meaning of this?

The temptation is strong to identify the epoch of the beginning of the expansion with the "beginning of the world," whatever that may mean. Now astronomically speaking this beginning of the expansion took place only yesterday, not much longer ago than the formation of the oldest rocks on the earth. According to all our modern views the evolution of a star, of a double star, or a star cluster, requires intervals of time which are enormously longer. The stars and the stellar systems must be some thousands of times older than the universe!

What must be our attitude with regard to this paradox? It would appear that, if two theories are in contradiction, we must give up either the one or the other. The conflict apparently is between the modern theories of stellar evolution and the dynamical theories of the evolution of double stars and star clusters on the one hand, and the general theory of relativity on the other hand. If this were the real contest, there could be no doubt about the issue: the theory of relativity would come out of the trial victorious, and the theories of evolution would have to be revised. This seems to be Sir Arthur Eddington's standpoint, as he writes: "we must accept this alarmingly rapid dispersal of the nebula with its important consequences in limiting the time available for evolution." I am afraid, however, that very few astronomers, not to speak of geophysicists, will be prepared to accept this drastic reduction of the time scale.

It is possible to relegate the epoch of the starting of the expansion to minus infinity, e. g. by using instead of the ordinary time the logarithm of the time elapsed since the beginning. But this is only a mathematical trick. We call zero minus infinity, but that only means that we allow the universe an infinite time to get well started on its course of expansion, but it does not make the time during which anything really happens any longer.

Mention should be made of a suggestion that has sometimes been advanced, viz. that the observed shift of the spectral lines towards the red might not indicate a receding motion of the spiral nebulae, but might be accounted for in some different way. In fact, all that the observations tell us is that light coming from great distances — and which therefore has been a long time under way — is redder when it arrives than when it left its source. *Light is reddened by age*, it loses energy as it gets older, travelling through space. Or expressed mathematically: the wavelength of light is proportional to a certain quantity R , which increases with the time. By the general equations of the theory of relativity the naturally measured distances in a homogeneous and isotropic world are then necessarily proportional to the same quantity R , unless some extraneous cause for the increase of wavelength, or the loss of energy, is present. By extraneous I mean foreign to the theory of relativity and the conception of the nature of light consistent with that theory. Moreover this hypothetical cause should have no other observable consequences, especially it should produce loss of energy without any concomitant dispersion, which would blur the images and make the faint nebulae unobservable.

It would require an hypothesis *ad hoc*, and a very carefully framed one too, so as not to overshoot the mark. No such hypothesis, deserving serious consideration, has yet been forthcoming.

It appears to me that there is no way out of the dilemma. It is an unavoidable consequence of the equations that the time taken by the radius to increase from anywhere near its minimum to its present value is of the same order of magnitude as the present radius itself, if we adopt corresponding units of space and time, e. g. years for time and lightyears for space. The scale is determined by the observed rate of expansion. I am afraid all we can do is to accept the paradox and try to accommodate ourselves to it, as we have done to so many paradoxes lately in modern physical theories. We shall have to get accustomed to the idea that the change of the quantity R , commonly called the "radius of the universe," and the evolutionary changes of stars and stellar systems are two different processes, going on side by side without any apparent connection between them. After all the "universe" is an hypothesis, like the atom, and must be allowed the freedom to have properties and to do things which would be contradictory and impossible for a finite material structure. What we observe are the stars and nebulae constituting "our neighbourhood." All that goes beyond that, in time or in space, or both, is pure extrapolation. The conclusions derived about the expanding universe depend on the assumed homogeneity and isotropy, i. e. on the hypothesis that the observed finite material density and expansion of our neighbourhood are not local phenomena, but properties of the "universe." It is not inconceivable

that this hypothesis may at some future stage of the development of science have to be given up, or modified, or at least differently interpreted.

Our conception of the structure of the universe bears all the marks of a transitory structure. Our theories are decidedly in a state of continuous, and just now very rapid, evolution. It is not possible to predict how long our present views and interpretations will remain unaltered and how soon they will have to be replaced by perhaps very different ones, based on new observational data and new critical insight in their connection with other data.

Meanwhile the simple workers in science go on quietly, each working at his own particular problem, undisturbed by the many strange and contradictory things that are happening around them and in their own house. And it is on this quiet and unostentatious work that the great advances of science are based. Especially in astronomy two characteristics are common to all data on which the solution of the great problems depends. The first is the extreme minuteness of the quantities to be measured. The determinations of the distances, of the proper motions and radial velocities, which form the materials out of which the great structures of the theories of the galactic system and the universe have been built, all require very accurate measurement of extremely small quantities. It is always a struggle for the last decimal place, and the great triumphs of science are gained when, by new methods or new instruments, the last decimal is made into the penultimate. Thus, as we have seen, new epochs were inaugurated in the beginning of the seventeenth century by the invention of the telescope, and

in the last third of the nineteenth by the discovery of photography and spectroscopy.

The other characteristic is that astronomy always requires a very large number of data. As a consequence of the fact that direct experiment is impossible in astronomy, we must have observations of very many stars, and extended to ever fainter objects, in order to enable us to draw reliable conclusions.

These two characteristics of the data that the astronomer requires to build his science on make two things more necessary in astronomy than in any other science: patience and organised coöperation. Patience because many of the phenomena develop so slowly that a long time is necessary for them to become measurable, coöperation because the material is too large and too various to be mastered by one man, or even by one institute. And coöperation not only between different workers and institutions all over the world, but also coöperation with predecessors and successors for the solution of problems that require, by their very nature, more than one man's lifetime. The astronomer — each working at his own task, whether performing long calculations, making theories and hypotheses, or patiently collecting observations in daily routine — is always conscious of belonging to a community, whose members, separated in space and time, nevertheless feel joined by a very real tie, almost of kinship. He does not work for himself alone, he is not guided exclusively, and not even in the first place, by his own insight or preferences, his work is always coördinated with that of others as a part of an organised whole. He knows that, whatever his special work may be, it is always a link in a chain, which de-

rives its value from the fact that there is another link to the left and one to the right of it. It is the chain that is important, not the separate links.

It cannot be doubted that at the present moment science stands high in the popular esteem, much higher than it did even a short time ago. But I am afraid that often in the popular mind a sufficiently sharp distinction is not drawn between science and the applications of science to practical ends. Science is praised and esteemed because to it are due the wonderful technical developments that we have witnessed in the last century, the railway and the steamship, the motor car and the airplane, telephone and radio, matches, bathrooms, and so on. All this of course is true, and very wonderful and most important. The general standard of life has been enormously raised; every citizen now enjoys, and takes for granted, daily comforts and commodities that would have been undreamt-of luxuries even for the most fastidious millionaire a hundred years ago. If we ask, however, why this increase in our comfort and general standard of life is good, the only answer can be that it gives us more leisure to devote ourselves to those things that give real value to life, the pursuit of science and art. It is as Poincaré said: "*je ne dis pas: la Science est utile, parce qu'elle nous apprend à construire des machines; je dis: les machines sont utiles, parce qu'en travaillant pour nous, elles nous laisseront un jour plus de temps pour faire de la science.*" And, of course, as we have repeatedly pointed out, also the improved technique provides science with more powerful tools, and thereby stimulates its development. It is a significant fact,

proved by history, that all the great technical advances have been based on scientific discoveries which at the time appeared to be utterly useless, and were made by men who studied science for its own sake, without giving the slightest thought to the possibility of application. They are the reward accruing to mankind from the disinterestedness of its greatest representatives.

A moment ago I coupled science and art in one sentence. They are the highest manifestations of the human mind, science of the intellectual side of it, art of the emotional side. There is however much misunderstanding. It happens now and then that those in whom natural gifts or circumstances have developed an exclusively artistic or emotional view of life are filled with pity for the miserable, unimaginative scientists, crawling in the dust of matter-of-fact measurements and calculations, instead of soaring in the pure ether of mystic contemplation and understanding. On the other hand it also happens now and then, but perhaps not so often, that persons with a specially scientific outlook and education look down on the poor unpractical and illogical artists. Both these judgements, of course, are very much mistaken. Science and art approach the great problems of the understanding of nature each in its own way, but both require, and use, the full attributes of the human mind. Imagination is as indispensable for the physicist or astronomer as for the poet; logic is as necessary for the architect or the musician as for the mathematician.

The great men of science, as well as the great artists, are filled with a spirit of reverence, with a consciousness of the presence of mystery and sublimity in the simplest and small-

est as well as the greatest of things and phenomena, and with faith in the order and unity of all things. Only the way in which they try to comprehend this order and penetrate to its deeper meaning is different.

It is a rather common misapprehension that science, by analysing and dissecting nature, by subjecting it to the rigorous rule of mathematical formulae and numerical expression, would lose the sense of its beauty and sublimity. On the contrary, even the purely aesthetical appreciation, say of a landscape, or of a thunderstorm, is, in my opinion, helped rather than impeded by the knowledge, so far as it goes, that the scientific beholder has of the inner structure and the connection of the phenomena. And the measurement and reduction to numbers, "pointer readings," as Sir Arthur Eddington says, is not the ultimate aim of science, but its means to an end. By the use of mathematics, that most nearly perfect and most immaterial tool of the human mind, we try to transcend as much as possible the limitations imposed by our finiteness and materiality, and to penetrate ever nearer to the understanding of the mysterious unity of the Kosmos.

