Why does the moon have no perceptible atmosphere? If the moon originated from the earth, we might expect that it had got its share of air. Can it have had air and later lost it? Such a loss of atmosphere could be conceived in terms of the kinetic gas theory. In 1870 Johnstone Stoney argued that a small percentage of gas molecules, which have a velocity much larger than the average, might escape into space, so that gradually such a gas would disappear entirely from the attracting body. The limiting velocity for particles to escape from the earth is 11 km./sec., from the moon it is 2.4 km./sec., whereas at 0° C. the mean velocity of the molecules of hydrogen is 1.6 km./sec., of water vapour 0.53, of oxygen 0.4. That hydrogen is absent in the earth's atmosphere proves that a velocity of escape only seven times the mean velocity allows for its disappearance. That water vapour is present proves that a ratio of 21 is too large to allow it. Later this theory was refined by Jeans, Milne and others. Applied to the moon, these data tell that all water vapour and all oxygen from the moon must have diffused into empty space.

For the astronomers there simply remained the task of making a careful topography of the moon. A fascinating programme to observe another world from far away and to picture it! A small telescope already shows such an immense wealth of minute details that a total picturing would require many years of patient and devoted efforts. First, in 1824, came the work of Lohrmann, himself a land surveyor; so, instead of publishing the original drawings as Schroeter had done, he worked them into topographical maps, in which the mountains were represented as on geographical maps. Only a few maps were published: eye disease prevented the observer from completing them. In 1828 a Berlin banker, Wilhelm Beer, and a young astronomer, J. H. Mädler, began their common study of the moon in a small observatory erected by the former, provided with a Munich telescope of 4 inch (108 mm.) aperture. Following the example of Tobias Mayer on a larger scale, they first laid down a network of precisely determined points as a basis for the topography; with the micrometer the positions of over a hundred well-identified primary and many more secondary points were measured, as well as the heights of a thousand mountain tops by means of their shadows. They formed the frame of topographical maps picturing the entire surface of the lunar disc on the scale of one metre in diameter, in four sheets (plate 11). It was published in 1836 after eight years' work in 600 nights of observation. Throughout all this work they saw no change in the mountainous formations, nor any change suggesting atmospheric phenomena. Stark and rigid stood the lunar surface, with its pitch-black shadows, a dead world of rock and stone.

Though Beer and Mädler's map of the moon was far more detailed and complete than any former representation, it did not contain more than was visible through any small telescope. So the task remained for astronomers and amateurs to make more detailed studies with their larger telescopes and to publish them as more detailed maps. Outstanding was J. F. Julius Schmidt (1825–84), who in Athens, with small instrumental means but a fine climate, applied himself most diligently to all kinds of observing work that is often left to amateurs. His drawings of the moon were sufficiently complete to combine them into a map of the moon two metres in diameter, which was published in 1878 (plate 11). Neison in England in 1876 published a book on the moon, with extensive descriptions and a number of detailed maps.

A solid basis was thus laid to deal with the question of whether small changes did occur on the moon. Julius Schmidt himself, in 1866, pointed out one case of unquestionable change: the small crater Linné, which formerly had been drawn by Lohrmann, by Beer and Mädler, and by himself, had disappeared, and in its place a large whitish spot or a shallow hole was seen. Other more dubious assertions of changes were published occasionally, resulting in much controversy, because it could never be decided whether the maps were sufficiently reliable in such details and especially whether an object missing on earlier maps was really new or had been overlooked.

For a body so rich in detail as the moon, photography meant invaluable progress. A single photograph picturing the entire disc at once replaced hundreds of drawings that would have taken months and years at the telescope; moreover it was trustworthy as a document. The first good photographic picture of the moon was made in 1850 by W. C. Bond, the director of Harvard College Observatory. This was soon followed by one by Warren de la Rue in England, who, from 1852 on, worked with a reflecting telescope provided in 1857 with clockwork motion. He obtained pictures of 28 mm. diameter, which were so sharp that they could be enlarged 20 times. Henry Draper in America, working with a self-made reflector, in 1863 took moon pictures of 32 mm. diameter which could be enlarged to 90 cm. All this was experimentation and could not increase our knowledge of the moon. Great progress was brought about here by the introduction of dry bromide plates in 1871, which reduced the exposure time to a few seconds or less. The exposure could be short because the pictures were taken in the focal plane of the objective, without enlargement, which would have diminished the brightness too much. Progress was now achieved by using long-focus telescopes. The great 36 inch refractor of the Lick Observatory, with its 57 foot focus, supplied moon photographs of 14 cm. diameter. The Paris *equatorial coudé* (2-elbow-telescope) was regularly used by Loewy and Puiseux after 1894 to take moon pictures of 18 cm. diameter. Enlargements of special parts of these negatives were reproduced and published, with the Lick pictures on a scale of 1 metre, with the Paris atlas on a scale of 2.6 metres to the moon's diameter. At this size the limit was reached at which the silver grain precluded the visibility of finer detail. More delicate detail was reached by photographs made with special care with the 40 inch refractor of Yerkes Observatory, and with the 100 inch reflector at Mount Wilson. They were smaller in number and did not cover all the phases of the moon.

A photographic atlas of the moon differs from a visual atlas, in that it gives the direct aspect of the moment with all its shadows; it is not a topographic map constructed by the astronomer out of a number of drawings at different phases. For every landscape, instead of one map, an entire series of reproductions at different phases is needed, without ever giving completeness. Moreover, in comparing different representations of the same landscape (as, for instance, the crater Hyginus with surroundings, plate 11), we see that the photographs cannot compete in wealth of minute detail with the visual work of careful observers with smaller telescopes. Such was the work of Ph. Fauth, an amateur at Landstuhl, who since 1895 has published many maps of special lunar landscapes. This experience showed that visual work should not be abandoned; many amateurs with good telescopes, up to 4 or 6 inch aperture, continued their study of the details of special objects, chiefly to check the occurrence of small changes. Conversely, photography keeps its place because of its documentary value. Observers, like Krieger (in his maps of 1898–1912) and Goodacre, often made use of the photographs as basic work, in which to fill all further detail at the right spot.

The development of lunar topography in the nineteenth century was confronted with two problems. The first was: Is the moon's surface indeed rigid and invariable, or does anything happen there? In other words: What is the amount of reality and what is the character of the observed small differences appearing sometimes in a landscape invariable in its gross features? A close collaboration between different visual observers was necessary to answer this question. After-effects of earlier lunar forces might possibly produce or wipe out craterlets, small holes or hills. W. H. Pickering later reported narrow haze or cloud tracks and snowlike spots around craterlets, waxing and decreasing with the amount of sunshine, and ascribed them to thin water-vapour exhalations from rifts in the soil. He also paid close attention to the dark parts seen within some circular walls at full moon, when there are no shadows. He described them as growing and variable 'colouring', i.e. darkening of parts of the surface, followed later by bleaching. Photographs confirmed the changes in their general character, but, because of their coarse grain, could not test the minute details. To explain them, Pickering assumed a primitive, but crowding form of life, such as low vegetation performing its life-history of growing and fading in the 14 days of sunshine, fed by carbon dioxide coming from rock fissures; and perhaps herds of freely moving primitive organisms. 'We find here next door to us a living world, with life entirely different from anything on our own planet.'[198] This was his conclusion in 1921. According to these reports, studied in minute detail, the moon would not be the entirely dead world that at first sight it appeared to be.

The second problem relates to the origin of all these remarkable features, so different from those on earth, such as the great circular walls, craters, rifts, mountains and large plains (called 'seas'), that constitute the solid lunar surface. It is a problem of geology, here more correctly to be called 'selenology'. The essential problem is why the active geological forces on the moon produced forms so entirely different from those on earth. The answer was sought in two different directions. In 1873 Proctor suggested that the lunar craters had been produced by the impact of large meteorites breaking through the thin crust, after which strong lava waves built the surrounding circular walls. Similar impacts upon the earth must have been weakened by the resistance of the air and their effects were effaced by later climatic and organic influences. In later times the same idea was expressed by the well-known geographer, Kurt Wegener. More generally, however, the origin is sought in the inner forces of the moon itself; physicists and astronomers often tried to imitate the formation of its surface by experiments with molten matter. In connection with the Paris photographs of the moon, Puiseux in 1896 gave an elaborate theory of the formation of the lunar features. As had been done before by the physicist Ebert, he appealed to the tides in the body of the moon, generated by the earth, as the primary determining forces. Because of the larger mass of the earth, these were far greater than those generated by the moon in the earth. This could explain why the effects in both cases are so different.

When in the nineteenth century interest in the moon lessened because it was uninhabitable, attention turned mainly to Mars as presenting the best chances in this respect. The study of the planet's surface, however, lies on quite a different level from that of the moon; it is more difficult and more restricted. In its most favourable perihelion oppositions, Mars presents a disc of 26″ diameter only, where 1″ represents 260 km. What at the end of the eighteenth century was seen in the reflecting telescopes by Herschel and Schroeter, besides the sharply-cut white polar spots, were only vague, ill-defined shades, wherein clear shapes could rarely be recognized. They were taken to be mostly fogs and clouds. With the introduction of the Munich telescopes, another epoch opened. In 1830, with a favourable autumn opposition, Beer and Mädler began to make drawings with their four-inch. They were the first to recognize a definite picture of darker and lighter parts. The difficult nature of the work is evident from their words: 'Usually some time had to elapse before the indefinite vague mass at first seen dissolved itself into clearly distinguishable forms.'[199] Since these forms periodically came back in nearly the same characteristic aspects, it was evident that they were not passing cloud formations but topographical markings on the planet's surface. For the first time, as a result of their work there appeared a world map of Mars in two hemispheres. In order to include the north polar regions, it was necessary to extend the observations over additional oppositions, from 1832 to 1839; because Mars in aphelion shows a small disc, they made use, from 1835 on, of the 9 inch refractor of the Berlin Observatory. Now that the markings could be identified the time of rotation of Mars could be derived by comparing their earlier and later drawings. From the observations in 1830 and 1832 it was found to be $24^{h}\ 37^{m}\ 23.7^{s}$.

Thus a foundation had been laid for later progress. Many observers made and published drawings of Mars (plate 12); their often very dissimilar aspect shows how difficult was the precise delineation of the markings and how strongly dependent it was on personal style. The opposition of 1862 showed some progress; the drawings of Lockyer in England and of Kaiser at Leiden—the latter condensed into a globe of Mars—both working with a 6 inch telescope, show more details. Secchi at Rome, in 1858 and after, working with a 9 inch telescope, made drawings in colour, which show subtle hues of green in the dark parts and of yellow in the light regions. Kaiser identified ancient drawings of Huygens and Hooke (from 1666 and 1667) with some main features of his globe and derived a rotation time of $24^{h}\ 37^{m}\ 22.6^{s}$ to $\frac{1}{10}$ second.

All these new observations showed that, on the basis of a fixed and permanent topography, a number of minor variations of detail occurred. Sometimes details were washed out by hazy cloudlike appearances; dark spots bleached, or light spots temporarily darkened, and also the precise figures and borders showed changes not entirely attributable to different directions of view. The observers generally agreed in calling the yellow or ruddy parts land or desert, and the dark, mostly greenish, hues water, perhaps shallows or vegetation. They assumed the white polar caps, which changed in size periodically with the seasons, to be snow and ice. From these data it was concluded that life-conditions on Mars are not greatly different from those on earth: the temperature

somewhat lower, the atmosphere thinner and drier, the water scarcer, conceivable as a further-developed stage in planetary evolution.

The perihelion opposition of 1877 brought new progress and a new surprise too: the discovery of the canals of Mars. Schiaparelli (1835–1919), working at Milan with a 9 inch refractor, at once raised areographic research to a higher level by making micrometric measurements, first of the small southern polar cap in order to fix precisely the position of the rotation axis, and then of 62 well-identified points as a basic network. He perceived long dark lines, narrow and straight, chiefly traversing the yellow northern half of the planet. He called them 'canals', without asserting them to be water; the scale of a planetary image is so small that the finest visible line represents a width of many miles, and some of these 'canals' were broader than that and represented considerable actual width.

In the 1879 and later oppositions these observations were continued, confirmed, and extended. Then, from 1881 on, a new phenomenon was recorded, the doubling of the canals. From then on, Schiaparelli's results were an object of continual controversy on the part of his colleagues. Some of the drawings of other astronomers began to show, hesitatingly, a few canals. While the reality of straight canals did not seem impossible, for their doubling no natural cause could be imagined. So the habitability of the planet from a mere fancy became a practical argument of discussion. Is Mars inhabited by intelligent beings, and are we justified in invoking their intelligence as a causal explanation of the phenomena? In that case even the doubling could be made understandable as deriving from the need for navigation or irrigation. Yet, with regard to the doubling, the general attitude was mostly sceptical. Perrotin of Nice and his assistant Thollon were the only ones who succeeded, in 1886, in seeing double canals with the 15 inch refractor. Yet it was curious that, at first, these observers saw nothing of this for weeks and then at last produced drawings entirely similar to Schiaparelli's. The question was posed as to how far bias could influence the delineation of objects at the limit of visibility. Or was it possible that the physiological effect of a minute, unsuspected astigmatism in the eye had played a role? Against these criticisms, Schiaparelli stuck to the reality of the forms he had reported. In 1892 he published a survey of his researches in the German review *Himmel und Erde*, accompanied by a coloured map of Mars in two hemispheres, so rich in detail that one could not fail to be impressed by the great progress of our knowledge of the planet. However, a second map with all the double canals threw the reader back into doubts about their reality and significance.

It was at this time that the giant 36 inch refractor was installed at the Lick Observatory. The Lick astronomers, Holden and Keeler, in 1888 directed it on Mars, expecting that their powerful instrument would solve the doubts. It was a disappointment, not so much because their drawings showed nothing of the canals and their duplication, but chiefly because they were so poor in detail that they looked like products of an earlier stage of research. It had often been asserted that the great telescopes are not superior but rather inferior to small ones for the study of planetary surfaces, as might be seen by comparing Beer and Mädler, Lockyer and Kaiser, with Herschel, Schroeter and Lord Rosse. Faint differences in shade, extending over wide surfaces, are more difficult to distinguish than when contracted to small dimensions. Moreover, air striae in the broad bundle of light entering a large objective will smooth the image more than in the narrower pencil of a smaller instrument. On the other hand, intricate complexes of objects cannot be distinguished by low powers and are conceived in simplified forms. In the case of the Lick astronomers, lack of practice certainly played a role. The observation of Mars was to them a short interruption in important other work on the stars, whereas only by continuous and assiduous occupation with the planet's surface as his sole object can the observer succeed in discovering all the minute details—though at the same time he clings more firmly to a personal style in interpreting and drawing them.

The problem of the possible existence of Martians exerted its fascination. The well-known French popularizer of astronomy, Camille Flammarion, in 1892 published a large monograph, *La planète Mars et ses conditions d'habitabilité*; the title indicated and the contents clearly expressed his conviction that next to the earth there travels a brother-world, with life-phenomena of its own and of a nature as yet unknown to us. He wrote: '... the considerable variations observed in the network of waterways testify that this planet is the seat of an energetic vitality. These movements seem to us to take place silently because of the great distance separating us; but while we quietly observe these continents and seas slowly carried across our vision by the planet's axial rotation, and wonder on which of these shores life would be most pleasant to live, there might at the same moment be thunderstorms, volcanoes, tempests, social upheavals and all kinds of struggle for life. . . . Yet we may hope that, because the world of Mars is older than ours, mankind there will be more advanced and wiser. No doubt it is the work and noise of peace that for centuries has animated this neighbouring home.'[200] He further says: 'The present inhabitation of Mars by a race superior to ours is very probable.'[201] Then he indulges in praise of the beauty and greatness of the conquests of modern astronomy: 'It is the first time since the origin of mankind that we have discovered in the heavens a new world sufficiently like the Earth to awaken our sympathies';[202] and he thinks of a far distant future in which the peoples of both planets, united, will

proceed to greater common works. He wrote: 'The Earth becomes a province of the Universe,' and 'we feel that unknown brothers are living there in other fatherlands of Infinity.'

For the present, mankind on earth was restricted to a lot of stories about the Martians and to the assiduous and continuous study of science. Percival Lowell, a wealthy New England aristocrat, enthusiastic, gifted and well-versed in many fields, in 1894 founded an observatory at Flagstaff, in the bright desert climate of Arizona, at a height of 6,000 feet, especially to study Mars and its inhabitants. He used first an 18 inch Clark refractor, later one of 24 inches. A number of drawings, with about 180 mostly new canals discovered by him and his assistants, was published. He emphasized the changes in the aspect of continents, seas and canals, varying periodically with the Martian seasons, and explained them by the yearly movement of the water to and fro, from the melting towards the growing ice cap. The canals, serving as expedients for irrigation to ensure the most economical use of the scanty water supply, disclose control by intelligent beings.

More fruitful and dependable was the work of the French astronomer Antoniadi, who from 1909 onwards made studies of Mars with the 83 cm. (33 inch) refractor of the astrophysical observatory at Meudon, near Paris. With this instrument, surpassing in aperture all European refractors, and in a climate distinguished by steady telescopic images, Antoniadi was able to demonstrate that under such conditions and in the hands of a good and trained observer, large instruments surpass the smaller ones for the study of the planets. What Schiaparelli had drawn as straight or double canals was now dissolved into minute details of small spots, sometimes arranged in regular series, also often irregularly, or simply boundaries of regions of different shades. For twenty years he was engaged in drawing and describing the features of Mars; his work was notable progress and enrichment after Schiaparelli, just as the latter had surpassed his predecessors. Antoniadi sometimes saw white clouds, which he took to be composed of water, and at other times yellow dust clouds, obliterating the markings; but the prominent phenomena to him were the colouring and discolouring according to the season—the dark greenish or bluish parts, probably the lowest regions, which by drying up in summer became yellow, brown or red, each in its own manner. The movement of the water from pole to pole, here melting and evaporating, there condensing and freezing, he likewise considered to be the basis of the visible changes. Antoniadi's work opened the way to further progress by showing how the increase in telescope sizes and the observers' skill can advance the study of the planets if aided by perfect atmospheric conditions.

Could not photography also open new ways here? At the close of the nineteenth century many photographs of planetary surfaces were made; but what they presented was at most a blurred copy of the drawings. Taken directly in the chief focus, the images were very small, of some few millimetres in extent, and the grain coarsened the details. In recording enlarged images, the details, owing to longer exposure, were effaced by atmospheric vibrations, whereas an observer's eye could wait for the short still moments when minute details stand out sharply. Photographs of Mars taken by Barnard with the 40 inch telescope in 1909, by Slipher at the Lowell Observatory in 1922, by Wright and Trumpler with the 36 inch Lick telescope in 1924, and by Ross with the 60 inch telescope at Mount Wilson, clearly show the main shape of the dark markings, as well as the polar caps. Of the slender lines of the canals nothing of course was visible. Yet in other respects these photographs were most instructive. Wright took photographs on plates sensitized for limited ranges of colour. The images in infra-red and red light showed the surface topography clearly, though this was totally absent from the plates sensitive to violet and ultra-violet light; besides the white polar spots, these showed vague shadows only. By comparison with earthly landscapes, Wright could establish that this was an effect of the Martian atmosphere, entirely transparent to red light, while reflecting blue light diffusely. This reflected light altogether exceeded the blue light from deeper layers.

Remarkable progress was reached in recent years by an ingenious method by Lyot to eliminate the grain in the photographic emulsion. By rapidly taking a number of small focal images and combining them in one enlarged picture, the grains are smoothed into a continuous background. These figures show almost the same amount of detail as the best drawings; it may be expected that they will surpass them after further improvement has been made to the new method.

In the nineteenth century the other planets gave rise to much less sensational discoveries. On Mercury no distinct spots were perceived until Schiaparelli in 1881 saw a system of vague and dim streaks, which always kept the same position relative to the terminator, the boundary between the dark and the illuminated parts of the disc. He concluded that the planet always keeps the same hemisphere towards the sun, as the moon does towards the earth, so that it rotates about its axis in 88 days, its period of orbital revolution. This was later confirmed by Lowell at Flagstaff, but was opposed by Leo Brenner, who worked on Martian canals with a small telescope in the beautiful climate of Lussinpiccolo in Istria. The question was definitely settled by Antoniadi, who in 1924–29 made a large number of drawings with the great Meudon telescope. He confirmed the rotation time of 88 days; his

drawings showed a system of vague broad bands in which he could recognize most of Schiaparelli's streaks. They are fixed markings on the planet's surface, of course visible only on the illuminated half, since the opposite hemisphere always remains dark. Moreover, now and then he saw extensive and variable white patches, which he supposed to be clouds of dust. Water clouds cannot exist on Mercury because its low gravity would cause water vapour to escape; for dust clouds, however, a small amount of atmosphere must be present.

No astronomer observing Venus ever saw anything more than vague and ill-defined differences in shade, not suitable for ascertaining a period of rotation. Schroeter at Lilienthal said he had seen one of the horns of the Venus crescent sometimes blunted—as also a horn of Mercury—which he explained by shadows of high mountains carried around in a rotation period of about 24 hours. Herschel, however, did not see this. In 1839 Vico at Rome thought he recognized definite forms and derived a rotation period of $23^h\ 21^m\ 22^s$, by comparing them with drawings of Bianchini, made a century earlier. On the other hand, Schiaparelli found a rotation time of 225 days, so that Venus, like Mercury, would always turn the same face to the sun. The result raised doubts and contradictions; Villiger at Munich in 1898 pointed out that an entirely smooth matt-white globe illuminated sideways shows gradual differences of shade, which by contrast give the impression of faint markings keeping their place relative to the shadow boundary. Observers always agreed that with Venus we do not see the planet's surface but only the upper side of a dense layer of clouds. That Venus has an atmosphere was established first at the transit of 1761, when the part of the dark disc away from the sun was surrounded by a luminous ring. Moreover in the nineteenth century it was often seen that the horns of the crescent extend over more than 180°.

Observation of the planet Jupiter in the nineteenth century afforded nothing so exciting and new as with Mars or Mercury. It confirmed the chief features already established in the past. They were the two dark belts parallel to and separated by a brighter equatorial zone, all showing small spots and irregularities that indicate a rapid rotation in about 10 hours. Since we see only the upper side of a cloud envelope, the moving spots can only roughly disclose a rotation period; it was found to vary at different latitudes, $9^h\ 55^m$ in the dark belts, $9^h\ 50^m$ in the equatorial zone. Since the big planet shows a wealth of ever changing details in small telescopes, it was an interesting object for astronomers with limited equipment. During the entire nineteenth century, a number of observers assiduously made drawings of the spots in the cloud bands, without learning more about them than that they were steadily changing, appearing and decaying. Larger instruments could give information about the different colours, mostly yellow or ruddy in between greenish parts. The only sensational news was the appearance, in 1878, of a large spindle-shaped red spot; it could be found on drawings as far back as 1859 and was indicated even in 1831. After its discovery it remained visible for many years, gradually losing its red colour, with a changing rotation period of nearly $9^h\,55^m$.

Saturn, with a greater flattening than Jupiter's, hence a certainly rapid rotation, showed far less interesting surface features. There was a faint dark belt along the equator, sometimes accompanied by other fainter belts at higher latitudes. At times irregular spots appeared therein that offered the possibility of determining a period of rotation. Thus in 1876 Asaph Hall found $10^h\,14^m\,24^s$ for the equator; in 1894 Stanley Williams found $10^h\,12^m\,13^s$ at the equator and $10^h\,14^m\,15^s$ above 20° of latitude; whereas in 1903 a spot at 36° latitude, observed by Denning and Barnard, gave $10^h\,38^m$ to 39^m. So here also we see only a cloud layer.

Uranus and Neptune are of course far more difficult objects; the flattened figure (flattening about $\frac{1}{12}$) of the greenish-tinged Uranus indicates a rapid rotation about an axis situated nearly in the ecliptic. Drawings made by Lowell showed some spots, the reality of which is suspect.

As objects equivalent to our moon, the satellites of other planets must be mentioned. Their smallness and resultant lack of atmosphere ruled them out as abodes of life; but they completed the world aspect for their primaries. With the improvement of the telescopes in the nineteenth century, more and more of them were discovered. Venus and Mercury remained without satellites, though deceptive reflex images in the telescope were often announced as moons. In 1877 Asaph Hall, with the Washington 26 inch telescope, discovered two moons of Mars, certainly the smallest known of their kind, with estimated diameters of 9 and 12 miles, visible only because of the proximity of Mars to the earth. The period of revolution of the innermost, $7^h\,29^m$, shorter than the rotation period of the planet, offered a distinct problem in the theory of tidal action.

In 1892 Barnard, with the 36 inch Lick refractor, discovered a fifth moon of Jupiter in close vicinity to the planet, a small object perhaps 100 miles in diameter. Outside the system of the four large Galilean satellites, four small satellites were discovered between 1904 and 1914 on photographic plates; this number was increased later. Owing to their remarkably deviating orbits, strongly inclined and eccentric, some even retrograde, strongly perturbed by the sun, they are interesting objects of celestial mechanics. Thus the question of whether they could be former planetoids captured by Jupiter was posed and answered, in the affirmative for some of them, by Moulton.

To the seven satellites of Saturn known at the end of the eighteenth century an eighth was added in 1848, discovered nearly simultaneously by W. C. Bond with the Harvard refractor and by Lassell with his reflecting telescope. A very small ninth satellite, found photographically in 1898 far outside the realm of the others, resembles the outermost Jupiter satellites in having a retrograde motion and great eccentricity. Uranus, which had two satellites discovered by Herschel, received two more discovered by Lassell in 1851, all very faint objects, with their orbits nearly perpendicular to the ecliptic. It was again Lassell, who, shortly after the discovery of Neptune, discovered its moon, which also has a retrograde motion in a strongly inclined orbit.

With respect to size the satellites join the planets to form one continuous series. The third and fourth Jupiter moons, with a diameter of 5,000 km. (3,200 miles), are equal to Mercury; the Neptune satellite is somewhat larger; our moon and the sixth of Saturn are somewhat smaller; so there may be just a trace of overlapping. Then the others follow in a decreasing series down to the Mars satellites running parallel to the similar series of planetoids. Whereas the preceding centuries distinguished the sun, the planets, and the satellites as three different types of size, the discoveries of the nineteenth century combined the latter types into one series, continued in the still smaller meteorites. The sun as a solitary, quite different, body stands separated by a gap from the ten times smaller Jupiter.

Nineteenth-century astronomy did not restrict its findings to the outer surfaces of the celestial bodies. It could also penetrate into their interior, because it was chiefly a theory of attraction, and attraction proceeded from all the matter in the interior. The first new datum derived was the mean density of this matter. When the masses of the planets relative to the earth's mass are known, whether from the motion of their satellites or, with more difficulty, from the perturbations they exert, measurement of their diameters reveals their volume, hence their mean density, relative to the density of the earth.

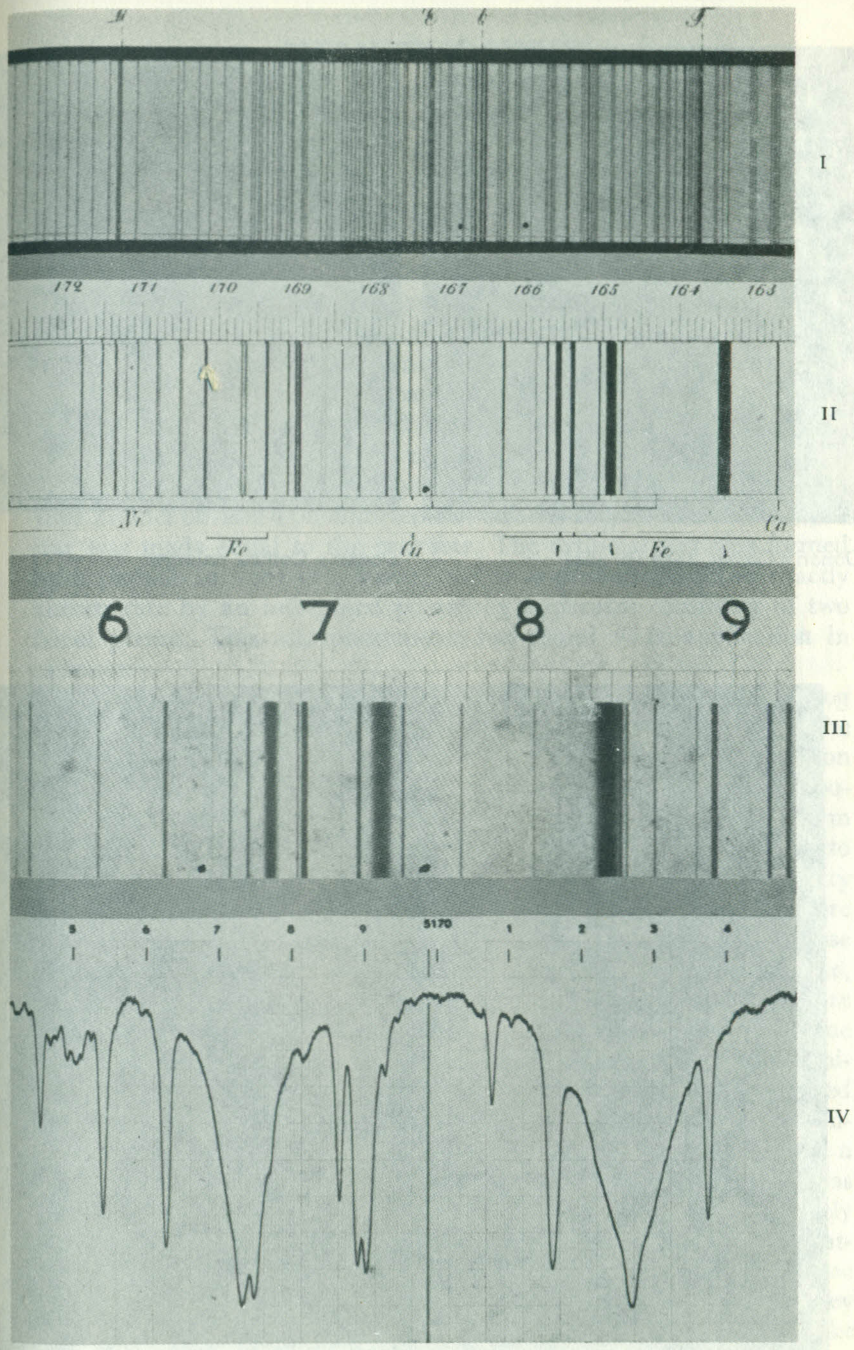
The mass and the mean density of the earth have been the subject of a series of the most subtle investigations extending over the entire nineteenth century, based on the measurement of the hardly perceptible attraction exerted by bodies in laboratories. The result was a mean density of 5.50, larger than the densities of rocks and minerals constituting the crust of the earth. Hence its interior must have a far higher density, such as is found only for metals.

Now the absolute densities of the other celestial bodies could also be stated. For the moon it is 3.3; for Mercury, Venus, and Mars, 3.8, 4.9, and 4.0, of the same order of magnitude but somewhat smaller than with the earth, decreasing parallel with the planet's size. For the major planets the mean density is found to be far smaller: for Jupiter 1.3, for Saturn 0.7, for Uranus 1.3, for Neptune 1.6. That Saturn has a smaller mean density than water can be understood only if a large part of its apparent volume consists of gas. The surface of its solid or fluid body is situated far deeper, and what we see as its surface is the upper side of a cloud mass floating in an extensive atmosphere. The same, in a lesser degree, must hold for the other major planets.

Celestial mechanics could give still more information. Clairaut had derived a formula by which the visible flattening of a planet depended on the period of rotation, combined with the increase in density toward the centre; thus the inequality in the distribution of matter in the interior was revealed by visible phenomena. Celestial mechanics met here with seismology, which had organized all over the earth the exact registering of the small vibrations and seismic waves pervading its body. This branch of geophysics had derived analogous conclusions, but in a different form. Wiechert in 1897 found the earth to consist of two clearly distinguished parts, a metal core, chiefly iron, with a density of 7.8, surrounded by a layer of rock, chiefly silicates of iron and magnesium, with a density of about 3.3. In 1934 Jeffreys expressed the plausible opinion that the same holds for the other earthlike planets and that their decrease in mean density with volume is due to the iron core diminishing in size until it is entirely lacking in our moon. From the flattening of Jupiter and Saturn it was derived that the increase in density with depth was greater here than with the earth, so that, as with Jupiter, a considerable part of the volume beneath the visible cloud surface must be occupied by atmospheric gas.

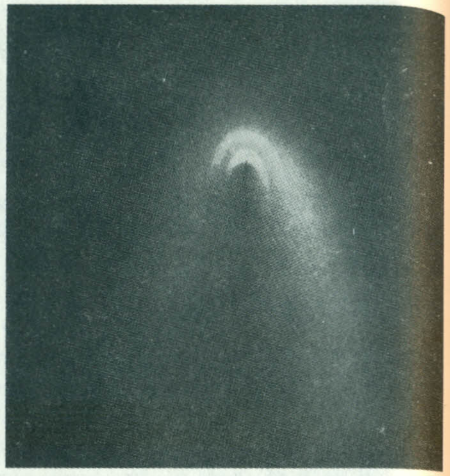
The knowledge about the conditions on other worlds that was acquired merely by using telescopes was not very satisfactory. Happily support came from another source; in this realm of study, astronomy had no longer to rely solely on its own forces. The progress of physics in the nineteenth century provided new instruments and new methods of research that were applied to the celestial bodies with increasing success. They consisted chiefly of photometry, spectrum analysis, thermometry and polarimetry.

The principles and methods of photometry—the measuring of quantities of light—are so simple and obvious that they could have been used in earlier centuries; but interest was lacking, as well as the restless urge to investigate everything. Some first attempts were made in the eighteenth century, by Bouguer, who in 1729 made measurements, and by Lambert, who in 1760 gave a first theory of the diffuse reflection of light by even matt surfaces. The fraction of the incident light reflected diffusely by a surface was called by Lambert its *albedo*, the Latin word



13. Sections of the solar spectrum. I. Fraunhofer; II. Kirchhoff: III. Rowland; IV. Utrecht. The wave-lengths increase to the left on I, and to the right on II, III, and IV. The dots mark the limit of the part reproduced on the next illustration (pp. 406–9)

14. Donati's Comet, 1858 (p. 424)



Donati's Comet, enlarged

Morehouse's Comet, September 29, 1908 (p. 425)

for 'whiteness', since his book was written entirely in Latin; this Latin word has remained the usual scientific term. Practical photometric work in astronomy began with John Herschel, who in 1836, during his stay at the Cape of Good Hope, measured the brightness of a number of stars. He compared the star with the pointlike image of the moon formed by a small lens; by varying the distance between the lens with the light point and the eye, he could make it equal to the star. The most fundamental law of optics then says that the intensity decreases as the inverse square of the distance. Because the results of different days were discordant, he supposed that unknown systematic sources of error had vitiated his measures. The real cause was shown afterwards by J. C. F. Zöllner (1834–82): Herschel had computed the brightness of the moon on different days by means of formulae from Lambert which did not fit here. In reality, Herschel's measurements were quite concordant. In 1861 Zöllner devised a photometer in which the brightness of an artificial star was made equal to the real star. The artificial star was formed by a pinhole in front of a flame, and it is diminished at an exactly known rate by an interposed polarizing apparatus consisting of two Nicol prisms. This astrophotometer has found wide application in astronomy.

The same principle of varying the intensity by the use of polarizing apparatus was used by Zöllner in determining the brightness of the moon in different phases. He discovered that this brightness at full moon had a sharp maximum; before and after the exact moment of opposition the brightness rapidly increased and decreased, almost uniform with the phase-angle (the angle at the moon between the directions to the sun and the earth). According to Lambert's theory, the intensity should show a flat top like a hill, slowly varying near the top and more rapidly only at larger phase angles. Zöllner soon understood the cause of the discrepancy: Lambert's theory was devised for an even surface, whereas the surface of the moon is full of irregularities. Directly before and after opposition, when the phase angle is small, shadows become visible and diminish the light; they occupy a part of the surface, uniformly increasing with this angle. The decrease in brightness (expressed in stellar magnitudes) for 1° of phase angle, the so-called 'phase coefficient', for which Zöllner found the value 0.025, may be taken as a measure of the irregularity of the surface. Herschel's results, now used as measurements of the moon's brightness by means of known stars, entirely confirmed Zöllner's results. In 1923 Barabashev measured the brightness of apparently flat parts of the lunar surface; they showed the same dependence on phase angle as the total light, an indication that they were full of small irregularities, holes and knobs, and might be like pumice stone.

The comparison of the brightness of the full moon and of the sun was far more difficult, because the difference is so enormous. Zöllner found a ratio of 618,000. G. P. Bond shortly before had found 470,000 by a different method. Several physicists and astronomers afterwards found widely varying values, and the usually adopted average of 465,000 is uncertain by a tenth of its amount. For the albedo of the moon, Zöllner derived the value 0.17. Later researches resulted in a lower value, 0.07, corresponding to rather darkly coloured stone.

Zöllner, the founder of astrophotometry, also measured in 1862–64 the light ratio of the full moon and some bright stars, as well as the brightness of the planets. He detected some main characteristics; for Mars he found a considerable phase coefficient; for Jupiter, none. From the brightness of the planets relative to the sun, he derived an albedo of 0.27 for Mars, 0.62 for Jupiter, 0.50 for Saturn, 0.64 for Uranus, and 0.47 for Neptune, hence a rather strong reflective power for the major planets.

The Zöllner photometer was used from 1877 to 1893 by Gustav Müller and Paul Kempf at the Astrophysical Observatory at Potsdam, in extensive accurate measurements of the brightness of the planets. They may be condensed in the results for the two characteristic quantities, the phase coefficient and the albedo. For Mercury the former is 0.037 (larger than for the moon), the latter 0.07; for Venus they are 0.013 and 0.59. In these simple numbers the difference in character between the two planets is more distinctly expressed than by any description. The large phase coefficient and the small albedo for Mercury indicate a surface directly seen, unimpaired by an atmosphere, similar to the moon, dark rock with shadow-producing inequalities. The high reflective power of Venus, which makes it our brilliant evening star, indicates that we do not see its solid surface but the white upper side of a cloud layer; this is confirmed by the small phase coefficient. For Mars, Müller and Kempf found a small albedo, indicating that we see its not very dark surface, and a small phase coefficient, indicating that this surface is rather smooth, without large irregularities. The large albedo 0.56 and the small phase coefficient 0.015 of Jupiter show the same conditions as for Venus. For Saturn and Uranus it is difficult, owing to the small phase, to determine a phase coefficient, but the large albedo indicates similar conditions.

For the ring of Saturn photometric measurements have been of special importance. Laplace had demonstrated that flat solid rings could not keep in equilibrium when freely floating in space. In 1859 the famous physicist James Clark Maxwell, the founder of the theory of electricity, dealing with the same problem, pointed out that the ring would be broken by the tensions of attraction and centrifugal force. He suggested that Saturn's rings consist of numerous small bodies which, like a ring of meteorites, freely circle the planet. This theory was sustained by the rift observed between the outer and inner ring, which Kirkwood explained in the same way as the gaps in the ring of planetoids, by a commensurability with the motion of the largest among Saturn's satellites. Another confirmation consisted in the faint 'crape ring' discovered by Bond in 1850, a semi-transparent extension of the ring beyond its inner border.

Müller's photometric measures of Saturn showed that when the ring was invisible, Saturn had a very small phase coefficient. When, however, the ring was widely open, the total brightness increased to $\frac{8}{3}$ times the former amount (the ring giving more light than the ball), and now the phase coefficient rose to 0.044, due to the ring; so that, for the ring alone, it must be still greater. Seeliger at Munich gave an explanation similar to Zöllner's for the moon. If we look at the particles of the ring from exactly the direction from which the solar rays are coming, every particle covers its own shadow upon the particles behind. However, as soon as the two directions deviate by a small angle, the borders of the shadows will appear. Thus, the photometric measurements prove that Saturn's rings do indeed consist of small meteoric bodies.

Turning now to the other physical methods named above, we note that polarimetry is a special refined extension of photometry. The rate of polarization of the diffusely reflected light, positive (parallel to the plane of incidence) or negative (perpendicular to that plane), depends for different substances in a different way on the angle between incident and reflected light (called 'phase angle' above). This method of research could acquire practical importance only when instruments for its measurement became sufficiently accurate. Bernard Lyot in Paris succeeded in improving them so much that the rate of polarization could be measured to 0.1 per cent. In 1924 and thereafter he measured the polarization of the light reflected by the planets and the moon in its dependence on the phase angle. For the moon this dependence had a special irregular course, partly positive, partly negative; of all earthly matter investigated, it was volcanic ash only that showed exactly the same dependence; this gave direct information about the character of the lunar surface.

For Mercury the curve representing this dependence was of exactly the same shape. For Venus it was entirely different; nor did it wholly agree with the light of white clouds. It corresponded rather to light reflected by a fine haze, and it was even possible to recognize the size of the droplets. So looking at Venus we do not see the upper side of a cloud layer but a deep atmosphere, non-transparent, through a thin whitish haze. For Mars the polarimetric curve had some likeness to that of the moon, but it fitted a sand surface better. With Mars, variations and disturbances occur for many weeks, bearing the character of dust clouds. Dust storms from visual observations had also been supposed to occur, indicating a dry atmosphere. The study with polarimeters has in this way afforded with unexpected certitude important knowledge of the physical nature of planetary surfaces. For Jupiter and Saturn, with their small phase angles, only a difference between the equatorial and polar regions could be discerned.

Measuring the heat of distant bodies was always far more difficult than measuring their light, because man does not possess for heat perception so sensitive an organ as is the eye for light perception. It was not until the nineteenth century that physical instruments were constructed of such sensitivity that the heat of celestial bodies other than the sun could be measured. The thermopile of Melloni was applied to the moon by its inventor in 1846, and in a larger investigation in 1869–72 by Lord Rosse. Part of the radiation received from the moon is reflected solar radiation, which consists mainly of visual wavelengths below 0.7 micron. The other part is direct radiation of the heated lunar surface, which, as low-temperature radiation, consists chiefly in long waves above 1 micron. By interposing a sheet of water, which entirely absorbs the long waves, the two can be separated, so that the temperature of the moon can be derived from the intensity of the long-wave radiation. Lord Rosse found that 14 per cent of the lunar radiation was reflected solar radiation, 86 per cent was moon radiation proper; and he derived that the temperature of the illuminated moon surface was 300° C above that of the parts in the dark. His results were confirmed in 1874 by Very's careful measurements at the Allegheny Observatory; in full sunshine the moon's surface was hotter than 100° C; in the dark it fell to little above absolute zero. In 1886 Boeddicker, assistant to Lord Rosse, observed that during a total eclipse of the moon the temperature fell rapidly to far below zero; it shows that the heat is absorbed superficially and is rapidly lost.

The use of Crookes's sensitive radiometer and the rapidly increased sensitivity of modern thermo-elements made it possible, by putting them in the focus of giant telescopes, to measure temperatures of the planets. By interposing sheets of different absorbents like glass and fluorite, it was possible to make a more detailed division in the dark radiation between longer and shorter wavelengths, in order to derive the temperature. Nicholson and Pettit at Mount Wilson Observatory observed the lunar eclipse of 1927 with a radiometer and found that the temperature fell from 70° C to −120° C at the end of the eclipse. From the measurements of Mars made by Coblentz and Lampland in 1926 at the Lowell

Observatory, Menzel derived that the temperature at noon is about, or some degrees above, 0°, the dark parts somewhat warmer than the yellow plains. In the morning the temperature rapidly rises from −100° C; the southern pole, when emerging from the darkness of the winter night, is −100° and in the course of the long summer day rises to above 0°. Former theories that Mars, as a result of the twice times smaller solar radiation, should be covered with a thick sheet of ice, the cracks in which appear to us as canals, are refuted by these measurements. Our earth after all seems to be the better place to live.

For Mercury a temperature of 400° C was measured. The illuminated part of Venus showed a temperature of 60° C and the dark part −20°; the rather high value of the latter shows that it cannot be permanently devoid of sunshine, so that the rotation period cannot be as long as 225 days. The temperature of Jupiter was unexpectedly as low as −130° C; for Saturn it was −150° C; and for Uranus it was still lower, −200°. Because the rapid changes in the cloudy features on Jupiter suggested turbulent processes in its atmosphere, it had often been supposed that the surface of the planet below was hot and, because of its size, had not yet sufficiently cooled. Jeffreys in 1923, assuming radioactive matter to be the source of the observed interior heat of the planets, had expressed his doubts of this idea. The temperatures measured agree with the supposition that the solar radiation, weak on account of the great distance, is their sole effective source. Here, as well as in the case of Venus, we have to remember that the temperatures thus measured hold for the highest layers of the atmospheres, the outer surface of the clouds.

Spectrum analysis was the most powerful aid that physics offered to astronomy in the nineteenth century. Soon after its discovery it was applied to the stars and the planets, first by Huggins in England and Secchi in Rome, soon followed by Rutherfurd in America and Vogel in Germany. What information could the spectra of the planets afford? Their light is reflected sunlight that has twice passed through their atmospheres; so their spectra must show the same Fraunhofer lines as the solar spectrum, but increased by the planets' atmospheric absorptions. If the reflecting surface is coloured, the distribution of intensity over the different colours will deviate from the pure solar spectrum. When accurate determination of wavelengths later became possible (from about 1890), chiefly by the use of photography, velocities in the line of sight, e.g. velocities of rotation, could be measured.

Such was the case with Jupiter. If the slit of the spectrograph is placed on Jupiter's disc along its equator, its eastern and approaching limb has the Fraunhofer lines displaced towards the violet, its western and receding limb towards the red; the lines are inclined, and this inclination is a measure of the rotation period. Keeler, at Lick Observatory in 1895, made spectrograms of Saturn in this way; the spectrum of the disc showed such inclined lines, indicating an equatorial velocity of 10 km. per sec., a confirmation of the rotation period of $10\frac{1}{4}$ hours. At both sides of this disc spectrum, two bands showed the spectrum of the ring; they had the absorption lines inclined in the opposite direction. This corresponded exactly with the velocities of small satellites describing free orbits about the planet: 20 km. at the inner border, 16 km. at the slower outer one. Thus the spectrum of Saturn afforded new confirmation of Maxwell's meteoric theory, more directly conspicuous than the photometric proof.

These were not new results. For Uranus, Slipher at the Lowell Observatory in 1911 could determine the unknown rotation period of 10.7 hours from spectrograms; this was in accord with the strong flattening of $\frac{1}{12}$ derived from measures of the diameter. Astronomers hoped also that the spectrum of Venus would determine whether the rotation took place in about one day or in 225 days. It was a delicate problem, because in the case of the one-day rotation the equatorial velocity, as with the equal-sized earth, is only 0.4 km. Slipher at the Lowell Observatory, by careful study of spectrograms taken in 1903, could not find any trace of an inclination of the spectral lines; this meant that the velocity could be at most 0.02 km. or that the rotation period was at least some weeks and certainly not in the region of a day.

For the moon, which has no atmosphere, the only difference from the solar spectrum to be expected would be due to the colour of the surface. Wilsing at Potsdam, by spectrophotometric measurements, i.e. by measuring the relative intensity of different colours in the moon's and the sun's spectra, found that the light of the moon is more yellowish, as light reflected by ash or by sand.

For the planets, the main object of study was the constitution of their atmospheres. The absorption of the earth's atmosphere produces in the observed solar spectrum and in all celestial spectra three strong absorption bands in the red, denoted by the letters *A*, *B*, and α, produced by oxygen, and many other groups produced by water vapour. If oxygen or water vapour exists in a planetary atmosphere, it must show in the spectrum as an enhancement of these bands. On the question as to whether this is the case in the spectra of Mars and Venus, there has been much controversy. It was an extremely difficult observation not only to see whether such bands could be ascertained and measured in the faint, hardly visible red but, moreover, to decide whether they were stronger than in the moon's spectrum. In the 'seventies the observers Huggins, Maunder (at Greenwich), Janssen (at Meudon) and Vogel (at Bothkamp) agreed that oxygen and water vapour were present in the atmospheres of Mars and Venus; and in 1894 they confirmed their former conclusion. Certainly some preconceived opinion of a natural similarity in the planetary atmospheres must have played a role. Gradually, when better telescopes and spectroscopes were used with greater resolving power, doubts arose. In 1894 Campbell stated that with the powerful Lick instruments he could not perceive any difference between the bands in the spectrum of the planet and of the moon when they stood at equal altitude.

Not until the spectra could be photographed on plates sensitized to red light could accurate decisions be made. Even then they were not easy. When Venus in greatest elongation is approaching or receding from the earth, the series of lines constituting these bands must be displaced relative to the same lines produced by the earth's atmosphere. Slipher could find no trace of doubling or displacement of these lines in the Venus spectra taken in 1908. Clearly, oxygen and water vapour are not present in the highest atmospheric layers of this planet. At Mount Wilson this result was confirmed; here another group of absorption lines was found which appeared to be due to carbon dioxide. For Mars the decision proved to be far more difficult. Slipher in 1908 had taken Mars spectra also; by measuring the intensity of the bands photometrically, Very found that the oxygen bands were 15 per cent stronger than in the moon's spectrum and that there was more water vapour than in the dry Flagstaff air. Campbell however, in 1910, could find no difference from the moon. In 1925 Adams and St John at Mount Wilson took plates with large dispersion to see whether the separate lines of these bands were displaced in comparison to the intermingled true solar metallic lines. Their result was that water vapour in the Martian atmosphere was 3 per cent of the Mount Wilson atmosphere, and oxygen was 16 per cent. Thus the question seemed to be settled. When, however, new sensitizers had been discovered, by means of which the oxygen band *B* could be photographed with great dispersion, Adams and Dunham in 1933 made a new study of the problem; now the absence of any trace of Martian oxygen could be stated with certainty. The quantity of free oxygen in the atmosphere of Mars must be less than 1 per cent, perhaps even less than 0.1 per cent of what it is on earth. A trace of carbon dioxide was found by G. P. Kuiper in the Martian atmosphere. So the result of all this atmospheric research is that neither Mars nor Venus is inhabitable for beings whose life-energy depends on the use of oxygen.

What, then, is the meaning of the greenish colour of the dark parts of Mars, called 'seas' and 'canals' but interpreted as vegetation? Kuiper succeeded in photographing their spectrum and found the infra-red absorption bands characteristic of cholorophyll were absent; as this is also the case with some mosses and lichens on the earth, a very primitive order of vegetation on Mars is not excluded.

The spectral investigation of the major planets brought results no less surprising. In the 'sixties Secchi and Huggins had already perceived in the Jupiter spectrum an absorption band in the red, wavelength 6180, not known from any substance on earth. In the 'seventies besides this one a number of other unknown dark bands in the red and yellow were perceived. Real progress came when modern sensitizers had opened to our view the extreme red and infra-red. Slipher then established that all major planets exhibited the same absorption bands, in such a way that they are faintest with Jupiter and increase in intensity with increasing distance from the sun. Their origin remained unknown until, in 1932, Rupert Wildt at Göttingen pointed out that a number of them were present in the spectrum of methane, the simplest of hydrocarbons, and others were due to absorption by ammonia. It was confirmed by exact comparisons made by Dunham at Mount Wilson. H. N. Russell deduced from a theoretical discussion of chemical equilibrium, that in the absence of oxygen the atoms of nitrogen, carbon and hydrogen under such low temperatures as below −100° must combine chiefly into these carbon-hydrogen (CH_4) and nitrogen-hydrogen (NH_3) compounds as their most stable molecules. The increase in intensity from Jupiter to Neptune is then connected with decreasing temperature. The structure of the two largest planets, Jupiter and Saturn, was thus described by Jeffreys in 1934: a core of rock minerals is enclosed by an envelope of ice and solid carbon dioxide of a density of about one; they are surrounded by an extensive atmosphere of nitrogen, hydrogen, helium and methane, in which float clouds of crystallized ammonia.

Thus the problem of the plurality of worlds has acquired a new aspect in the last century. Three centuries ago the similarity between the dark celestial bodies and the earth had to be strongly emphasized to give a firm emotional basis to the new heliocentric concept of the world. The conviction of a purpose in creation and the idea that the planets also are destined to be the abode of living beings contributed to its gradual acceptance as a harmonious world view. Nineteenth-century astronomy, however, has demonstrated their dissimilarity with convincing power. It has revealed a far greater wealth and variety in nature than the awakening insight of simpler times could suspect. The ingenuous faith in a purpose had to give way to the scientific view that the conditions of other worlds are determined by natural influences and forces, different for each planet according to size and position. The problem of the reactions and the equilibrium of different atoms and molecules in Jupiter's atmosphere, resulting in various forms and colours, loses nothing of its charm by the exclusion of any idea of living organisms.

Yet this means a fundamental change in our concept of the world. The dream of a plurality of inhabited worlds, the dream of other men living on neighbouring kindred globes, is over. It is true that as yet we know nothing of possibilities in inaccessibly distant stellar systems; but as far as our own solar system is concerned, no other mankind exists than on earth. We do not feel the solitude, divided as we still are into hostile peoples who look upon one another as foreigners and try to exterminate one another. When, however, mankind on earth has become united, the consciousness of our being the sole humanity in our solar system—separated from other systems by impassable distances—will exert a determining influence upon our life-concept.

That the earth occupies this exceptional place among the planets is due to the presence of free oxygen in its atmosphere. Oxygen as a main constituent of a planetary atmosphere is an anomaly; since this element is easily bound by many other elements, it will be rapidly absorbed and fixed in rock minerals and other compounds. It would also have disappeared on earth had it not been continually supplied by the photo-chemical processes in the plant cells, the dissociation of carbon dioxide by chlorophyll under absorption of sunlight. So the kernel of the problem confronting us here is: how did chlorophyll originate on earth? What special conditions—elsewhere absent, or perhaps present in the future for the atmosphere of Venus with its content of carbon dioxide—brought it about first that protein molecules formed in the early tepid oceans and merged into living matter? What conditions caused the forming of this special compound that could utilize absorbed solar energy for the splitting of carbon dioxide molecules into carbon for themselves and oxygen for the future animal world, and so could open the road leading to man? The study of this fundamental problem as yet has hardly been started. The development of astronomy in the last century has brought it to the fore as a problem specific for the earth.

CHAPTER 36

COSMOGONY AND EVOLUTION

WHAT was the origin of the world? In the earliest dawn of civilization and even before, in the stage of barbarism, this problem already occupied the minds of men. In myths and legends, man gave various answers, according to the conditions of his life and the state of his knowledge. Such cosmogonies do not belong to astronomy but to folklore and religion, or, as in later days in Greece, they were the work of poets and philosophers. They were based on the simplest ideas of world structure, in which heaven and earth were its two equivalent halves. They usually proceeded from an original chaos into which the gods, by their act of creation, brought order and structure. A creation out of nothing was beyond their imagination.

In the Middle Ages and the following centuries cosmogonic ideas were determined by the Christian doctrine laid down in the Book of Genesis. The scientific contests of these centuries dealt with the structure and not with the origin of the world. This changed with the eighteenth century, the century of rationalism. The Kant and Laplace theory of the origin of the solar system out of primeval nebula has been dealt with above. Here, for the first time, primitive legend gave way to scientific theory. Theory could not go beyond the science of the time. The only science, apart from astronomy, that had developed solid foundations was mechanics, the theory of forces and motions. Eighteenth-century cosmogony therefore could do no more than apply mechanics to the limited problem of the solar system. A true all-encompassing scientific cosmogony was for the first time possible through the development of physics in the nineteenth century.

Atomic theory at the beginning of this century awakened in the scientists the consciousness that matter is indestructible. It implied the notion of an eternity of the world in the past and the future, dominated by natural laws, in which there was no room for an act of creation. Thus the concept of 'cosmogony'—though the word remains in use—was superseded by the concept of 'evolution'. The science of the evolution of the universe is intimately connected with the development of astrophysics.

The origin and growth, with increasing rapidity, of astrophysics, was the most important renewal of astronomy in the nineteenth century. Gradually it outstripped the old astronomy of position and motion. Indeed it is the real astronomy, since it is the knowledge of the world bodies themselves. It deals with their physical nature, whereas the science of their motions deals with their outer behaviour under the compulsion of gravitational forces. Most important now are not the small and dark secondary bodies which can only be seen when they are near, but the large self-luminous millions of stars, which, as the sources of light and heat, fill the universe with radiant energy. Among them our sun, the basis of our life, is most important as the foremost object of scientific research.

In this development astronomy was no longer autonomous as in former centuries, when it had to seek its own way. It now has other sciences to lean on, especially physics; the name 'astrophysics' indicated that it is the application of physics to the stars. So it had to wait until, about the middle of the century, physics had developed first the theoretical principles and laws valid for the entire world of phenomena, and secondly the practical methods for studying the distant bodies. The first condition was realized by the rise of the mechanical theory of heat, including the doctrine of energy and entropy, the second by the discovery of spectrum analysis.

The rise of theoretical physics, especially the theory of heat, was intimately connected with the development of the steam engine as the technical basis of nineteenth-century industry. Practical and theoretical study of the working of the steam engine was a matter of first importance for increase in knowledge as well as increase in productivity. Here it was seen every day how the engine produced power and performed work while heat was expended. The obvious conclusion was that heat was transformed into work or working power. From the experience that steam engines had a greater output when the steam had been heated to a higher temperature, Carnot derived another explanation: the work effect is produced because heat (at the time considered as a special weightless matter) descends from high to low temperature, just as in a water-mill it is produced when water descends from a higher to a lower level. On the other hand, machines in operation showed numerous cases where work used to vanquish frictions or resistances turned into heat. Thus in the mind of the physicists the concept of energy was born as the capacity of performing work, which in later days was increasingly to dominate the realm of physics. In the law of conservation of energy, Julius Robert Mayer in 1842 stated that energy is indestructible; it appears in many different forms—heat being one of them—which in physical processes pass into one another, the total amount always remaining the same. At the same time James Prescott Joule at Manchester determined by various experiments how many units of work were equivalent to one unit of heat: 1 calorie, the quantity of heat increasing the temperature of 1 kg. of water by 1° C, is equal to 425 kilogram-metres, raising 425 kg. by 1 metre. In Germany, Helmholtz in 1847 demonstrated the transformation of energy (called *Kraft*, 'force') in all different physical phenomena.

The truth in Carnot's apparently inconsistent view came to light when subsequently (1850–54) Rudolph Clausius in Zurich and William Thomson in England brought it into line with the conservation of energy. Energy in the form of heat can pass into the form of work only when at the same time a quantity of heat is allowed to descend from higher to lower temperature. Of itself heat goes from warmer to colder bodies, through conduction or radiation; of itself mechanical energy passes into heat. This is the automatic positive course in nature. The reverse negative processes can take place only when compensated by a simultaneous positive process. This law, called the 'Second Law of Thermodynamics', was brought into mathematical form by Clausius through the introduction of the concept of *entropy*, which name was given to the 'transformation value' of the heat contained in a system; this 'value' of heat is larger as its temperature is lower. The Second Law states that the entropy (of a closed system or of the world) can only increase. Through the automatic processes the entropy increases; all opposite changes (from lower to high temperature or from heat to work) must be compensated for by simultaneous positive processes. These two laws of thermodynamics—that the total energy in the world remains constant and that the transformations of energy take place in one special direction—became highly important for astronomy.

Immediately after the first law of energy had been established for the phenomena on earth, it was extended to the heavens and applied to the problem of the sun. This was only reasonable. All the transformations of energy on earth from one to another of its different forms—energy of motion, potential energy of gravity, heat, chemical and electrical energy, life-energy—which form the contents of all phenomena, have their source in the energy poured out upon the earth by solar radiation—the tidal motions excepted. Every discussion of the conservation of energy has to include the sun and to pose the problem of to what other forms we have to look to for the origin of solar energy. Practically, it presented itself in the astonishing question of how the sun can maintain its radiation in undiminished power, notwithstanding the continuity of its enormous losses through radiation.

Mayer had an answer directly to hand: heat proceeds from mechanical energy. The earth is continuously battered by meteors, and so to a

greater intensity is the sun. The energy of motion of the meteors when they are arrested is transformed into heat. Qualitatively this seemed to be a satisfactory explanation; when, however, a numerical test was applied later on, it failed. The yearly solar radiation amounts to 2.9×10^{33} calories; the energy of 1 gm. of matter coming from afar and falling into the sun is equal to 4.4×10^{7} calories; hence, to make up for the radiation, a mass of 6.5×10^{25} gm. yearly has to fall upon the sun, one thirty-millionth of the sun's mass (19.8×10^{32} gm.). It seems little, but it is much, far too much. By such a regular increase in the sun's mass, the planets would gradually revolve more rapidly and in ever smaller orbits. As a year is thirty millions of seconds, the earth's time of revolution would be shortened by something of the order of 2 seconds per year. This is impossible; a thousandth of that amount would have been detected centuries ago.

Helmholtz in 1853 explained the constancy of the solar radiation in a different way, by framing his theory of contraction. Contraction produces heat. When the sun contracts so that all its particles come nearer to the centre, their potential energy of attraction decreases and an equivalent amount of heat is generated. The yearly radiation is covered by a contraction of 75 metres, i.e. one eighteen millionth ($\frac{1}{18000000}$) of the diameter. It is so little that even after a thousand years it would be imperceptible to our instruments. So the contraction theory offered a satisfactory explanation of the problem of the solar heat. Moreover, it could be connected with the Kant and Laplace theory of the origin of the solar system through the contraction of an original nebula. In the eighteenth century this theory had to serve to explain the planets and their circular orbits. Now attention was directed to the contraction of the matter of the nebula—except the insignificant fragments forming the planets—into the small body of the sun. This contraction from a widely extended nebula into the sun's present volume must have produced an amount of heat sufficient to keep up the present radiation for 18,000,000 years. Thus, for the first time, physical astronomy afforded an estimate of the time scale in which the age of the solar system and of the earth must be expressed.

Proceeding along the same line of thought, the American mathematician James Homer Lane pointed out in 1871 that the sun had indeed acquired its high temperature through this contraction, so that an original high temperature of the nebula was not necessary. He demonstrated that, free in space, a sphere of gas able to expand or to contract can remain in equilibrium when its temperature changes in inverse ratio to its radius. If it contracts in losing energy by radiation only part of the generated heat is needed for the radiation, and the remainder increases the temperature. It sounds like a paradox to say that by losing energy a body should become warmer instead of colder. The same paradox is demonstrated in celestial mechanics, that a body, subjected to resistance in its orbit, increases its velocity in contracting its orbit; here the decrease in potential energy of attraction is only partly needed to neutralize the resistance, and the remainder is used to increase the velocity. Lane's law now offered a theory of how the sun had acquired its heat by regular evolution: a cool, widely extended mass of thin gas develops by its radiation into a more and more contracting solar sphere of increasing temperature. Of course this development does not continue indefinitely; Lane's law is based on the validity of the simple gas laws. When the density has increased so much that they no longer hold, Lane's inverse proportionality of temperature and radius loses its validity. Gradually the contraction becomes more difficult and smaller, the energy of contraction becomes less and is not sufficient to increase the temperature or even to compensate for the radiation. Then the sun will begin to cool down and finally turn into a small, dark and cold body. Lane's results afforded a first theory of continuous evolution of the stars, which was to dominate the ideas of the astrophysicists for many years.

It was not by chance that the idea of evolution here appeared in a form entirely different from the ideas of the eighteenth century. We saw that the eighteenth century knew of development only as preparation for a remaining final condition. In the human world this was the growth from barbarism and ignorance towards a society based on nature and reason; we have already pointed out that the nebular theory of Kant and Laplace appeared as the cosmological image of this world concept. The nineteenth century brought the idea of a continually progressive social development. The Industrial Revolution, starting in England at the beginning of that century, gradually spread over the adjacent countries, France and Germany, over the United States, and ever wider over the earth. All trades were involved, all the life-conditions of man were fundamentally changed, and in half a century the aspect of our world was transformed more profoundly and more rapidly than in all preceding centuries together. Thus the human mind became accustomed to seeing the world as a continual process of development in which no end and no goal could be discerned.

This new view expressed itself in various scientific and philosophic theories. Hegel's philosophy had already presented the world as a 'dialectic' process of unfolding of the Absolute Idea. In biology the doctrine of the development from lower to more highly organized forms, after Charles Darwin's publication in 1859 of his *Origin of Species*, found wide recognition. The idea of progressive development found its physical expression in Clausius's Second Law of Thermodynamics: all autonomous processes in nature go in one direction; the entropy of the world can only increase, never decrease. Physicists sometimes expressed this law by saying that the entropy of the world tends to a maximum. This maximum would be reached when all mechanical energy had been transformed into heat, and the temperature would be the same everywhere. The final world would consist of a dispersed nebula with nothing but molecular motions. In this alleged final condition no account is taken of the powerful world force of gravitation; on the contrary, we saw that according to Lane's results a dispersed isothermal nebula must condense into a small, dense and hot sphere: intensification instead of smoothing of the temperature differences.

The idea of evolution now took increasing hold on astronomy; everywhere in the universe it discovered processes of development. George H. Darwin, the son of the great biologist, in a series of theoretical studies after 1879, dealt with tidal friction as an important agent of evolution. It was generally accepted that the earth and the moon had originally been soft and deformable through their tide-raising mutual attractions. If there were no friction and the earth could immediately take the equilibrium figure, the high-tide humps would daily go around the earth, with the moon always exactly in zenith (or nadir). By inner friction this course was retarded; the humps followed the moon at some distance, dragged after it, or, expressed in another way, the rotating earth dragged along the humps from beneath the moon over a certain distance. Then the attraction between the high-tide humps and the moon had a retarding influence upon the earth's rotation and, simultaneously transferring rotation momentum to the moon, widened its orbit and increased its period of revolution. Darwin traced this process back into the past and found that originally the daily rotation and the monthly revolution had been equal, between 3 and 5 hours; in this condition the moon and earth continually faced and almost touched each other. When originally the earth then shrank by cooling and accelerated so that the moon fell behind, tidal friction, considerable on account of the close proximity, began to operate in driving the moon away and retarding the rotation. Day and month both increased, the latter most, until at last 29 rotations came in one revolution. Later on, the moon's increase slowed up, so that now this number is 27⅓ only. This will lead to a final state in which day and month will be equal, but now amounting to 55 of our present days; at a great distance, the two bodies will once more always show the same side to each other. What is treated here is a simplified theoretical case, because in practice the tidal action of the sun will disturb this course of things. The question of how, assuming this theory, a moon can revolve more rapidly than the planet rotates (as in the case of the inner Mars satellite) was answered by

Moulton: while the solar tides retard the rotation of the planet, the satellites of Mars are so small that they practically raise no tides and are therefore not subjected to tidal influences either.

More important than this distant future was the initial condition to which Darwin was led by theoretical reckoning into the past. It indicated that originally the moon and the earth had probably been one body that had split into two unequal parts. Here the conclusions coincided with the results of an investigation by the great French theorist, Henri Poincaré, in 1885 on the equilibrium and stability of rotating fluid bodies. In the eighteenth century, in 1740, Maclaurin had deduced how the flattening of a rotating fluid sphere increases with the velocity of rotation. In 1834 Jacobi, at Königsberg, demonstrated that a flattened ellipsoid of rotation, such as we have with Jupiter and with the earth, is not the only equilibrium figure of a rotating fluid; an ellipsoid with three considerably different axes is another. Poincaré now investigated the conditions of stability of these different forms. When the rotation (e.g. by shrinking) gradually becomes more rapid, we have first an ellipsoid of rotation with increasing difference between the polar and the equatorial diameters. With more rapid rotation, this form loses its stability and two unequal equatorial axes appear. With still more rapid rotation, this three-axial ellipsoid also becomes unstable; then it is constricted asymmetrically into the shape of a pear, indicating that it is going to split into two unequal parts. If the period of the solar tides in the earth, corresponding to the rotation time, is equal to the period of the body's proper vibration time, which is estimated at between 3 and 5 hours, then the deformations will increase unhampered and a catastrophic development may set in. In this way Darwin explained how the original body of the earth may have split into two unequal bodies; the resulting system is the starting point for his discussion of evolution through tidal friction.

These investigations provided the first response based on exact scientific analysis to the old mysterious and impressive cosmogonic problem of the origin of the present world. Afterwards Jeans, in a series of researches on cosmogony published in 1922, confirmed these results and extended them over more general cases of gaseous masses of different distributions of density. An interesting result was that, with a rapidly rotating mass of thin gas, the surplus matter flows out at two opposite points of the equator, forming two spiral arms. Could it give a clue towards explaining the spiral nebulae?

Most noteworthy, however, is the fact that all this nineteenth-century research on cosmogony is not applicable to the problem that is of foremost interest to mankind; it is silent regarding the origin of the planetary system. Repeatedly critical remarks pointed out that a system like ours could not have originated in the way expounded by Kant and Laplace. The chief difficulty is that 60 per cent of the total moment of momentum (what in a popular way might be called the 'quantity of rotation') of the solar system is supplied by the orbital motion of Jupiter, which possesses no more than one-thousandth of the mass of the system, and only 2 per cent by the rotation of the sun. It is not easily conceivable that in the original division the rotational momentum could thus be partitioned between the outer ring and the condensing central mass. Nor was it possible that such an amount of rotational momentum could be transferred by tidal friction from the sun to the circulating small body. The astronomer Moulton and the geologist Chamberlin, at Chicago, who put these arguments forward strongly, concluded that the bulk of the rotational momentum must have been brought into the system from outside. So in 1906 they formulated the theory that the planetary system had originated from the passage of another star close to the sun. By its attraction eruptions of solar matter had taken place, which, because of the sideways attraction of the passing star, acquired orbital motions constituting the final rotation momentum. In the surroundings of the sun it condensed into small bodies, called 'planetesimals', which afterwards, by their collisions, coalesced into planets and thereby produced the high temperature in the solid interior of the earth.

This encounter theory of the origin of the planetary system, with many variations as to details, met with the approval of Jeans, Eddington and many other astronomers. Jeans pointed out, and Eddington agreed, that such close encounters of stars, because of their large mutual distances, must be so extremely rare that perhaps no other case has happened among the millions of stars of our galactic system. 'The calculation shows that even after a star had lived its life of millions of millions of years, the chance is still about a hundred thousand to one against its being surrounded by planets' (Jeans).[203] Then it follows that possibly our planetary system is unique in the stellar universe and that hence the earth, as an abode of living beings, is unique in the world; 'not one of the profusion of stars in their myriad clusters looks down on scenes comparable to those which are passing beneath the rays of the sun' (Eddington).[204]

This startling conclusion surely cannot be considered as a final dictum of science; if one supposed cause is found to have so infinitely small a probability that it could happen once only in millions of millions of cases, numbers of other apparently improbable possibilities remain. That it was proclaimed notwithstanding its hypothetical character as a probable conclusion must be considered not only as a result of scientific reasoning but also, and even more, as the expression of a new concept of the world. From the very first, the special and unique place of man in the world was closely knit with religion. When Giordano Bruno proclaimed his doctrine of a plurality of inhabited worlds, he put himself in opposition to the accepted church doctrine. In the middle of the nineteenth century the doctrine of a multitude of planetary systems, all inhabited by intelligent beings, formed part of a world concept often expressed in materialist and rationalist forms, strongly antagonistic to the dominant religious creeds. Afterwards, in Europe, the materialist and rationalist ideas lost much of their prestige against the increasing influence of religion, doubtless in connection with the social crises, catastrophes and world wars, engendering uncertainty of life and the future. It is not surprising that in such times the rejection of former materialist doctrines also involved a critical mood towards belief in a plurality of inhabited worlds.

CHAPTER 37

THE SUN

WHAT is the sun? The old belief that the sun was a globe of pure fire and light was shaken when the telescopes revealed the existence of dark spots in 'the eye of the world'. By analogy with molten metal, they were compared with scoriae. Gradually their pitch-black appearance engendered the idea of a black interior of the sun surrounded by an ocean of fire. Lalande took the solar spots to be mountains protruding above this ocean. In 1774, however, Alexander Wilson at Glasgow observed that regular round spots, on approaching the sun's border, show the surrounding half-shade ring to be broad at the outer side and narrowing at the inner side, like the inclined edge of a hole. He concluded that solar spots are holes in the brilliant shell through which we see the dark interior.

William Herschel adopted this idea and in a paper dated 1795 extended it to the supposition that the dark solar surface could quite well be inhabited by living beings. They had to be protected against the heat from above by an interposing layer of clouds; classicists may be reminded here of Aristotle's world structure, with the fire above the air. Herschel produced good arguments against the objection that the heat from above would scorch the dark surface beneath. He said: 'On the tops of mountains of sufficient height . . . we always find regions of ice and snow. Now if the solar rays themselves conveyed all the heat we find on this globe, it ought to be hottest where their course is least interrupted. Again, our aeronauts all confirm the coldness of the upper regions of the atmosphere . . .' To explain this phenomenon, he assumed that the sun's rays produced heat only by contact with 'the matter of fire' contained in the substances that are heated; then we had further to assume that the lower atmosphere and the dark surface of the sun are not 'capable of any excessive affection by its own rays'. Thus he could find similarity with the other globes of the planetary system, which 'leads us on to suppose that it is most probably also inhabited, like the rest of the planets, by beings whose organs are adapted to the peculiar circumstances of that vast globe'.[205] In the first half of the nineteenth century this notion of a dark solar body was generally accepted; it shows how far physical consciousness lagged behind the extent of astronomical knowledge. A more profound insight could not force its way until the concept of energy was established. Then, in the second half of the century, the study of the sun advanced with rapid strides.

At first this study proceeded along the old paths, though with more breadth and perseverance. It was restricted chiefly to observation of the sunspots. In 1826 Schwabe, a chemist at Dessau in Germany, began to notice and register the sunspots regularly, chiefly with a view to discovering an eventual new small planet within the orbit of Mercury. What happened to him he compared to the experience of Saul, who went out to search for his father's asses and found a kingdom. After many years, comparing all his notes in 1843, he found (and published in 1851) that the number of sunspots showed a periodic variation. In 1828 and 1829 and also in 1836–39, the sun was not without spots for a single day, whereas in 1833 and 1843 on half the days of observation no spot was present. The total number counted in 1828 was 225; in 1833 only 33; in 1837, 333; in 1843, 34. So the maxima and minima returned after nearly 10 years. Rudolf Wolf at Bern, later at Zurich, by investigating all the historical data about sunspots, could trace the periodicity through the former centuries; the mean value of the period was $11\frac{1}{9}$ years, but with large irregularities between 7 and 17 years. Numerous attempts have been made to find the cause of this curious phenomenon, mostly by connecting it with the course of the planets, chiefly Jupiter, but without result. The forces producing sunspots must be located in the sun itself.

Still more remarkable was a discovery by Lamont, a Scottish scientist at Munich, published in the same year (1851), that the irregular perturbations of the magnetic instruments and the earth's magnetic field, also in a 10 year period, were alternately stronger and smaller; the aurorae connected with them showed the same periodicity. Sabine in England and Wolf in Switzerland immediately pointed out that both periods coalesced; the magnetic perturbations and the aurorae followed the sunspots not only in their periodicity itself but also in all its irregular variations. Even the appearance of single large spots produced magnetic storms and aurorae on earth. Thus a most remarkable and mysterious effect of solar disturbances upon terrestrial phenomena came to light.

More important than the simple counting of sunspots was the determining of their position and motion. The first object was, as with the planets, to find the rotation period of the sun. One of the foremost workers in this field was Carrington, at Redhill, who determined the positions of sunspots in the years 1853–61. He belongs to the numerous groups of British amateurs who, by the quality of their work, count as fully-fledged astronomers in the sense that William Herschel, South, Lassell and Lord Rosse were also amateurs. England from olden times had produced a class of wealthy landowners and merchants, later on also industrialists and businessmen, who expected nothing from government but had to do everything on their own initiative. They founded libraries and academic chairs and, if attracted by astronomy, built their own observatories and did scientific researches for themselves. Carrington had installed a meridian circle with which at night he made observations for a catalogue of northern stars to complete Bessel's zone work, and in the daytime he observed the sunspots. He found that the period of rotation increased with the distance from the solar equator. Near the equator the period was 25.0 days; at 20° it was $25^{d}\ 18^{h}$; at 30° it was $26^{d}\ 11^{h}$ and it increased to 27½ days at 45°, where the spots are scarce. Hence the spots cannot be fixed parts of a solid solar body. His results were confirmed by similar work (in 1860–73) by Spoerer, a German amateur at Anklam in Pomerania. Both observers perceived another peculiarity. During the years of most numerous spots these came gradually nearer to the sun's equator; their latitude decreased from about 25° to about 10°, and then they became extinct about latitude 5°. At the same time (the time of minimum number), at a high latitude of 25°–30°, the first numbers of a new cycle appear, which in the next years of increasing number expand towards lower latitudes. Thus the periodicity of the sunspots consists in a succession of consecutive series or waves, all starting at high latitude, swelling as they descend towards low latitude and being extinguished there.

Photography, of course, was used to make pictures of the sun immediately after its discovery. Here, contrary to other celestial objects, it was the abundance of light that had to be neutralized by special devices. Now it was possible in a split second of exposure to register all the details of the sun's surface—the spots, faculae and other markings—so that afterwards their number, their total surface, their shapes and movements, could be studied. A spectacular effect was obtained by combining two pictures in a stereoscope, one taken a short time after the other, when the sun had rotated to a small extent; in this way the sun was seen as a globe, with the dark spots as pits and the faculae floating at high level. In 1858 Warren de la Rue devised and erected a photoheliograph at Kew, afterwards transferred to Greenwich and used for the regular photographing and measuring of solar pictures as routine work.

Besides this work, which was used chiefly for statistical purposes, there were refined technical methods for the study of the minute details in the structure of the spots as well as of the fine granulation of the undisturbed photospheric surface. Here it was the French astronomer P. Jules C. Janssen, at Meudon, who in the seventies excelled in his enlarged photographs of the granulation and the sunspots. With the rapid variations in the spots, the finest detail was not so important here as with the planets, except for the granulae. Through a careful comparison of a number of photographs taken in rapid sequence, Hansky at Pulkovo in 1905 was able to determine the average lifetime of the separate granulae at 2–5 minutes; then they dissolved and were replaced by others.

Up to the middle of the century only the sun itself was observed by the astronomers. In 1842 for the first time, at a total eclipse of the sun visible in Southern France and Northern Italy, their attention was directed to those luminous phenomena around the dark disc that make total eclipses of the sun the most wonderful sight and the most important source of knowledge: the far-extending glory called the 'corona' and the small pink 'prominences' protruding at different points outside the dark limb of the moon. They had already been observed earlier, in 1733, by Wassenius at Gothenburg in Sweden, and described as red clouds in the moon's atmosphere. They were mentioned even in the medieval Russian monastery chronicles, where, e.g. on May 1, 1185, is written: 'The sun became like a crescent of the moon, from the horns of which a glow similar to that of red-hot charcoals was emanating. It was terrifying to men to see this sign of the Lord.'[206]

The eclipse of 1851, visible in Sweden, made it possible to ascertain that the prominences belonged to the sun, not to the moon, and that they are the highest parts of a narrow pink ring (afterwards called the 'chromosphere') surrounding the sun, which also had occasionally been mentioned before. At the eclipse of 1860, observed in Spain, photography was used on a big scale to ascertain the objective reality of all these phenomena. From now on, at every total eclipse astronomers travelled to the zone of totality, in whatever distant country it might be, to make, in the few minutes available, the various observations that by startling discoveries extended our knowledge of the sun.

Then in the years 1859–62 spectrum analysis arose, chiefly through the work of Kirchhoff and Bunsen. It was, so to speak, in the air. Several scientists, such as Stokes, Foucault, Ångström and others, had perceived that Fraunhofer's double line D in the solar spectrum coincided with a bright-yellow double sodium line. The conclusion that sodium must be present in the sun was obvious. Stokes described the process in this way: the particles that had absorbed light of this special wavelength from a light source were afterwards able to emit it. It was Gustav Kirchhoff (1824–87), eminent theoretical physicist, who gave a solid scientific basis to spectrum analysis. He demonstrated that for any wavelength the ratio of emission (quantity of emitted light) and absorption (the fraction absorbed from incident light) is the same for all bodies and is equal to the emission of a 'perfectly black body' (supposed to absorb 100 per cent). The latter is a continuous function of wavelength and temperature. So a hot gas that absorbs one special wavelength and thereby produces a strong black line in the spectrum of a light source will emit the same wavelength strongly, whereas adjacent wavelengths, which are little or not at all absorbed, are emitted little or not at all. The significance of the Fraunhofer lines in the solar spectrum was now at once clear; they indicated what absorbing particles were present in the sun. Kirchhoff measured, on an arbitrary scale, the position of some thousands of Fraunhofer lines and established their coincidence with lines emitted by diverse chemical elements, such as hydrogen, iron, sodium, magnesium, calcium, etc. He concluded that these elements were present in the sun's atmosphere, absorbing their special wavelengths from the continuous spectrum emitted by the solar body. Ångström in 1868 replaced Kirchhoff's arbitrary scale by the natural scale of wavelengths; he expressed them in a unit of one ten-millionth of a millimetre, afterwards called by his name; for the visible colours the wavelengths in this scale are numbers of four figures (red, 6,500; green, 5,000; violet, 4,000).

Spectrum analysis miraculously realized what shortly before had been declared forever impossible: establishing the chemical composition of distant inaccessible bodies. The French philosopher of positivism, Auguste Comte, in 1835 in his *Cours de philosophie positive*, in order to emphasize that true science is impossible if it is not based on experience, wrote of the celestial bodies: 'We understand the possibility of determining their shapes, their distances, their sizes and motions, whereas never, by any means, will we be able to study their chemical composition, their mineralogic structure, and not at all the nature of organic beings living on their surface.' And some pages later: 'I persist in the opinion that every notion of the true mean temperature of the stars will necessarily always be concealed from us.'[207] It is apparent here, just as it was with Descartes and with Hegel, that philosophy must stumble when it tries to prescribe and predict the results, or even the methods, of science; its task is to use them as materials for its own theory of knowledge, epistemology.

The new discoveries brought new ideas on the nature of the sun. It was no longer possible to believe that the sun's interior was a dark and cold body. Kirchhoff considered the sun to be a red-hot sphere, solid or liquid—because its spectrum is continuous—surrounded by a less hot atmosphere containing the terrestrial elements in a gaseous state, which produce the Fraunhofer lines. He considered the sunspots to be cooler clouds in this atmosphere; he was well aware that relative darkness was bound to a lower temperature and that the often-heard opinion that darkness was due to a smaller power of emission was refuted by physical law. His ideas on the sun's interior were corrected in 1864 by Secchi and John Herschel, who assumed it to be gaseous also; they ascribed the continuous spectrum to droplets floating like a cloud layer deep in the atmosphere. Experiments showed that strongly compressed gases also emitted a continuous spectrum; moreover, the physicist Andrews in 1869 discovered that any matter above a certain 'critical temperature' cannot exist in a fluid but only in a gaseous state. Yet the cloud theory was generally adhered to. 'It seems almost impossible to doubt that the photosphere is a shell of clouds,' was written by Young in 1882 in his book *The Sun*.[208] Father Secchi, as well as the French astronomer Faye, explained the sunspots as openings in the cloud layer which at such places was volatilized by an uprush of hotter gases expelled by pressure from below. More complicated accessory explanations were needed to meet the objection that in that case the spot would appear hotter than the surroundings. Faye was especially struck by whirling phenomena observed in the spots, and he compared them with tornadoes on earth.

To establish what elements are present on a luminous celestial body is, in principle, very simple; it demands nothing but the observation of exact coincidences of spectral lines or the equality of exactly measured wavelengths. Since instrument making and glass techniques were highly developed, excellent prism spectroscopes for various purposes could soon be constructed. The introduction of photography not only afforded the same advantages as in other domains of astronomy but, moreover, gave access to a new field of wavelengths invisible to the eye—the ultraviolet part of the spectrum between 4,000 and 3,000 Å (Ångströms). Thus spectrographs increasingly superseded spectroscopes.

A prominent achievement in this development was the construction of concave gratings, in 1887, by Henry A. Rowland at Baltimore. A concave metal mirror produces a sharp image of a slit without the need of lenses. On such a mirror a grating of fine parallel lines, 25,000 per inch, was engraved, which produced a series of diffraction spectra of great dispersion and high resolution. Rowland worked for many years on the perfecting of an engraving machine capable of automatically cutting the fine grooves at exactly equal distances. The reward of this painstaking work consisted in spectra which remained unequalled for dozens of years. With these gratings Rowland photographed the solar spectrum and published it in 1888 as an atlas of maps, on a constant scale of 1 Å = 3 mm. so that the entire spectrum between 3,000 and 6,900 Å forms a band of 40 feet in length. It contains more than 20,000 Fraunhofer lines in all intensities, from barely visible traces up to heavy dark bands. By measuring the original photographs Rowland was able to publish in 1896 a catalogue of all these lines, with their wavelengths given to three decimals (hence in seven figures) and their estimated numerical intensities, a standard work used henceforth by every astrophysicist. From the lines he could ascertain the presence of 36 terrestrial elements in the sun. In a later revision of his catalogue by St John at Mount Wilson in 1928, this number was increased to 51. The most important later progress was the publication, in 1940, by Minnaert and his co-workers at Utrecht, of a photometric atlas of the solar spectrum. Instead of by separate lines, the entire spectrum is pictured here by a continuous intensity curve, so that, in addition to the place, the curve also shows the width, the character and the distribution of intensity (the profile) of every line. In the figures in plate 13 the progress of our knowledge of the solar spectrum can be seen.

We have now to revert to the nineteenth century for the first application of the new method of spectrum analysis to the special solar phenomena. The occasion of the first total eclipse after the introduction of spectral research was of course seized upon to satisfy the curiosity as to the nature of the newly discovered prominences and corona. Numerous observers with their spectral apparatus went to India for the eclipse of August 18, 1868. In their results they all agreed, with small differences in detail only, that the spectrum of the prominences consisted in some few bright lines and that, hence, the prominences were glowing masses of gas. The brightest lines were the red and green emission lines of hydrogen (designated $H\alpha$ and $H\beta$) coinciding with Fraunhofer's C and F, and, moreover, a yellow line, first taken for the sodium line D, afterwards seen to be different and denoted D_3. It was not known from any terrestrial spectrum and was hence ascribed to an element present only on the sun and called 'helium'.

The observed emission lines were so brilliant that one of the observers, Janssen from Meudon, immediately understood that the darkness of an eclipse was not needed to make them visible. The next day he placed the slit of his spectroscope just outside the sun's limb and without difficulty could observe the emission lines in full daylight. For some weeks he could study the sun 'during a period equivalent to an eclipse of 17 days', as he wrote in a report to the Paris Academy. 'I have made charts of the prominences which show with what rapidity (often in a few minutes) these immense gaseous masses change their form and position.'[209] At the same time, Lockyer was working in England on the same lines; two years before, supposing that the prominences consisted of glowing gas emitting bright spectral lines, he had already contrived that they could be made visible if the continuous spectrum of the sky obliterating them could be weakened by strong dispersion. Had not the instrument-maker left him waiting a long time for the spectroscope he had ordered, he would have made his discovery long before the eclipse. Now the discoveries of both astronomers were presented to the Paris Academy on the same day in 1868.

The situation had now changed considerably. To study the spectrum of the prominences, it was not necessary to wait for an eclipse; it could be done on any bright day. Moreover, having plenty of time now, the observer could derive the extent and shape of a prominence by moving the slit. This was done still better, first by Huggins, by opening the slit widely; if the continuous background was sufficiently weakened by strong dispersion, the entire prominence could be looked at and followed during all its rapid changes. Several observatories soon put the regular observation of prominences on their programmes of work. They were found to be most frequent in the years of sunspot maxima and at the same latitudes as the spots, though they occurred less frequently in high latitudes, up to the poles, where the spots are absent. As to form and character, two main types could be distinguished: the quiet prominences, floating like pink clouds in the atmosphere, and eruptive prominences, shooting up like fountains of fire to great heights and then dissolving or being sucked into the pits of the sunspots. Sometimes quiet clouds were suddenly torn into a mass of fragments as if dispersed by a violent storm. The same velocities of hundreds of miles per second were also indicated by distortions, i.e. local displacements of the lines in slit spectra. All these fascinating scenes of events on a scale of thousands of miles presented new problems; they helped to sustain the idea of eruptions as the cause of sunspots.

At the eclipse of 1870 in Spain, Young discovered in the faint continuous spectrum of the corona a narrow green emission line of a wavelength first given as 5,315 but afterwards found to be 5,303; as it was found not to occur in any known spectrum, a second unknown solar element was assumed, called 'coronium'. Another important discovery was made by Young; having set the slit of his spectroscope nearly tangent to the sun's limb, at the moment of the eclipse he saw, as in a flash, the flaring-up of an innumerable number of emission lines; after one or two seconds they disappeared, when the moon covered the thin layer of only $1''$ width (representing 500 km.) that emitted them. It seemed as if all the Fraunhofer lines were for a moment reversed into bright lines, because the gas layer absorbing them was seen sideways without the background of the sun; so it was called the 'reversing layer'. In order to observe it without an eclipse, Young sought the steady air on the top of Mount Sherman; here he saw that the spectrum consisted of the same numerous metal lines as the solar spectrum, but mostly with different relative intensities. The brightest of these lines also appeared in the eruptive prominences. In the quiet prominences, in addition to the hydrogen lines and the helium line D_3, a small number of unknown lines were seen; when Ramsay in 1895 had discovered helium in terrestrial sources, so that its entire spectrum could be investigated, they appeared to be other helium lines.

When photography was applied to these phenomena, Huggins in 1875 discovered the harmonic series of ultraviolet hydrogen lines, in continuation of the four lines $H\alpha$ to $H\delta$. At continuously decreasing separations these lines merge at last at a limiting wavelength, below 3,700 Å. On the photographs it appeared, moreover, that the two violet calcium lines coinciding with H and K of Fraunhofer, with wavelengths 3,968 and 3,934 Å, surpassed all other chromospheric lines in brightness; indeed, H and K themselves are the strongest absorption lines in the Fraunhofer spectrum.

For the purpose of photographic observation of solar eclipses the appropriate instrument was devised and constructed by Lockyer. Joseph Norman Lockyer (1836–1920) was also an amateur, an official at the War Ministry, who could devote only his spare time to astronomy, yet was esteemed a first-rate scientist. He constructed the 'prismatic camera' simply by placing a prism before the camera lens. A self-luminous object is then pictured in as many images as it emits separate wavelengths, each image showing by its shape the distribution of its atoms emitting this line. A photograph taken at the exact moment of the 'flash' shows all the metal lines of the reversing layer as short narrow arcs, whereas the hydrogen lines H and K present the prominences in their true form; the green corona line is pictured as a faint and broad luminous ring, sharply cut off at the inner edge by the limb of the dark moon. In 1893 the first imperfect photographs had been obtained—showing, because the exact moment of the flash was missed, only the arcs of hydrogen, helium, H and K. After 1896 prismatic cameras were in regular use at every eclipse because of the vast amount of information they afforded. Mitchell in 1905 perfected the method by using a Rowland grating instead of a prism.

The mysterious coronal lines remained an important object of study at solar eclipses, because all attempts to make them visible in full daylight had failed. The curious fact here was that different observers at different eclipses reported new and different coronal lines not perceived before. A few of the brightest lines appeared regularly: besides the green line, a red and a violet line; but for the others there remained doubt as to whether they were real and whether the coronal spectrum was variable. For all these lines the origin was unknown; perhaps there were more 'coronium' elements.

The regular photography of prominences led to a new method of research. The bright lines of hydrogen or calcium were visible not only outside the sun's border but also at different strongly disturbed places on the solar disc, mostly in the vicinity of sunspots. They appeared here as bright, narrow emission lines in the centre of the broad, dark absorption lines and were considered to be due to luminous gaseous masses in the upper layers of the atmosphere. In 1890–91 Deslandres at Paris and George E. Hale at Chicago, independently and in somewhat different ways, constructed an instrument, called a 'spectroheliograph', to picture these high emissions. The slit of a spectrograph is made to slide over the sun's image; a firmly connected second slit behind the prism, allowing only the narrow emission line to pass, slides over the photographic plate. Such pictures of the sun in the light of one wavelength were made first in Chicago and at the Yerkes Observatory; to conduct a more profound study, Hale founded the 'Solar Observatory' at Mount Wilson, where later the restricting word 'solar' was dropped. The regular study of these pictures, most in the light of the calcium K line and the hydrogen Hα or Hγ, revealed an abundance of remarkable structures, especially about the sunspots, reminiscent of spiral arms or whorls and vortices, sometimes connecting two adjacent spots by curves reminiscent of iron filings over two magnetic poles.

The spectrum of the sunspots was also an object of many researches. Though sunspots appear black by contrast, they radiate a strong light. In 1866 Lockyer found, as a cause of the relative darkness, the broadening of most Fraunhofer lines and the appearance of numerous additional fine lines. Their detailed investigation had to wait for the use of more perfect spectrographs. It was after 1920 that photographs with such great dispersion were taken at Mount Wilson that a catalogue of spot lines could be published (in 1933) by Charlotte Moore, which in their completeness approached Rowland's solar spectrum. Careful comparison showed interesting differences; the so-called 'high-temperature' lines were weaker, the low-temperature lines were stronger in the spot spectrum, whereas numerous fine lines belonging to molecular bands appeared. Both gave clear evidence that the spots are regions of low temperature.

Most metallic lines in the spot spectrum with great dispersion show a reversal in their centre, a bright line separating the dark line into two components. They were considered to be of the same character as the bright emissions in the centre of the hydrogen and H and K lines, i.e. as due to high layers of hot gases. In 1908, however, Hale discovered at Mount Wilson that their origin was entirely different, viz. the magnetic doubling of the lines through the Zeeman effect. The two components were circularly polarized in opposite directions, showing that strong magnetic fields were acting in the sunspots. It was remarkable that pairs

of spots following one another in the solar rotation, which in the Hα light often show opposite directions in the vortex structure, also showed opposite magnetic fields. More remarkable was it that the sequence of polarity in such a pair in the northern hemisphere was opposite to the sequence in the southern hemisphere. It was suggestive of the opposite direction of rotation of tornadoes on earth at both sides of the equator. When the whirling charged particles in the two spots forming a pair had opposite directions, Hale considered them as opposite ends of one vortex tube situated in the deeper parts and producing spot phenomena where it ended at the surface. The many new ideas and problems raised by these phenomena acquired a still more curious and mysterious aspect when, after the sunspot minimum of 1912, it appeared that the polarities of the northern and southern hemispheres had interchanged. After the minimum of 1922, a new interchange took place. So it might be said that the real period of the sunspots, especially in their magnetic and rotational phenomena, was not 11 but 22 years.

Gradually a wealth of knowledge on solar phenomena had been acquired by a long series of patient researches and startling discoveries. Yet all this could still only be called a prescientific period of solar physics. Half a century after the discovery of spectrum analysis, astrophysics was in the same stage as was the old astronomy before Kepler and Newton. It consisted of an abundance of data and facts but without the basis of a solid theory. New theories of the sun were now and again devised by prominent observers to explain new facts just discovered, but they were based on the imperfect general ideas of the time. There were the more general theories of keen outsiders, like August Schmidt in Germany, who in 1891 explained many phenomena, such as, for example, the sun's sharp border, by strongly curved light rays. Proceeding from the same idea, the Utrecht physicist W. H. Julius gave an explanation of the Fraunhofer and the chromospheric lines as consequences of anomalous dispersion. In all these attempts the rigid certainty of unfailing principles was lacking.

This was due not to astronomy, which could not make experiments with its objects, but to physics. Astrophysics could not become a real science until physics had developed the phenomena of radiation into a perfect theory. This was accomplished at the beginning of the twentieth century. It was possible only by such a fundamental revolution in the principles of physics that—as was then often said—to a physicist of 1890 not only were the new physical laws entirely incomprehensible but their terms even formed a foreign language which he could not understand.

This development began with the establishment of the general laws of radiation. In 1870 they were so utterly unknown that Secchi put the temperature of the sun's surface at some millions of degrees, whereas the physicist Pouillet found 2,000°. Both results were based on almost the same experimental value of the sun's amount of radiation, but in the one case the radiation was assumed to increase proportionally, in the other exponentially, with the temperature. In 1879 the Austrian physicist Stefan, from accurate measurements over a large range of temperatures, deduced that the total radiation was proportional to the fourth power of the absolute temperature. In 1884 Boltzmann gave a rigid theoretical demonstration of this law. Following along the same paths, W. Wien in 1893 deduced that the radiation of a perfectly black body was given by one single function of the wavelength, which with changing temperature shifted in such a way that the wavelength of its maximum was inversely proportional to the temperature. These two laws, applied to observational data, established that the temperature of the sun's surface was nearly 6,000°. In 1906 they found their final completion in the radiation formula of Max Planck, which expressed the black radiation as a combined function of wavelength and temperature. More than by its own contents, this formula became famous because Planck had to assume for its derivation that radiation energy is emitted and absorbed not in a continuous stream but in discrete portions. His theory of the 'quantum' of action placed its stamp upon all later physics.

At the same time the knowledge of the regularities in the monochromatic radiations of the atoms, which we observe as the emission line spectrum of each element, also made great progress. In 1885 J. J. Balmer at Basle published his famous formula for the wavelengths of the four visible hydrogen lines. They can be exactly expressed by the value 3645.6 multiplied by the ratios $\frac{9}{5}$, $\frac{16}{12}$, $\frac{25}{21}$, and $\frac{36}{32}$, hence by the fraction $n^2/(n^2-4)$, in which n is taken as 3, 4, 5 and 6. This formula was confirmed by the ultraviolet lines which Huggins and Draper had photographed in the spectra of the star Vega and of the solar chromosphere. They correspond to the larger values of n from 7 on increasing indefinitely. The wavelengths decrease with decreasing intervals, so that the lines crowd ever nearer together, until they coalesce in the limiting wavelength 3645.6. After this lucky find, analogous but more complicated formulas for other spectra were devised, first by Kayser and Runge and then in a more promising form by Rydberg. The abundant lines in these spectra were full of regularities and numerical relations; their grouping in doublets, triplets and other multiplets was connected with the position of the element in the periodic system of elements. Too little was known, however, of atomic structure to reduce these spectrum relations to atomic properties. The spectra were, according to a saying of those days, an answer of nature to which we did not know the question. In the light entering our instruments the distant celestial bodies transmit messages telling of their condition; but the messages are in code, and, as long as there was no key, they could not be deciphered.

Niels Bohr's atomic model, constructed in 1913 on the basis of the structure derived by Rutherford in 1911 from his experiments, was the key enabling the physicists to break the code and to decipher the light messages. In the next dozen years the rapid progress in theory and experiment made the entire field of spectral structure and corresponding atomic structure into an assured domain of science. Every line emission or absorption is produced by the transition between two atomic states of different energy; its wave number (the reciprocal of the wavelength) is the difference between two 'terms', corresponding to the energies of these two states. Thus hundreds of spectral lines could be reduced to some few tens of terms.

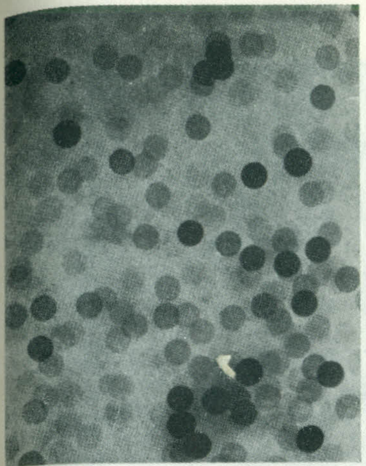
Based on this theory of atomic spectra, the Fraunhofer lines could give information on the interchange of energy of the atoms and hence on the physical conditions in the sun's atmosphere. The solar spectrum, of course, could only give information on the outer layers, the only ones from which light enters our instruments; but this information was complete. The state of these layers could now be ascertained. In 1905 Karl Schwarzschild (1873–1916), a pioneer here as in so many fields, gave a theory of the sun's atmosphere based on the principle of radiative equilibrium, according to which the temperature at any point is a result of the radiation it receives from all sides as the main mechanism of heat interchange at these high temperatures. Here the cloud theory of the photosphere has disappeared completely; temperature and density regularly increase in going downwards. Building on this foundation Milne from 1921 on and Eddington in 1923 gave a thorough theoretical treatment of such an atmosphere as the gradual thinning-out of the dense and hot deeper layers. The continuous spectrum, as well as the Fraunhofer lines, now appeared as the combined product of all the deeper and higher layers of atoms; the distinction between photosphere, reversing layer and atmosphere no longer corresponded to real qualitative but only to practical quantitative differences.

The structure of a Fraunhofer line now also became an object of investigation. What was formerly called a 'line', with one wavelength, in reality had width and structure. Theory had shown that the atomic processes existed not only in taking up and emitting radiation (as Stokes originally had assumed) but also in collisions, in which energy of thermal motion was transferred to or from the atoms (corresponding to Kirchhoff's law). Adjacent wavelengths participated in these processes, though to an amount decreasing with distance. After the example given by Unsöld's theory in 1927, the entire intensity profile of a Fraunhofer line could be derived in the following years, so that the conditions in the solar atmosphere—temperature, pressure, ionization and the abundance of the different elements—could be derived from the line intensities. A remarkable unexpected result was that the abundance of hydrogen atoms surpasses a thousand times the abundance of all the metal atoms.

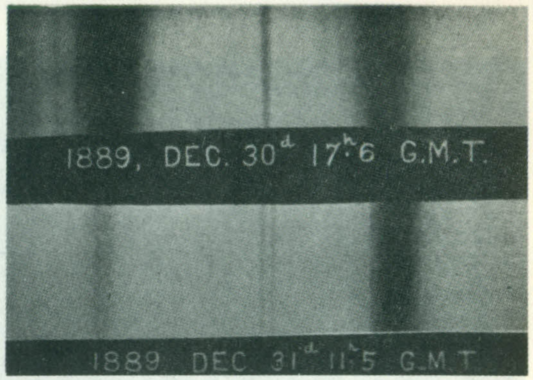
The origin of the continuous spectrum now began to present difficulties. About 1870 it was easily explained by the cloud theory, and about 1910 again by the radiation from deep gases of high density. However, since atomic theory had shown their radiations to be the result of jumps of energy giving line spectra, these explanations became doubtful. In 1939 Rupert Wildt pointed out that hydrogen atoms are able to bind and to release a second electron; in the absorbing and emitting of the needed small amount of energy, all the wavelengths of the visible solar spectrum are involved. These processes are rather infrequent; yet they play an important role because hydrogen atoms are in enormous abundance compared with other atoms. It is a curious fact that the incomparable beauty of the deep hues shading off in the continuous solar spectrum—the source of all light and colour on earth—has its origin in such an accidental and secondary process.

Astronomers could now (about 1930) be satisfied that the nature of the sun was known in its main outlines and that only details had to be added. Nature, however, in her inexhaustible richness, ever again surprised us by unexpected phenomena. We mentioned how photographing the sun with a spectroheliograph had become everyday routine at Mount Wilson. Photographing, with all its documentary value, is in some respects a new work of registering; one misses the direct contact with the happenings of every moment. Hale, himself the inventor of the spectroheliograph, was not satisfied with it, because he missed the charm of seeing what happened; so in 1926 he constructed a 'spectrohelioscope', on the same principle, adapted to visual observation. This new method of observing led, in the following years, to the discovery of sudden explosions occurring on the sun, bright flashes of hydrogen radiation, mostly in the disturbed spot regions, which lasted about ten minutes and after half an hour had disappeared. It was then remembered that Carrington in 1859, without spectroscope, had observed a short-lived brilliant starlike flare-up lasting only five minutes, which he ascribed to a large meteor falling into the sun. In 1933 it was perceived that unexpected radio fadings coincided with these flares on the sun, a new case of solar disturbances influencing the phenomena on earth. But what was their cause?

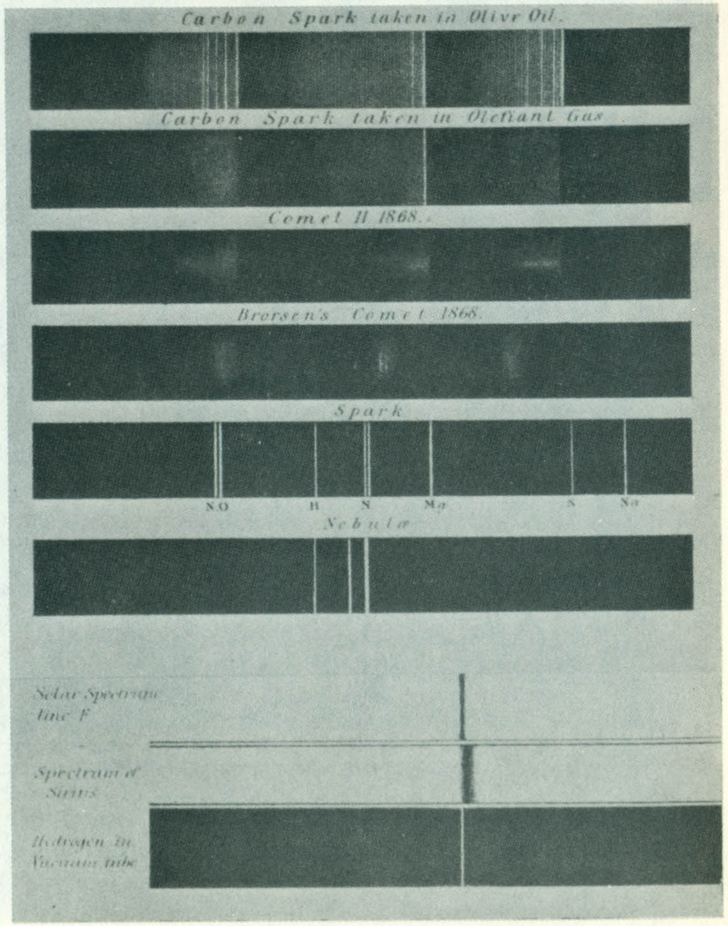
Moreover, some phenomena which did not fit into the general picture had been detected. The presence of a line of ionized helium in the



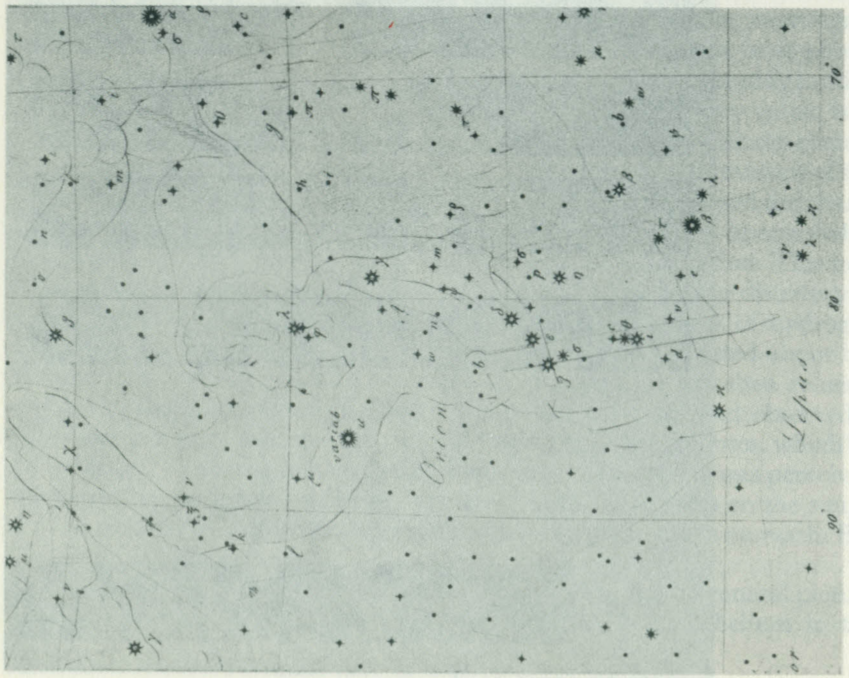
15. Part of Orion, extrafocal exposure (p. 441)



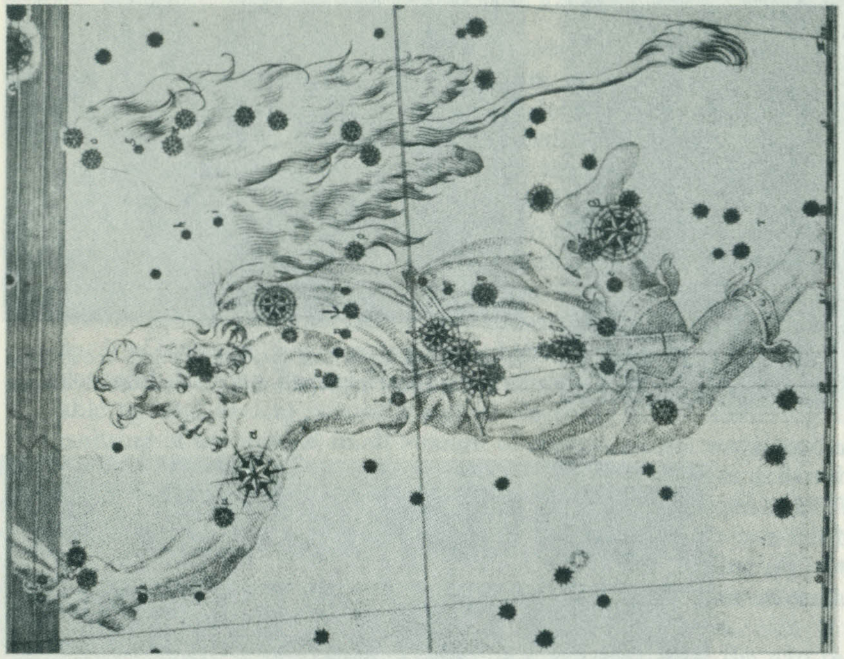
Doubling of the K line in β Auriga (the broad lines to the left and right are the 5th and 6th hydrogen lines) (p. 435)



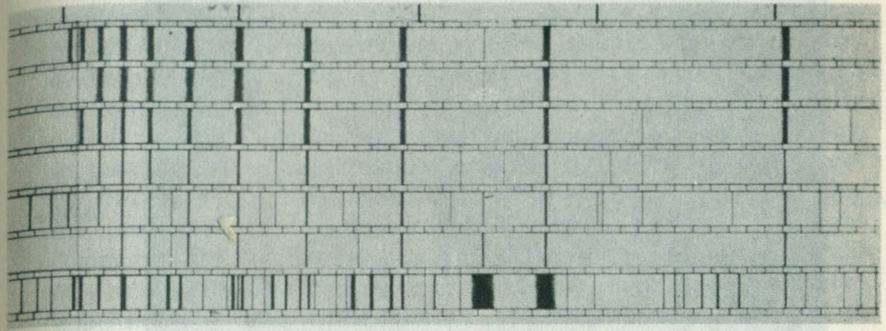
Spectra by W. Huggins (pp. 426, 451)



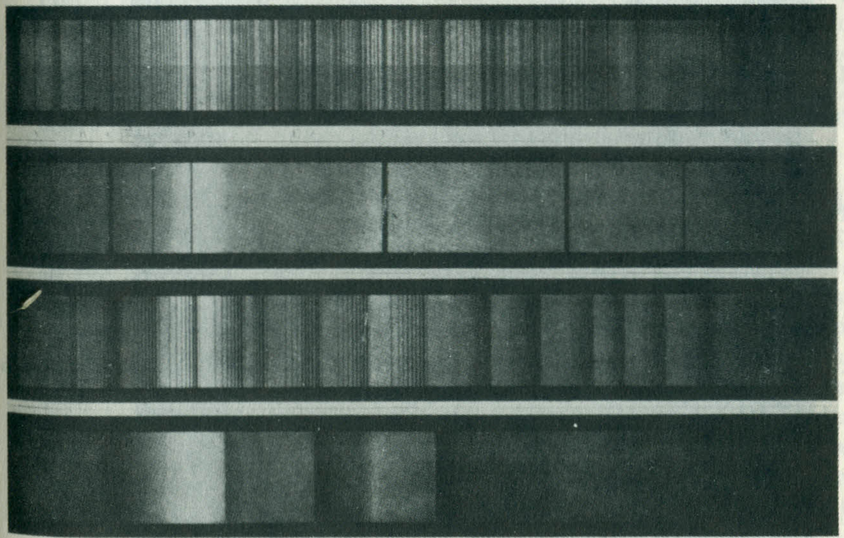
16. Constellation of Orion by Argelander from *Uranometria Nova* (p. 444)



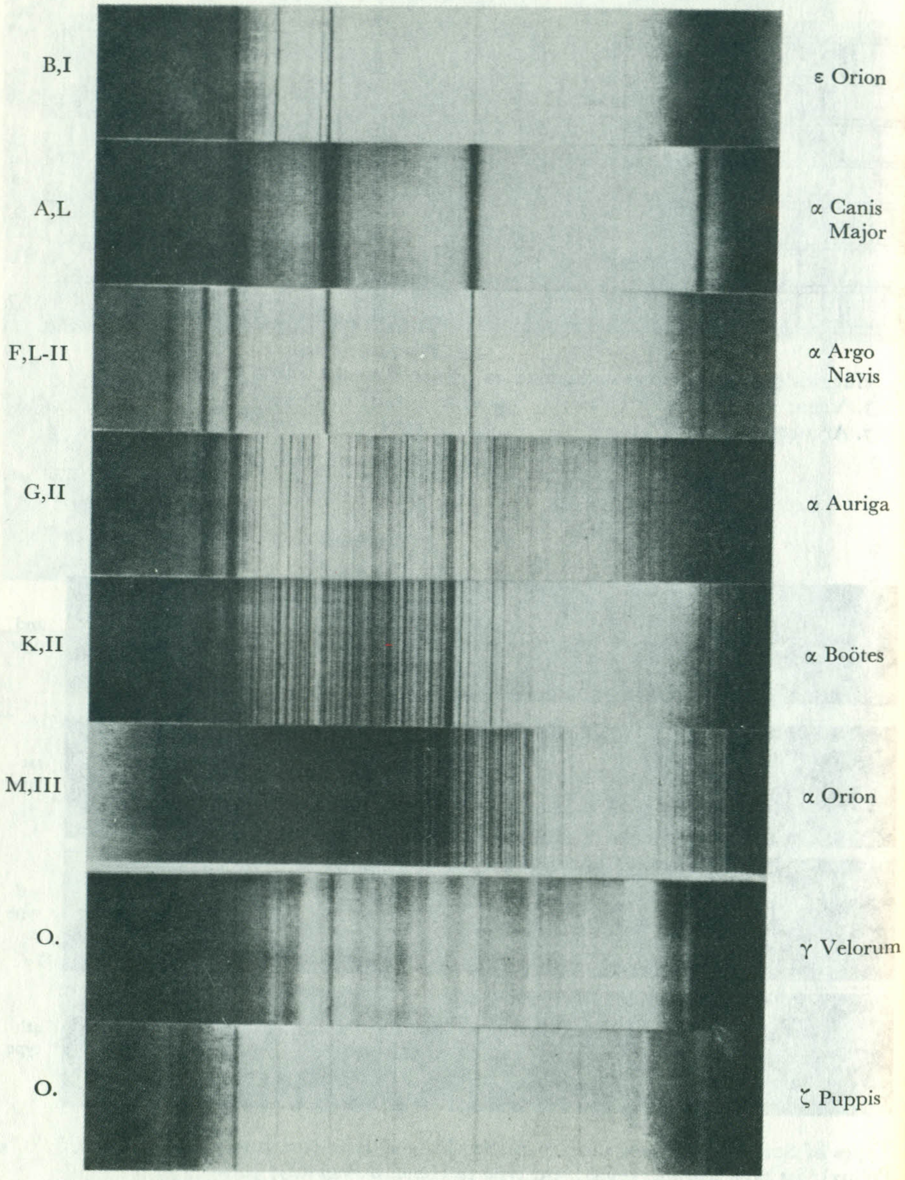
Constellation of Orion from Bayer's *Uranometria* (p. 215)



17. Huggins' first ultra-violet stellar spectra (*Philos. Transact.* 1880)
1. Vega; 2. Sirius; 3. η Ursae; 4. Spica; 5. Altair; 6. Deneb;
7. Arcturus (p. 452)



Types of Star Spectra according to A. Secchi. 2nd type (the sun, Pollux); 1st type (Sirius, Vega); 3rd type (α Henc. β Pegasus); 4th type (152 Schjellerup) (p. 450)



18. Harvard types of stellar spectra (p. 453)

chromospheric spectrum indicated as its cause the presence of radiations of far greater ionizing power than could be produced by a surface temperature of 6,000°. The emission lines of the chromosphere were too broad to correspond to this temperature and indicated a greater velocity of the particles, i.e. a higher temperature, of the order of 30,000°. In the continuous spectrum of the corona, which is reflected sunlight far outside the sun, the Fraunhofer lines were invisible; the possibility was suggested that they could be smoothed out by rapid molecular motion. In 1930 Lyot had succeeded in solving the difficult problem of observing the coronal lines in daylight by carefully excluding all disturbing light coming from diffraction, impurities and atmosphere. The lines could now be studied at ease, and they appeared to be so broad that the atom velocities producing them demanded temperatures of millions of degrees. It seemed incredible, until in 1941 came the unravelling of the mystery of the 'coronium' lines. Following an indication by Grotrian, the Swedish physicist Bengt Edlén established that the strongest among them were emitted by 9, 10 and 13-fold ionized iron atoms, 11 and 12-fold ionized calcium, and 11 to 15-fold ionized nickel atoms. To strip off all these strongly attached electrons, temperatures of many hundred thousands or millions of degrees were needed. Thus it has been ascertained that the sun, with its surface at 6,000°, is surrounded up to a great distance by a widely extended atmosphere of extremely thin and hot gases. It contributes hardly anything to the total radiation of the sun but adds to it an appendage of strongly ionizing radiations of very short wavelength. How this heating from above influences the stormy processes in the prominences remained an object of further research.

A complete answer to the question, What is the sun? has to include knowledge of its interior. Important progress was made on this subject in the same years about 1920, on the basis of the above-mentioned discovery of the exact laws of radiation. In the years 1878–83 the German physicist A. Ritter had already tried by theoretical deductions to derive results about the interior of the sun. Since, however, the phenomena of radiation constituted at the time an unknown chapter of physics, the right basis to his work was lacking. In 1907 R. Emden constructed a general numerical theory of gaseous spheres in space, applicable to the sun and the stars. He again knew of no other mechanisms of heat transfer than convection and conduction; so the results were unsatisfactory. Yet he could give a good explanation of the granulae visible at the sun's surface, considering them as vortex elements in ascending and descending streams. Where Ritter and Emden had failed, Eddington succeeded when our knowledge had matured. In 1916 Eddington (1882–1944) started his series of theoretical researches, extending Schwarzschild's earlier work on radiative equilibrium to the inner layers of a star and making use of the numerical computations of Emden. Now it was possible to compute exactly the physical conditions of matter at any point of the interior as a function of its distance from the centre: its temperature, density, pressure, ionization and coefficient of absorption. The values found looked fantastic: at the centre of the sun a temperature of 18 million degrees and a pressure of 9,000 million atmospheres; yet they were products of exact computation. This does not mean that they are final and absolute; in such researches there remains a certain arbitrariness in the assumed basic data which must be introduced in a simplified way. Matter, under these extreme conditions not directly observable and manageable, because they far exceed our experimental possibilities, now became an object of science because its effects present themselves in the very existence of the sun and the stars.

CHAPTER 38

PASSING LUMINARIES

Aristotle had placed fire as the highest sphere of the unsteady earthly world; all the transitory luminous phenomena not admissible into the celestial realms of eternal ether were given a place there above the air. Here his intuition was not so very different from what later science taught. All that was above he called meteors, the higher things; this name was later restricted to the phenomena which were popularly called 'shooting stars'.

For many centuries shooting stars did not attract the attention of scientists; they were considered to be a kind of lightning. To learn something about their character, one has first to answer the question about their place. At what height do they flare up and become extinguished? In the last years of the eighteenth century, in 1798, two young scientists, Benzenberg and Brandes, students at Göttingen, tried to answer these questions by simultaneous observation at two different stations. Identical times of observation, after carefully co-ordinating their watches, ensured that they observed identical objects; the different apparent courses as projected among the stars made it possible to compute the heights of the objects. The observers soon perceived that these heights were far greater than they had suspected, not some few hundreds or thousands of feet but tens, even hundreds of miles, so that they had to make their observing points much farther apart. Shooting stars appear and disappear far above the cloud region of the atmosphere, in and even above the layers called in modern times the 'stratosphere'. Though their duration was difficult to estimate, it was apparent that with such distances the real velocities amounted to tens of kilometres per second and were comparable to planetary velocities.

So the meteors were shown to be objects which from planetary space penetrate into our atmosphere, where, because of the resistance of the air, they lose their velocity and are burned up. Benzenberg supposed that they might be little stones erupted in former times by the volcanoes of the moon, which after various wanderings through space at last hit the earth. That they came from the farther depths of the solar system was demonstrated in 1833 by a magnificent meteor shower. For more than cores in 1846 and had disappeared after 1852. It was assumed that the former split had been the beginning of a wholesale dissolution into small invisible fragments. But meteors from Andromeda were now often seen, not in December but at the end of November. Computation of its perturbations by the planets showed that the node of its orbit regressed rapidly, so that the dates of meeting came earlier. Because the comet would come near to the earth in 1872, Galle predicted an abundance of Andromedid meteors for November 28th of that year. And, indeed, on the evening of November 27th a fiery rain, lasting for many hours, burst over Europe. Thirteen years later, after two more revolutions, the spectacle was repeated with equal splendour; the comet itself remained invisible, following the observed swarm by 60 days according to the computation. At the next appearance in 1892, on the nights of November 23rd–26th, only a small number of meteors was seen.

Thus those dozens of years of exciting celestial fireworks came to an end. In 1899 many observers on November 14th kept watch for the Leonids, but no meteor shower appeared; only a somewhat greater number of shooting stars than on other nights. Computations by Berberich and by Downing disclosed that the perturbations through the great planets had displaced the orbit, so that the swarm now passed the earth at twice the former distance. So meteor showers from the Lion can never again be expected. Swarms of meteors become visible as showers of shooting stars only under special conditions; and numbers of them may describe their orbits through the solar system unperceived by us. Of course, perturbations may change other orbits so as to make them intersect the earth's orbit and produce a new meteor shower.

Observation of shooting stars during the entire nineteenth century was a regular field of work for amateurs. It consisted partly in corresponding observations at two stations, to derive the height of their appearance and their extinction; partly it was one-man work to find out radiants. Denning at Bristol, after many years of observing, in 1899 published a catalogue of 3,000 radiants, but they only partly represented real swarms; conspicuous separate fireballs were among them. Olivier in 1920 made a list of 1,200 radiants, half of which he assumed to be real swarms. Every night among the many isolated meteors there are some belonging together in groups. All the swarms in space are gradually dispersed along, as well as outside, their orbits and become less marked, while the number of solitary objects increases. That the space between the planets is full of small bodies condensed in the vicinity of the sun is manifest in the zodiacal light, mentioned first by Cassini in 1683 as a special phenomenon. Its brightness decreases with increasing angular distance from the sun; but in 1854 Brorsen pointed out that at 180° distance, just opposite the sun, the brightness increases in the so-called

Gegenschein ('counterglow'), which, as a faint nebulous patch, wanders along the ecliptic. Zöllner explained it by means of the sharp maximum of brightness which every planet shows at opposition.

Meteor tracks sometimes appear on stellar photographs, but then the exact instant is uncertain. Photography was used at the Harvard Observatory to determine the most difficult datum—the duration and velocity of the phenomenon. For this purpose the track was interrupted every tenth of a second by a rapidly rotating screen before the camera. The velocities found, the result of atmospheric resistance, are employed to study the conditions in the highest layers of the earth's atmosphere.

The connection found in the nineteenth century between meteors and comets had disclosed that the meteors come from the vast spaces outside the planetary system and that the comets consist of extensive thin clouds of smaller and larger meteoric bodies. Is this all that can be said?

We have already met with comets, but chiefly as objects of fear and objects for computation of orbits. From olden times comets impressed man as imposing celestial luminaries, mostly appearing as big stars with an appendage which we call its 'tail' but which in former times was mostly described as hairs(the word *kometes* means 'hairy') or a beard. Their uncommon aspect, sometimes a bright band of light spanning half the sky, sometimes a faint ghostly torch or only a wisp, combined with their sudden appearance and arbitrary course between the stars, made them, especially in perturbed times, awe-inspiring phenomena. Chiefly at the end of the Middle Ages and in the centuries that followed they were, apart from regular astrology, considered as heralds of calamities and signs of divine wrath. The appearance of Halley's comet in 1456, a few years after the fall of Constantinople, with its scimitar-like shape, was considered as announcing the Turkish peril for Europe. Even after Tycho had shown that comets were celestial bodies in world space, comet fear remained. It did not abate until, after Newton and Halley had computed their orbits and predicted their return, the prediction was verified in the rationalist climate of the eighteenth century. Yet an instinctive uneasiness remained, which, in the nineteenth century when large numbers were discovered, became the fear of a catastrophic collision between the earth and a comet. Most of them, certainly, were innocent faint nebulae; only the bright ones developed the tails which were feared so much.

Where did these tails come from? As early as 1531 the 'Imperial Mathematician', Bienewitz (latinized 'Apianus') remarked that comet tails are turned away from the sun. Kepler in 1618 said that the luminous matter was driven away from the comet's head by a repulsive power from the sun. With the great comet of 1811 Olbers perceived with his telescope a starlike nucleus, surrounded at some distance by a parabolic envelope passing into the tail. In a spraying fountain the separate jets, curving downwards, also show a parabolic envelope; so he supposed that matter ejected from the nucleus in different directions and repelled by the sun, in streaming into the tail showed the aspect of the luminous envelope. He pointed out that different repelling forces might produce tails of different forms: 'The longer and straight tail [of the comet of 1807] must consist of particles repelled more strongly by the sun than the matter forming a curved tail.'[210] And on the nature of the repelling force, he said: 'It is difficult to restrain oneself from thinking of something analogous to our electrical attractions and repulsions.'[211]

The study of these luminous phenomena was continued at every appearance of a great comet. When Halley's comet returned in 1835, Bessel observed a fan-shaped jet of luminous matter directed towards the sun, oscillating in a period of 4½ days, obviously driven sideways by solar repulsion. Bessel developed a theory of the shape of the tail and computed the repelling force of the sun as compared with its gravitational attraction. That the tail is formed continuously by new rapidly outflowing matter was shown by the brilliant comet of 1843, which in two hours made a turn of 180° near the sun's surface, the tail remaining directed away from the sun. In the head of Donati's comet, a beautiful appearance in the autumn in 1858, several luminous envelopes were visible, which proceeded from the nucleus, expanded while moving outward, and then dissolved; sometimes luminous fans connected them with the nucleus (plate 14). Two phenomena always stood out clearly; the streaming of luminous matter out of the nucleus on the sunward side and the repulsion of the matter by the sun. The Russian astronomer Th. Bredichin in Moscow, in a number of papers from 1862 on, continued Bessel's work on the repulsive force. He found that long, straight tails are produced by a repulsion 15 to 20 times greater than the attraction; the best-known curved tails, like a scimitar, develop if the repulsion is about once or twice the attraction, whereas in the case of heavy particles, for which the attraction far surpasses the repulsion, short, strongly-curved tails appear.

The physical character of the repulsive force could not be derived from the observations and remained a matter of speculation. Most investigators spoke of electrical forces, since other cases of repulsion were not known. In 1900 the prominent Swedish chemist Svante Arrhenius, author of a couple of remarkable astronomical works, pointed out that light pressure, theoretically derived by Maxwell, and afterwards confirmed by subtle experiments, could play a role. A theoretical discussion

by Schwarzschild in 1901 showed that this pressure, for particles smaller than one micron, could amount to 18 times the gravitation at most. That stronger repulsive forces had sometimes been observed, Arrhenius explained by assuming the particles to be porous like smoke puffs, combining great surface with small weight.

Photography was used here with unexpected results, after the right method had been found. Though an amateur with a common portrait camera in 1858 had made a good photograph of Donati's comet in seven seconds, Draper in 1881 with his telescope needed an exposure of 2½ hours to get a good picture of Tebbutt's comet 1881 III. And not until a professional photographer at Cape Town in 1882 had once more made a beautiful picture of Cruls's comet, did astronomers realize that it was not the linear aperture, but the angular aperture of the lens that mattered. A comet, unlike a pointlike star, is a faint luminosity extending over a great surface; in order to have a great surface intensity acting upon the plate, the aperture must be large relative to the focal distance. Once this was realized, astronomers began to use portrait lenses and doublets in photographing all kinds of faint extensive celestial luminosities. At the Lick Observatory pictures of Swift's comet were made in 1892 by E. E. Barnard and of Rordame's comet in 1893 by J. W. Hussey. With increased exposure time, comet tails, formerly smooth and ghostly, hardly visible phantoms, now appeared on the plates as brilliant torches with rich detail of structure, with bright and faint spots, never before seen or even suspected. Such photos, taken of every succeeding bright comet (like Morehouse's comet in 1908 and Halley's in 1910, often reproduced in scientific and popular reviews) gave a new impulse to the study of comet tails (pl. 14). Comparison of pictures on consecutive days showed that the bright condensations moved away from the head. Their velocity could be derived; Heber D. Curtis, also a Lick astronomer, found for Halley's comet that the velocity increased with the distance from the nucleus, from 5 to 10 up to 90 km. per second. The same phenomenon—the repulsion of cometary matter by the sun—formerly derived by theoretical discussion from telescopic structural changes in the comet's head, was now made directly visible in small-scale images of the total object.

The origin and nature of the comet's light could be ascertained only by spectrum analysis, which was developed in the same sixties in which comets were recognized as meteor swarms. It was immediately applied and used for every new comet. As faint nebulae at great distances, they showed only the reflected solar spectrum. When, however, in approaching the sun they became bright and formed a tail, a new spectrum appeared. Donati, observing Temple's comet in 1864, perceived a spectrum of three faint luminous bands, yellow, green and blue.

Huggins in 1868 succeeded in identifying these bands—which were sharply cut off at the red side—with the bands emitted by ethylene vapour and other hydrocarbon compounds made luminous by electrical discharges, described earlier by Swan (pl. 15). Huggins said the comet's light was emitted by luminous carbon gas. Meteors heated in a laboratory give off gases consisting of carbon monoxide and dioxide, hydrogen and hydrocarbons; so it could be understood that swarms of such stones, heated in approaching the sun, produced atmospheres about their nucleus consisting of the same gases.

Improvement of the spectral apparatus and its application to all the comets in the ensuing decades led to a continual increase of knowledge. Photographic spectra showed a number of ultraviolet bands, chiefly belonging to the spectrum of the carbon-nitrogen compound cyanogen. A new and unexpected phenomenon was presented by Wells's comet 1882 I; when it approached the sun, the hydrocarbon band spectrum disappeared and was replaced by the bright yellow line of sodium. The same sodium line, now accompanied by several iron lines, was shown by the brilliant September comet, 1882 II, which was visible in the daytime at the sun's limb and grazed its surface in perihelion passage; here the metal lines disappeared and gave way to the hydrocarbon bands when the comet had receded to a greater distance from the sun. Also later comets, 1910 I and 1927 IX, when near to the sun, showed the sodium emission; it was a regular phenomenon belonging to a great intensity of solar radiation.

The amount of matter involved in these phenomena was determined by Schwarzschild and Kron in an exhaustive study of some few photographs of Halley's comet, taken by chance, but standardized. They measured the surface intensity of the tail ½° behind the head. By assuming that the light here was not reflected sunlight but solar energy absorbed and emitted as new radiation (called 'fluorescence'), they deduced that 150 gm. of matter flowed per second from the head into the tail. The density of this luminous matter then was 10^{20} times less than that of our air, i.e. one molecule per cubic centimetre, scarcely more than is assumed for empty outer space. It can hardly be called a 'gas', since in a gas the billions of collisions of the particles, through interchange of energy, produce average conditions of equilibrium. In these comet tails, however, every molecule or atom, once ejected and propelled by solar repulsion, runs its course without being disturbed by mutual collisions.

The processes indicated in these older researches by the vague names of 'luminescence' and 'fluorescence' acquired determinate precision after 1914, through Bohr's atomic theory of spectral lines. Every atom, because of internal changes of energy, emits or absorbs certain determinate wavelengths. The rapid increase in knowledge of atomic line spectra was soon followed by an analogous increase in knowledge of the molecular band spectra. Molecules consisting of two or more equal or different atoms have far more additional freedom of internal change by absorbing or emitting energy in small and slightly different bits; so the spectrum shows an abundance of close lines, which in small instruments appear as bands. The unravelling of molecular spectra since 1920 found its application in the study of the ever-increasing number of various bands disclosed in cometary spectra by the improved spectrographs. This study revealed that the Swan spectrum had nothing to do with hydrocarbons; it was radiated by the carbon molecule C_2. Next to it the cyanogen molecule, CN, contributed most to the light of the comet's head. The fainter light of the tail is chiefly due to ionized carbon monoxide (CO^+). Fainter bands are due to such molecules was CH, CH_2, OH, NH, and (in the tail) N_2^+. Chemically, such molecules would be called free radicals, unable to exist in a gas in equilibrium where disintegration of compounds is neutralized by recombinations. In the extremely thin comet matter, collisions and recombinations are practically absent, so that what has been dissociated or ionized by solar radiation remains in this condition. When the swarm of meteoric blocks and stones comes near to the sun, they give out gases as they do when heated in a laboratory: hydrogen, nitrogen, carbon dioxide, hydrocarbons. When these molecules absorb the right wavelengths out of the solar radiation, they are split; CO_2 produces CO and CO^+; C_2N_2 produces CN; hydrocarbons produce C_2 and CH. They form the nebulous envelope of about 100,000 km. diameter called the 'comet's head', and by radiation pressure they flow into the tail.

Modern atomic theory has attributed a new character to this radiation pressure. Instead of the pressure exerted by the total solar radiation upon globular 'small particles' treated by Schwarzschild, we are now dealing with atoms and molecules pressed by only such special wavelengths as they are able to absorb. In absorbing this radiation coming from the sun, they get an impulse directed away from the sun; the opposite back-push, when they emit the same radiation, is, on the average, directed evenly to all sides and has no effect upon the motion. In 1935 Karl Wurm computed that in the case of the CO^+ molecule these impulses and emissions are strong and numerous; they explain the strong intensity of its spectrum in the tail and produce a great repulsive force, 80 times greater than gravitation, in agreement with the observed velocities.

It is a curious thing that such striking effects come from this simple mechanical basis: swarms of blocks and stones, not so very different from earthly matter, in their elongated orbits from far spaces coming near the sun are able, in consequence of simple, recently discovered physical processes, to produce those mysterious luminous phenomena that filled earlier generations with fright and awe and fill men today with surprise and wonder.

CHAPTER 39

PECULIAR STARS

With the nineteenth century, supremacy in astronomy switches over from the solar system to the world of fixed stars. Its realm widens with a big leap; its horizon broadens millions of times. In former centuries the stars served only as the fixed background behind the play of planetary motions in the foreground. True, their positions were carefully measured, but it was chiefly to supply a basis for the motion of the planets and the moon. Even in the first half of the nineteenth century the careful mapping of telescopic stars was done to promote the discovery of minor planets.

Now the fixed stars became increasingly the object and purpose of astronomical research for their own sake. After the pioneer work of William Herschel, the new age had to study the now opened field with more thoroughness and greater accuracy. Though it was known that the stars were distant suns akin to our sun, they were not entirely similar. The dissimilar objects, of course, attracted most attention, because here new things could be found; such peculiar objects were double stars, variable stars, and star clusters.

Double stars in the world view of that time were so peculiar that their existence in great numbers was at first doubted. People were accustomed to one sun in a system. Moreover, they were convinced that one sun, providing light and heat to the surrounding inhabitable planets, was the most suitable world system. What purpose could the Creator or Nature have with two, often unequal and differently coloured, suns? How would planets move about two such centres of attraction? Though we will never be able to see them, they yet presented an interesting theoretical problem which in later times was treated, by means of extensive numerical integrations by Ellis Strömgren at Copenhagen. The problem of the motion of the two component stars was of more direct importance. Nobody doubted that Newton's law of attraction applied also to the remote world of stars; it could be settled by observation of the double stars. Two bodies attracting each other must describe ellipses (or, more generally, conic sections) about their common centre of gravity, and the orbit of one star relative to the other has a similar figure. Seen obliquely, this (relative) orbit also appears like an ellipse, but one of different form and with different position of the focus. So it was the task of the astronomers, by means of accurate measurements of the relative positions of the components, to derive, first, their apparent orbit, and therefrom the real orbit.

Herschel had published lists of hundreds of double stars and in 1803 found relative motion of the components for about 50 of them. His pioneer work, however, was too coarse for the demands of a good determination of orbits. Double-star astronomy became the first field of application for the refined nineteenth-century instruments, with their higher standard of accuracy. It began when F. G. Wilhelm Struve (1792–1864) at Dorpat (now Tallin), who in 1819 had already measured double stars, in 1824 introduced the new 9 inch refractor, the biggest Fraunhofer instrument, to double-star work. First a survey over the new field was needed; a more complete list, containing about 3,000 double stars, was gathered together and was published in 1827. In the next ten years the distance and position angles for each of them were repeatedly measured with Fraunhofer's perfect filar micrometer. The high standard of accuracy is shown by the use of two decimals of a second in all the results. There were, or came, more workers in this field; John Herschel, South and Dawes in England, Bessel and Mädler in Germany, and Kaiser in Holland, either worked contemporaneously or followed his example; afterwards his son, Otto Struve, his successor as director of the new Pulkovo Observatory, carried on the work. Thus the motion within each pair could be followed year by year over increasingly larger parts of the orbits, and soon complete orbits could be derived for the most rapid binaries. In 1850 about 20 orbits had been computed; the smallest period of revolution (for ζ Herculis) was 31 years in an ellipse measuring 2.4″.

This kind of work continued throughout the entire century; the mere passage of time brought increase of knowledge. Nearly every astronomer with a good filar micrometer considered it his duty to take part in this work, and some among them, Dembowski in Naples and Burnham in Chicago, devoted all their time to double-star work. This general participation proved to be highly appropriate, because, firstly, smaller instruments in the right hands produced results nearly as good as big telescopes and, secondly, all results are subjected to systematic personal errors of unknown origin, which are diminished in the average of many observers. What, then, is the advantage of big instruments? It is their greater separating power. Stars which in smaller instruments appear single, or at most somewhat elongated, are seen with larger apertures as two well-separated stars which can be measured. Also faint companions are made visible by larger instruments. The construction of

ever bigger telescopes involved new discoveries, especially the separation of close pairs. First Otto Struve, with the Pulkovo 15 inch refractor, extended his father's catalogue. Then it was chiefly S. W. Burnham, who in 1873, with a small 6 inch Clark telescope, had already been able to discover, on account of his sharp vision, duplicity of stars not previously suspected. Afterwards, having at his disposal the Lick 36 inch and the Yerkes 40 inch, he added a thousand new objects to the lists, which were all carefully measured. They were mainly very close pairs, separated by less than 1″; this is often due to actually small distances involving small periods of revolution. The shortest period was 5½ years for δ Equulei, with a distance of 0.3″. In all, 5 per cent of the stars examined were found to be double.

After discovery and measurement, computation had to follow. The astronomers were faced with the problem of computing double-star

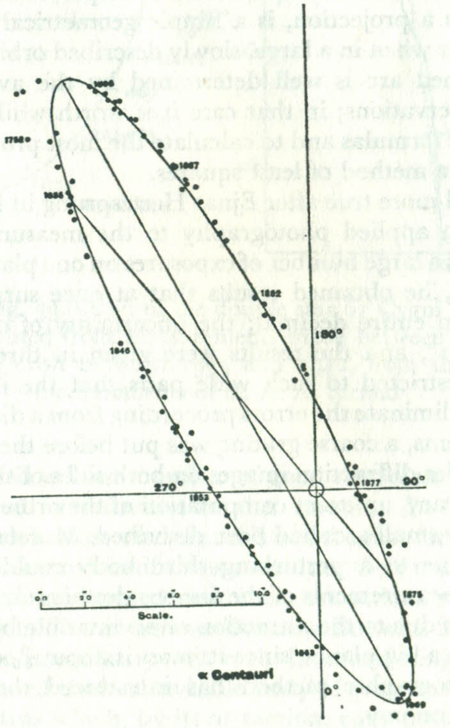


Fig. 32. Revolution of the double star α Centauri

orbits, while at the same time they were greatly occupied with planetary and cometary orbits. Naturally the solution was sought for along the same lines. An ellipse is determined by five points and can bep comuted from five measured relative positions. After Savary in 1828 had first derived and applied a method of computation chiefly by trial and error, Encke in 1830 developed a system of formulae somehow analogous to that of planetary orbits. Though often used, it was gradually found less satisfactory. The cases were too different; a small error of, say, 0.1″ on a distance of 4″ distorts the apparent course more strongly than an error of, say, 3″ does a 10° motion of a comet. So the shorter graphical method, especially for close pairs, gradually superseded the apparently more rigid algebraic method. Every measurement of distance and position angle, plotted as polar co-ordinates, determined a point on the graph; the totality of these points presents a clear picture of the apparent orbit, and an ellipse is drawn as well as possible through the assemblage of points. To construct or compute the true ellipse, of which the apparent ellipse presents a projection, is a simple geometrical problem. It is a different matter when in a large, slowly described orbit every short part of the performed arc is well determined by the average of a large number of observations; in that case it is worth while to make use of exact algebraic formulas and to calculate the most probable elements of the orbit by the method of least squares.

This was still more true after Ejnar Hertzsprung in Potsdam between 1914 and 1919 applied photography to the measurement of double stars. By taking a large number of exposures on one plate and measuring them carefully, he obtained results that at once surpassed the visual measures by an entire decimal; the uncertainty of one plate was no more than 0.01″, and the results were given in three decimals. The method was restricted to such wide pairs that the images were well separated. To eliminate the errors proceeding from a different brightness of the components, a coarse grating was put before the objective, which produced smaller diffraction images on both sides of the central image. With this accuracy, an exact computation of the orbit became possible, even when only small arcs had been described. Moreover, the reality of irregularities due to a perturbing third body could now be strictly decided. The measurements of the famous double star 61 Cygni clearly show deviations due to the attraction of an invisible body that perhaps might be called a big planet, since its mass is about $\frac{1}{60}$ of the sun's mass. Thus the photographic method has introduced a new epoch into double-star astronomy.

In the quiet progress of double-star astronomy in the nineteenth century, an exciting episode was the discovery of 'dark stars'. In 1844 Bessel showed that the bright star Sirius and also Procyon, for both of

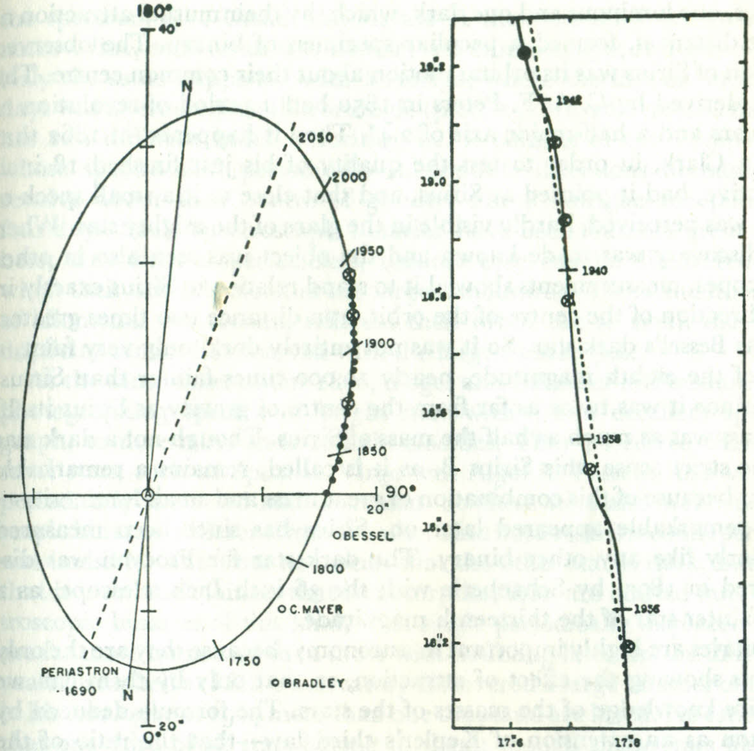


Fig. 33. Orbit of the double star 61 Cygni

Left: Ellipse computed from measurements made between 1830 and 1940. *Right:* Part of the orbit between 1935 and 1942, from the photographic measurements of K. A. A. Strand

which there existed a large number of good meridian observations of position, presented irregularities in their motion. In 1825 and 1833 Pond at Greenwich had asserted that this was the case for a large number of stars, and he was convinced it was a result of their mutual attraction. Bessel, however, demonstrated that such attractions must be imperceptible on account of the enormous distances. But with Sirius the matter was different; the irregularity was real. In considering various possibilities, he concluded that an invisible body of great mass must exist in the vicinity of Sirius which, by its attraction, caused the irregularities. It might seem a piece of irony that the brightest of our stars should be subject to a dark body, reversing the relation held until then as the only natural order of things. In reality, however, we had two equivalent