Cambridge University Press
978-1-107-66936-9 - The Cosmic Century: A History of Astrophysics and Cosmology
Malcolm S. Longair
Excerpt
More information

Part I

Stars and stellar evolution up to the Second World War

1 The legacy of the nineteenth century

1.1 Introduction

The great revolutions in physics of the early years of the twentieth century have their exact counterparts in the birth of astrophysics and astrophysical cosmology – these astronomical disciplines scarcely existed before 1900.

The history of the interaction between astronomy and fundamental physics is long and distinguished. From the birth of modern science, astronomy has provided scientific information on scales and under physical conditions which cannot be obtained in laboratory or terrestrial experiments. There is no better example than the history of the discovery of Newton's law of gravity, which provides a model for the process by which astronomical discovery is absorbed into the infrastructure of physics.[1] The technological and managerial genius of the great Danish astronomer Tycho Brahe (1546–1601) and his magnificent achievements in positional astronomy during the period 1575 to 1595 provided the data which led to the discovery of the three laws of planetary motion of Johannes Kepler (1571–1630) during the first two decades of the seventeenth century. The technical skill of Galileo Galilei (1564–1642) in telescope construction resulted in his discovery in 1610 of the satellites of Jupiter, which were recognised as a scale-model for the Copernican System of the World. Finally, in an extraordinary burst of scientific creativity, Isaac Newton (1643–1727) used Kepler's laws to discover the inverse square law of gravity and synthesised the laws of mechanics and dynamics into his three laws of motion. This story is too well known to need further comment, except to emphasise its astronomical roots – Newton had unified the laws of celestial mechanics with those of free fall on Earth. It is difficult to top this achievement in any branch of the physical sciences, but it illustrates beautifully the intimate relation between the astronomical and physical sciences – this is a theme which will be emphasised throughout this book.

Until the late nineteenth century, astrophysics as such did not exist. Astronomy meant positional astronomy, and the techniques of accurate observation had improved steadily since the time of Tycho Brahe. The accurate measurement of the motions of the Sun, Moon and planets against the background of the fixed stars had a practical application as a means of keeping track of time and of measuring position at sea. One of the early by-products of accurate time keeping was the first reasonably accurate measurement of the speed of light in 1676 by the Danish astronomer Ole Rømer (1644–1710), who observed that the interval between eclipses of Jupiter's innermost satellite, Io, by the planet was greater when the Earth moved away from the planet and was shorter when the Earth moved towards it. Interpreting these differences as resulting from the changing distance between the Earth and Jupiter, Rømer found a value for the speed of light of $c = 225\,000\ \mathrm{km\ s^{-1}}$.

All observations were made by eye using telescopes as large as the astronomers could afford. The revolution which was to take place at the beginning of the twentieth century can be traced to three important technical developments during the nineteenth century – the invention of astronomical spectroscopy, the first measurements of astronomical parallaxes for nearby stars and the invention of photography. To take full advantage of these developments, telescope design and operation had to be substantially improved, and the resulting instruments were to dominate the astronomy of the first half of the twentieth century. Let us review briefly these technical developments, since they were to provide the observational foundations for the great revolutions in astrophysics and cosmology in the first decades of the twentieth century.

1.2 From Joseph Fraunhofer to Gustav Kirchhoff

The first decades of the nineteenth century marked the beginnings of quantitative experimental spectroscopy. The breakthrough resulted from the pioneering experiments and theoretical understanding of the laws of interference and diffraction of waves by Thomas Young (1773–1829). It is said that his ideas on the interference of light waves were stimulated by observing the patterns of radiating ripples in the pond in the Paddock at Emmanuel College, Cambridge, where he was a Fellow Commoner. In his Bakerian Lecture of 1801 to the Royal Society of London, 'On the theory of light and colours', he used the wave theory of light of Christian Huyghens (1629–1695) to account for the results of interference experiments, such as his famous double-slit experiment (Young, 1802). In the same lecture, Young introduced the tri-chromatic theory of colour vision in its modern form. Among the most striking achievements of this paper was the measurement of the wavelengths of light of different colours using a diffraction grating with 500 grooves per inch. From this time onwards, wavelengths were used to characterise the colours in the spectrum.

In 1802, William Wollaston (1766–1828) made spectroscopic observations of sunlight and discovered five strong dark lines, as well as two fainter lines.[2] He interpreted the dark lines as delineating the four primary colours of sunlight, rather than the seven colours of the rainbow of Newton or the three colours of the tri-colour theory of colour vision (Wollaston, 1802).

The full significance of these observations only began to be appreciated following the remarkable experiments of Joseph Fraunhofer (1787–1826). Fraunhofer was the son of a glazier and he became one of the two directors of the Benediktbraun glassworks in Bavaria in 1814. The firm manufactured high-quality optical glass for military and surveying instruments. Fraunhofer's motivation for studying the solar spectrum was his realisation that accurate measurements of the refractive indices of glasses should be made using monochromatic light. In his spectroscopic observations of the Sun, he rediscovered the narrow dark lines which would provide precisely defined wavelength standards. His visual

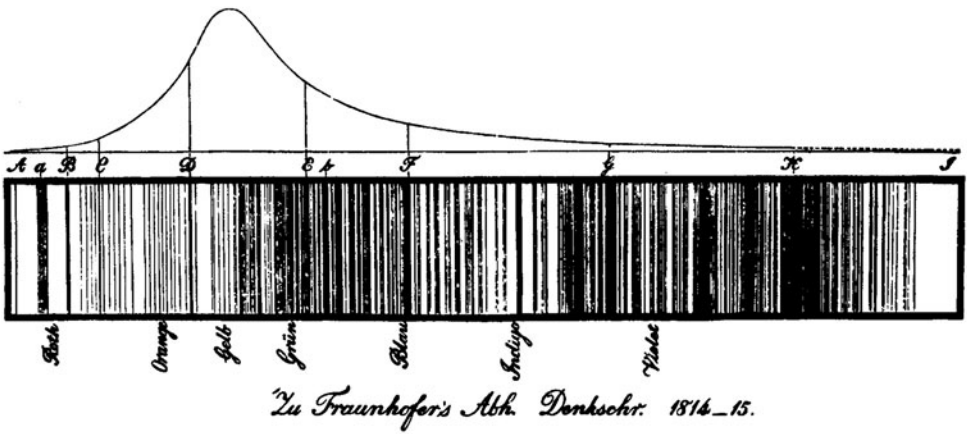

Figure 1.1: Fraunhofer's solar spectrum of 1814 showing the vast numbers of dark absorption lines. The colours of the various regions of the spectrum are labelled, as are the letters A, a, B, C, D, E, b, F, G and H, indicating the most prominent absorption lines. The continuous line above the spectrum shows the approximate solar continuum intensity, as estimated by Fraunhofer (Fraunhofer 1817a,b).

observations were made by placing a prism in front of a 25 mm aperture telescope. In his words,

> I wanted to find out whether in the colour-image (that is, spectrum) of sunlight, a similar bright stripe was to be seen, as in the colour-image of lamplight. But instead of this, I found with the telescope almost countless strong and weak vertical lines, which however are darker than the remaining part of the colour-image; some seem to be completely black.

He labelled the ten strongest lines in the solar spectrum by the letters A, a, B, C, D, E, b, F, G and H, and he recorded 574 fainter lines between the B and H lines (Figure 1.1); see Fraunhofer (1817a,b).[3] This notation is still used to describe the prominent absorption lines in the spectra of the Sun and stars.

From the technical point of view, a major advance was the invention of the spectroscope with which the deflection of light passing through the prism could be measured precisely. To achieve this, Fraunhofer placed a theodolite on its side and observed the spectrum through a telescope mounted on the rotating ring (Figure 1.2).

In a second paper, Fraunhofer measured the wavelengths of what are now referred to as the *Fraunhofer lines* in the Solar spectrum using a diffraction grating, which consisted of a large number of equally spaced thin wires (Fraunhofer, 1821) – he was one of the early pioneers in the production of diffraction gratings. He found that the wavelengths of these lines were stable and so provided accurate wavelength standards. In addition to his observations of the Sun, Fraunhofer was the first to make spectroscopic observations of the planets and the stars. In his papers of 1817, he reported the observation of Fraunhofer lines in the spectrum of Venus, inferring that the spectrum was the same as sunlight. In the case of the first magnitude star Sirius, he found, to his surprise,

> . . . three broad bands which appear to have no connection with those of sunlight.

Figure 1.2: A portrait of Fraunhofer with his spectroscope (Courtesy of the Deutsches Museum, Munich). This portrait is located in the Hall of Fame of the museum.

In 1823, Fraunhofer made further observations of the spectra of the planets and the brightest stars, anticipating by about 40 years the next serious attempts to measure the spectra of the stars (Fraunhofer, 1823). He concluded that the stars have dark lines in their spectra similar to those seen in the Sun, but that the lines present differ from star to star.

From the perspective of the glass industry, Fraunhofer was then able to characterise the chromatic properties of glasses and lenses quantitatively and precisely. These developments led to much superior glasses, as well as to much improved polishing and testing methods for glasses and lenses. These technical improvements also resulted in the best astronomical telescopes then available. Fraunhofer's masterpiece was the 24-cm Dorpat Telescope built for Wilhelm Struve at the Dorpat, now Tartu, Observatory in Estonia. In addition, he built a heliometer for Friedrich Bessel at Königsberg, to which we will return in Section 1.3.

The understanding of the dark lines in the solar spectrum had to await developments in laboratory spectroscopy. In his first report of the multitude of lines in the solar spectrum, Fraunhofer had noted that the dark D lines coincided with the bright double line seen in lamplight. In 1849, Léon Foucault (1819–1868) performed a key experiment in which sunlight was passed through a sodium arc so that the two spectra could be compared precisely. To his surprise, the solar spectrum displayed even darker D lines when passed through the arc than without the arc present (Foucault, 1849). He followed up this observation with an experiment in which the continuum spectrum of light from glowing charcoal was passed through the arc, and the dark D lines of sodium were found to be imprinted on the transmitted spectrum.[4]

Ten years later, the experiment was repeated by Gustav Kirchhoff (1824–1887), who made the further crucial observation that, to observe an absorption feature, the source of the light had to be hotter than the absorbing flame. From these considerations, Kirchhoff concluded that sodium was present in the solar atmosphere. These results were immediately followed up in 1859 by his understanding of the relation between the emissive and absorptive properties of any substance, now known as *Kirchhoff's law* of emission and absorption of radiation (Kirchhoff, 1859). This states that, in thermal equilibrium, the radiant energy emitted by a body at any frequency is precisely equal to the radiant energy absorbed at the same wavelength. From thermodynamics arguments, he was able to show that there must be a unique spectrum of radiation in thermal equilibrium, which depended only upon temperature and frequency.[5] This profound insight was the beginning of the long and tortuous story which was to lead to Planck's discovery of the formula for black-body radiation and the inevitability of the concept of quantisation over 40 years later.

Throughout the 1850s, there was considerable effort in Europe and in the USA aimed at identifying the emission lines produced by different substances in flame, spark and arc spectra. The fact that different elements and compounds possessed distinctive patterns of spectral lines was established, and attempts were made to relate these to the lines observed in the solar spectrum. In 1859, for example, Julius Plücker (1801–1868) identified the Fraunhofer F line with the bright Hβ line of hydrogen, and the C line was more or less coincident with Hα, demonstrating the presence of hydrogen in the solar atmosphere. The most important work, however, resulted from the studies of Robert Bunsen (1811–1899) and Kirchhoff. In Kirchhoff's great papers of 1861 to 1863 entitled 'Investigations of the solar spectrum and the spectra of the chemical elements', the solar spectrum was compared with the spark spectra of 30 elements using a four-prism arrangement with which it was possible to view the spectrum of the element and the solar spectrum simultaneously (Kirchhoff, 1861, 1862, 1863). He concluded that the cool, outer regions of the solar atmosphere contained iron, calcium, magnesium, sodium, nickel and chromium and probably cobalt, barium, copper and zinc as well.

1.3 The first stellar parallaxes

From the seventeenth century onwards, most astronomers assumed that the stars were objects similar to the Sun, but at vastly greater distances.[6] The method of distance determination used by Newton and others involved assuming that the Sun and stars have the same intrinsic luminosities, a procedure known as the method of *photometric parallaxes*. Then, the inverse square law can be used to measure the relative distances of the Sun and the stars. The major technical problem was that the Sun is so much brighter than the brightest stars that it was difficult to obtain good estimates of the ratio of their observed flux densities, or apparent magnitudes. An ingenious solution was discovered in 1668 by James Gregory (1638–1675), who used Jupiter as an intermediate luminosity calibrator, assuming that its light was entirely composed of sunlight reflected from the disc of the planet and that its surface was a perfect reflector. Then, the apparent magnitudes of Jupiter and the bright star Sirius could be compared, and the distance of Sirius from the Earth was found to be 83 190 astronomical units (Gregory, 1668). The same method was used by John Michell (1724–1793) in 1767 to estimate a distance of 460 000 astronomical units for Vega, or α Lyrae, from the Earth (Michell, 1767).[7] This distance was about a factor of 4 smaller than that found in 1838 by Wilhelm Struve, who used the method of trigonometric parallax. The problem with this approach is that it depends upon the assumption that the intrinsic luminosities of the Sun and the stars are the same.

Direct evidence for the large distances of the stars came from James Bradley's first definitive measurements of the effects of the aberration of light caused by the Earth's motion about the Sun in 1728 (Bradley, 1728). Ever since the time of Copernicus (1473–1543) it had been realised that a test of the hypothesis that the Earth moved about the Sun would be the observation of the annual parallax of the stars. Attempts to measure these small movements of the stars had been subject to a variety of insidious systematic errors. Instead of the expected effect, Bradley (1693–1762) discovered the phenomenon of the aberration of light due to the motion of the Earth, the effect amounting to about ± 20 arcsec for the star γ Draconis. A consequence of this remarkable result was that an upper limit could be derived for the annual parallax of γ Draconis and hence a lower limit to its distance of 400 000 astronomical units,[8] a figure consistent with Newton's estimate using the method of photometric parallax published in the same year. Bradley's pioneering observations ushered in a new epoch of precision astrometry.

The first definitive distance measurements were made in the 1830s by the method of *trigonometric parallax*, the apparent motion of nearby stars against the background of the distant stars due to the Earth's motion about the Sun. Priority for the first trigonometric parallax is accorded to Friedrich Bessel (1784–1846) at Königsberg. The instrument he used was a 16-cm heliometer custom-built by Fraunhofer. The heliometer consisted of a lens cut in half to form two D-shapes, after a design by John Dollond (1706–1761). The images of separated stars could be brought together and their separation measured by the reading on a micrometer screw. Bessel used this telescope to measure the movement of the high proper motion star 61 Cygni relative to distant background stars, and he announced its parallax in 1838 (Bessel, 1839). The parallax amounted to only about one-third of an arcsec, corresponding to a distance of 10.3 light-years. Three months later, Thomas Henderson (1798–1844) published a parallax of 1.16 arcsec for the southern star α Centauri (α Cen) (Henderson, 1840), and almost contemporaneously Wilhelm Struve (1793–1864) measured the parallax of α Lyrae to be 0.12 arcsec (Struve, 1840). Henderson was unlucky not to publish the first parallax – he had measured a parallax of 1 arcsec in declination a few years earlier, but delayed publication until he had reduced his data in right ascension as well. These observations set the scale of the Universe of stars and showed unambiguously that the stars are objects similar to our Sun.

One of the key programmes for the development of astrophysics in the late nineteenth century and the early years of the twentieth century was the gradual accumulation of trigonometric parallaxes for nearby stars, but it was a difficult and demanding task. By 1900, less than 100 parallaxes for nearby stars had been measured with any accuracy.[9] The measurement of parallaxes is still the only direct method of measuring astronomical distances for stars and it remains one of the great challenges of observational astronomy. Matters improved dramatically in the final decade of the twentieth century with the magnificent set of parallaxes measured by the *Hipparcos* satellite of the European Space Agency, which has measured precision parallaxes for many thousands of stars (see Figure 3.3).

1.4 The invention of photography

The third major contribution to the development of astrophysics was the invention of the photographic process by Louis-Jacques-Mandé Daguerre (1789–1851) and William Henry Fox Talbot (1800–1877). Daguerre began life as an inland revenue official and then became a scene painter at the opera. The search for methods of recording images by what was to become the photographic process began with the discovery that some natural compounds are rendered insoluble when they are exposed to light. In the course of his experiments, Daguerre discovered that iodine-treated silver paper was also sensitive to light. By 1835, he had made the important discovery of the *latent image* which was recorded on sensitised paper, even if the light was not intense enough to darken the paper. The latent image could then be developed by exposure to mercury vapour and fixed by a strong salt solution. The use of the latent image meant that exposures could be reduced to 20 to 30 minutes. Interestingly, the announcement of the discovery of what was called the *daguerreotype* process was made by François Arago (1786–1853), the director of the Paris Observatory, on 7 January 1839.

A similar announcement was made almost simultaneously by Fox Talbot in England. One of the earliest, and for me most moving, images is the picture taken in February 1839 by John Herschel (1792–1871) of his father's 40-foot telescope. In a two-hour exposure, the support for the large tube of the telescope can be clearly seen – the telescope was dismantled in the following year (Figure 1.3). John Herschel had a passionate interest in photography and invented much of its terminology, including the terms 'photography', 'positive', 'negative' and so on.[10]

The first astronomical images were taken in the succeeding years, but the process was slow. Isolated examples of successful daguerreotype images of astronomical objects were reported over the following decade and included the Moon, a solar eclipse and the Sun. Among the most significant images of these early years of photography was the first daguerreotype spectrum of the Sun obtained by Edmond Becquerel (1820–1891) in 1842 which showed the complete spectrum of Fraunhofer lines as well as many lines in the visually unobservable ultraviolet region of the spectrum (Becquerel, 1842). The problem with the daguerreotype process was that, even for terrestrial objects, the typical exposure times were about 30 minutes. This was greatly reduced by the invention of the wet collodion process by Frederick Scott Archer (1813–1857) in 1851 (Archer, 1851). This process produced finely detailed negatives, and typical terrestrial exposures were reduced to 10 seconds. Astronomical exposures were limited to 10 to 15 minutes because the plates had to remain wet during exposure.[11] The net result was faster, fine-grained plates which quickly superseded the daguerreotype process. These inventions sparked an enormous popular interest in photography in the 1850s and many commercial photographic studios were set up. The wet collodion process was used by Julia Margaret Cameron (1815–1879) in her spectacular portraits of great nineteenth-century figures, including her famous images of the aged John Herschel.

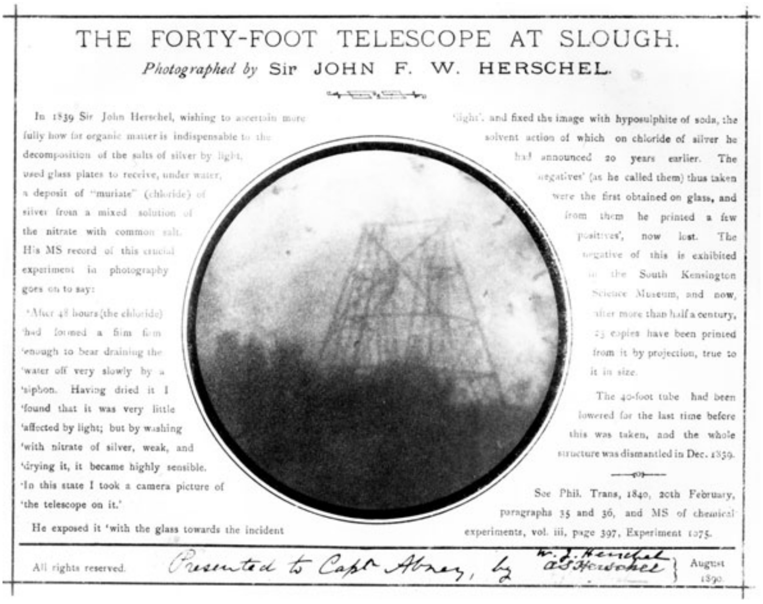

Figure 1.3: John Herschel's photograph of 1839 of part of the support structure of his father's 40-foot telescope just before the telescope was dismantled. The details of the photographic process are described in the text. (Courtesy of the Science Museum/Science and Society Picture Library.)

The story now diverges in two directions. Firstly, the wet collodion process was sufficiently fast for astronomical images and spectra to be recorded, and the search for improved photographic materials continued throughout the remaining years of the century. The boom in photography meant that there was no lack of plates for astronomical use. Secondly, telescope design had to be considerably improved. To take advantage of the use of photographic plates, it had to be possible to track and guide the telescope with very much improved precision as compared with a telescope used visually. In the latter case, the length of the exposure is determined by the response time of the eye, which is only about one-tenth of a second. Let us first complete the story of the development of photographic techniques.

The development of photographic astronomy was largely in the hands of inspired amateurs. Warren de la Rue (1815–1889) in England designed and built a photographic camera for taking daily images of the Sun from Kew Gardens in London using very short exposures. The result was a remarkably complete set of daily sunspot records for the period 1858 to 1872. The first photographic spectrum using the wet collodion process was obtained for the bright star Vega by Henry Draper (1837–1882) in 1872 (Draper, 1879). The spectrum showed the $H\gamma$ and $H\delta$ lines of hydrogen, as well as the first detections of the next seven ultraviolet lines in this hydrogen series. These ultraviolet lines were discovered by astronomical spectroscopy seven years before they were measured in the laboratory. Subsequent observations of the spectra of Vega and Sirius by William Huggins were used by the Swiss schoolmaster Johann Jakob Balmer (1825–1898) in his remarkable papers of 1885 on the *Balmer formula*, which describes the wavelengths of these lines in the spectrum of hydrogen.