Theme 7: Astrophysics

The situation at the end of the 19th century can be pictured by reading Agnes Clerke's authoritative *Popular History of Astronomy During the Nineteenth Centu*ry. There was much factual knowledge, and a start on classification, but very little understanding. **Sir Norman Lockyer** (1836–1920) had begun to argue, based on observations of solar and stellar spectra, that chemical elements could be in some way broken up—for example, that calcium "which at low temperatures gives a spectrum with its chief line in the blue, is nearly broken up in the sun into another or others with lines in the violet" (Clerke, 4th ed., p206). We now recognise this as ionisation, but at the time, given that JJ Thomson had discovered the electron only in 1897, and Rutherford had not yet discovered the nucleus, it was entirely empirical, based on comparing laboratory spectra in flames (coolish), arcs (hotter) and sparks (hottest). It was known that white stars were hotter than red, but Clerke follows Maunder in believing "that the average solar star is a weightier body than the average Sirian star" (though there were others who held the contrary—and correct—opinion). We have also seen that Clerke was confident that all nebulae were smallish objects within the Milky Way.

This situation was to be radically revised in the following 30 years. The main driver for the revision was the two revolutionary new ideas of 20th century physics: general relativity (Einstein, 1915), which paved the way for cosmology, and quantum mechanics, which led to an understanding of stars. Another contributor was the advent of the silver-on-glass (later aluminium-on-glass) mirror, using a technology first introduced in 1853 by Justus von Liebig. Glass mirrors were easier to figure than speculum metal, and the reflectivity of the silver film was superior to speculum. The 36-inch Crossley reflector, built by **Andrew Ainslee Common** (1841–1903) for the amateur Edward Crossley (1841–1905) who donated it to the Lick Observatory in 1895, was extensively used for photography by **James Keeler** and **Heber Curtis**, and established silver-on-glass reflectors as viable astronomical tools.

7.1 The structure and evolution of stars

In the late 19th century there were essentially two schools of thought relating to stellar evolution: the followers of Zöllner, who assumed that stars started out white and cooled to red, and those who felt on thermodynamic grounds that stars must start out red and then contract and heat to white, followed by a return to red as they cooled. Both approaches were essentially guesses: although the thermodynamics of self-gravitating gaseous spheres was beginning to be worked out in Germany by **Emden** and **Ritter**, who along with Lockyer was one of the early proponents of the red→white→red evolutionary model, and in the USA by **Lane**, there was an understandable reluctance to believe that stars like the Sun, with its mean density of about 1300 kg m^{-3}, could be gaseous. Relics of the Zöllner model survive in the unfortunate astronomical habit of calling OBA spectral classes "early" and GKM "late".

The first great stride towards understanding the physics of stars was taken in the early 20th century when **Ejnar Hertzsprung** (1873–1967) and (independently) **Henry Norris Russell** (1877–1907) plotted what would later become known as the Hertzspring-Russell Diagram, see figure 7.1. Hertzsprung (1911) used open clusters, where it could safely be assumed that all the stars were at the same distance; Russell (1914) used stars with measured parallax. Both demonstrated that, while (almost) all white stars were intrinsically bright, red stars came in two very distinct varieties: the luminous **giants** and the faint **dwarfs**. This finding was quickly mapped on to the red→white→red model of stellar evolution to produce the **giant-and-dwarf** evolutionary model, in which stars are born on the red giant branch, contract and heat up under their own gravity until reaching the top of the main sequence, and then gradually slide down the main sequence as they cool. Note that in all these early evolutionary models it is implicitly assumed that all stars follow essentially the same evolutionary path, although the rate at which they do so, and the brightness at which they switch from contraction to cooling, may depend upon the mass (there was no unanimity about whether more massive stars evolved faster or more slowly). Because some blue stars, such as those in the Pleiades, were surrounded by nebulosity and therefore looked as though they might be newly formed, some astronomers favoured a combination of Lockyer's and Zöllner's models, in which some stars started out white and some red, but there was no obvious theoretical motivation for this idea.

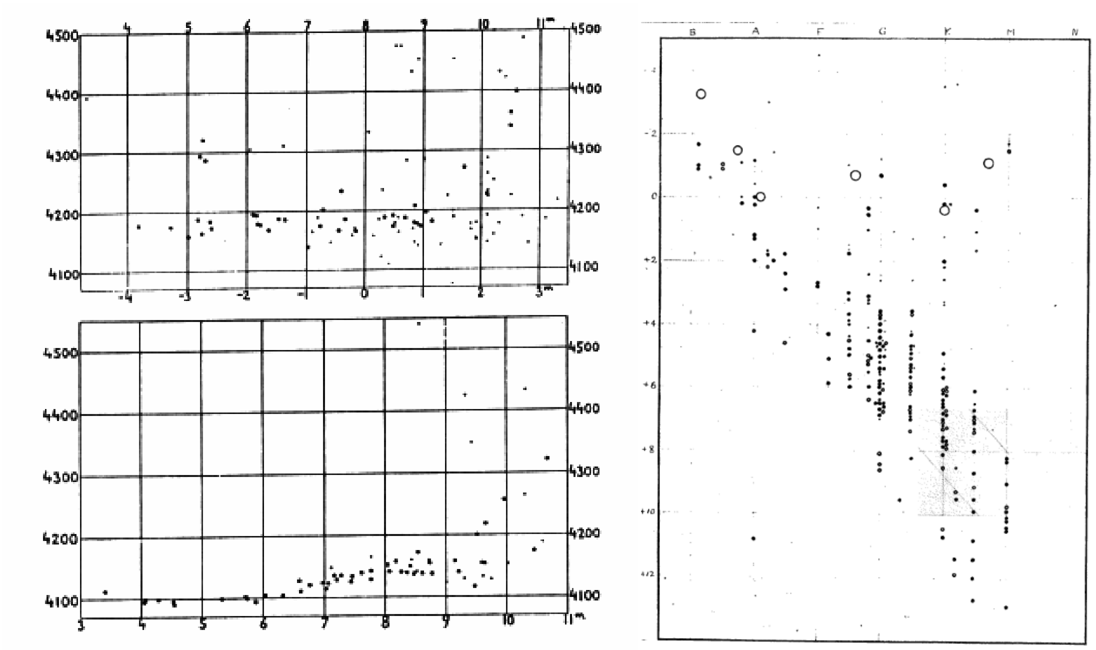

Figure 7.1: Hertzsprung's (left) and Russell's (right) Hertzsprung-Russell diagrams. Hertzsprung (1911) is plotting peak wavelength in ångströms against apparent magnitude for the Hyades (top; this must be magnitude compared to a reference star, since the brightest stars in the Hyades are certainly not magnitude −4!) and the Pleiades (bottom); Russell (1914) is plotting absolute magnitude against spectral class.
*Original sources: E Hertzsprung, Publ. Astrophys. Obs. Potsdam **22** (1911) 1–40;*
*HN Russell, Pop. Ast. **22** (1914) 331–351.*

At about the same time that Hertzsprung and Russell were discovering red giants, **R.G. Aitken** (director of the Lick Observatory) was using binary stars of measured parallax to determine stellar masses. The results had fairly large (if unquantified) error bars, since the parallaxes of the time were not particularly accurate (nor indeed were some of the orbital determinations), but should have been sufficient to demonstrate that mass and luminosity were correlated. Nobody, however, seems to have taken much notice of this until 1924, when **Arthur Stanley Eddington** (1882–1944) collected all the known stellar masses to produce the first **mass-luminosity diagram**, see figure 7.2. Eddington's main focus in this paper was the agreement of the data with his theoretical model; he was among the first astrophysicists to recognise that the fact that the material in stellar interiors is completely ionised means that it can still be treated as an ideal gas even at high densities. Notice that this realisation could not have taken place until after Rutherford's scattering experiments established the nuclear model of the atom: it's the very small size of an atomic nucleus compared to a neutral atom which makes the ideal gas approximation tenable.

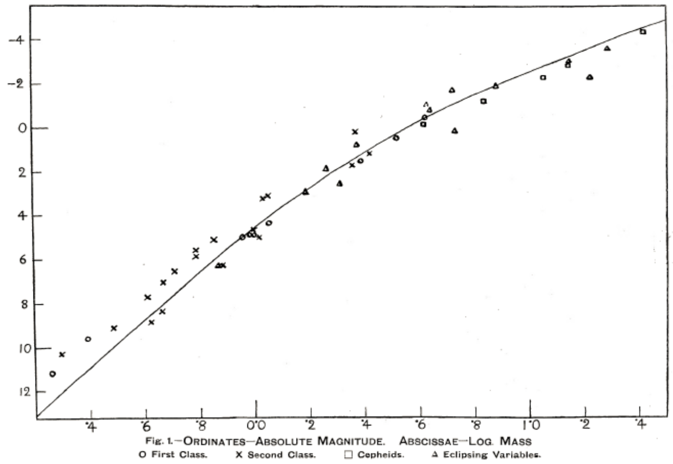

*Figure 7.2: Eddington's mass-luminosity relation, from MNRAS **84** (1924) 308–332. The curve is Eddington's theoretical calculation, normalised to Capella. Most of the data ("first class" and "second class") are from a list by Hertzsprung, Bull. Ast. Inst. Neth. **2** (1923) No. 43; the Cepheids are from Shapley (1914), and the eclipsing binaries use data from Plaskett analysed by Shapley. Eddington notes that the reason that the theory works for main-sequence stars is that "the atoms in a star are very much smaller than ordinary atoms."*

The mass-luminosity diagram immediately falsifies the giant-and-dwarf evolutionary model: if all stars trace the same evolutionary path (even at different speeds), then the idea that stars of a given mass all have the same luminosity is untenable unless they are all of the same age. The fact that Eddington's data came from field stars (not stars in a cluster) made this explanation very difficult to sustain. Eddington notes that "[the main sequence], instead of being a line of evolution, becomes, according to our view, a locus of equilibrium points". He further states that "the assumption of the giant and dwarf theory that a deviation from the gas-laws sets in at density 0.1 [g cm^{-3}] is based on a false analogy" between the hot, ionised material of stars and neutral matter at room temperature in the laboratory. Both of these arguments are correct.

Meanwhile, the emerging theory of quantum mechanics, starting with the **Bohr atom** of 1913, had begun to establish a theory of atomic spectra, and therefore a toolkit for using stellar spectra to understand the physics and chemistry of stars. By 1921, **Saha** had derived his well-known equation relating the relative populations of excited and ionised states to the temperature. In 1924, **Cecilia Payne** (1900–1979), in what Otto Struve called "the most brilliant thesis of 20th century astronomy", used Saha's equation to demonstrate, for the first time, that stars were composed primarily of hydrogen. She was sufficiently intimidated by the disbelief of senior astronomers that she initially dismissed this result as probably spurious! In fact, some other calculations being done at this time—notably Eddington's theoretical mass-luminosity relationship—also pointed to a high hydrogen content, and were likewise dismissed. This is a clear case of the ability of an existing paradigm to persist for some time in the face of contrary evidence.

By the late 1920s, however, the chemical composition of stellar surfaces (mostly hydrogen and helium with only ~1% everything else) was becoming established (though the full extent of the hydrogen content took even longer to sink in: in 1929 Atkinson and Houtermans[1] are only willing to say that "Hydrogen makes up perhaps 10% of the total mass of "early" type stars", and in 1940 McLaughlin[2] says "the sun contains probably about one-third (by weight) of hydrogen"—in contrast to the present estimate of around 67%) and the mass-luminosity relationship had provided the first unequivocal evidence that stellar evolution depended critically on mass. Meanwhile, the use of radiochemical dating in geology was beginning to provide direct evidence of the age of the Earth's surface (supporting previous indirect evidence

[1] RDE Atkinson and FG Houtermans, *Z. Phys.* **54** (1929) 656–665.
[2] DB McLaughlin, *PASP* **52** (1940) 358–372.

from geomorphology and evolutionary biology). The remaining unsolved piece of the puzzle was the continuing problem of stellar energy generation: what physical process could maintain the Sun's luminosity approximately unchanged over timescales of at least a billion years?

7.2 The problem of stellar energy generation

Anaxagoras (~500-428 BC), one of the early Greek philosophers, thought that the Sun was made of red-hot metal; his views on what prevented it from cooling down have not been recorded. By the 19th century, the issue of how the Sun could maintain its light output over geological time had become a pressing problem. Chemical reactions were clearly incapable of sustaining the Sun's luminosity over the timescales demanded by geologists (starting with James Hutton's *Theory of the Earth* in 1788 and continuing with Lyell's *Principles of Geology* in 1830) and evolutionary biologists (from Darwin's *Origin of Species* in 1859). In the absence of any other obvious energy source (radioactivity not being discovered until the last years of the 19th century), gravitational accretion seemed to be the only possible solution to the problem.

There were two main approaches to the use of gravity as a source of stellar energy, both of which essentially rely on the conversion of gravitational potential energy to heat and radiation.

- ***Mayer's meteoric hypothesis***
  A theory put forward in 1848 by JR Mayer, and independently in 1853 by James Waterston (who deserves to be remembered as an unsung pioneer of kinetic theory, rather than for this), was that comets and meteors falling into the Sun could generate energy. Assuming that the meteors fall from infinity to the Sun's surface, the gravitational potential energy lost is $GM_{\odot}m/R_{\odot} = 1.9\times10^{11}$ J/kg: to account for the Sun's luminosity, the rate of infall must be 2.0×10^{15} kg/s, or about 6.4×10^{22} kg/yr (that's about 1% of the Earth's mass, but only 3×10^{-8} of the Sun's). The problem with this model is that it increases the Sun's mass (recall that this is long before relativity—we are not converting the mass of the meteors into energy), and this affects the orbits of the planets. As early as 1854, Lord Kelvin (then plain William Thomson) showed that the change in the length of the year caused by this would be of the order of a second per year, easily detectable with mid-19th-century technology. This model is therefore unsatisfactory.

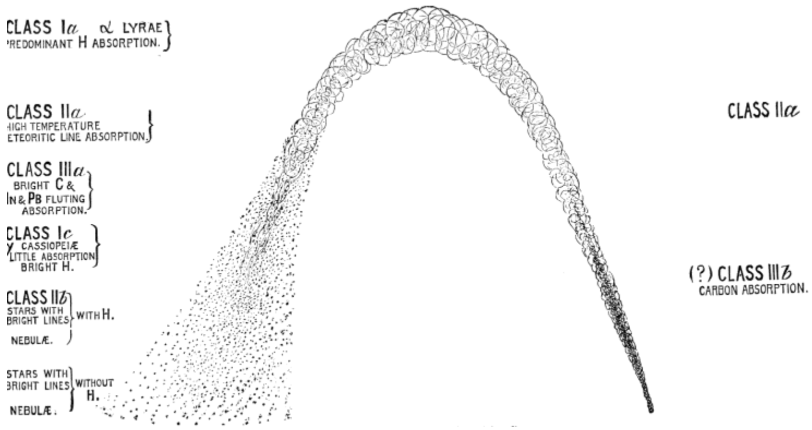

Figure 7.3: Lockyer's sketch of his meteoritic hypothesis, from The Meteoritic Hypothesis (Macmillan, 1890). Roughly, the y axis is temperature and the x axis is time: the width of the line represents the diameter of the star. The class names are Vogel's: later, Sir Norman proposed his own classification system embodying his theory (it did not catch on).

- ***Lockyer's meteoritic hypothesis***
  Norman Lockyer's approach (see figure 7.3) is summed up by his statement (*The Meteoritic Hypothesis*, 1890) that "All self-luminous bodies in the celestial spaces are composed either of swarms of meteorites or of masses of meteoritic vapour produced by heat." In Lockyer's

theory, the swarms of meteorites collapse under their own gravity, and the heat produced by this vaporises the meteorites, which subsequently condense into a single solid globe. Stars form from nebulae, and heat up while they are still condensing: Lockyer believes that such stars "do not resemble the Sun, but consist chiefly of discrete meteoritic particles". Once the meteorites have condensed into a single lump, the star has lost its power source and will gradually cool down—Lockyer regarded the Sun as being in this stage. Lockyer thought that the differences in spectral features were caused partly by temperature and partly by the spacing of the individual meteors: relatively wide spacing produced bright lines from gas, whereas condensation into a solid globe with a gaseous atmosphere produced a continuous spectrum with absorption lines.

The principal objection to this scheme is that the spectroscopic studies of nebulae did not, even at the time, really seem to be consistent with the idea that they were all swarms of meteorites: for example, planetary nebulae, which Lockyer considered to be protostars, have spectra wholly dominated by emission lines (which require hot gas).

- ***Helmholtz and Kelvin's contraction hypothesis***
  The alternative formulation of the contraction hypothesis assumes that stellar material is always gaseous, and makes use of the thermodynamic studies of self-gravitating gas spheres that were being carried out in the second half of the 19th century by Lane, Emden and Ritter. In this model, proposed by Hermann von Helmholtz in 1853 and supported by Kelvin, the Sun is always a gaseous sphere contracting under gravity, and its radiation comes from conversion of the lost gravitational potential energy. If the Sun is assumed to be a uniform sphere (which is not realistic, but only affects the numerical constant), its gravitational potential energy is $-3GM_{\odot}^2/5R_{\odot}$, and if it contracts by an amount ΔR the energy lost is therefore $3GM_{\odot}^2\Delta R/5R_{\odot}^2$. According to the work of Emden and Lane, half of this goes into heating the gas and the other half is radiated away, so to maintain the current solar luminosity we need a contraction rate of 2.4×10^{-6} m/s, or about 74 m/yr. This would not be measurable by 19th-century techniques, so could not be disproved directly. However, the total energy available from this mechanism is "only" $\sim10^{41}$ J, and this would power the Sun for only ~10 million years. By the late 19th century, this timescale was looking unreasonably short from a geological perspective.

The situation at the end of the 19th century was therefore deeply unsatisfactory. Geologists argued, from physically and empirically motivated models of geological processes (e.g. sediment deposition, salinity of seas, rate of accumulation of volcanic ejecta), that the Earth must be hundreds of millions of years old; physicists argued that the Sun could be at most a few tens of millions of years old; neither calculation seemed obviously at fault. In the early 20th century, astrophysical arguments were added to the geology: once Cepheids were recognized as pulsating variables (Shapley 1914), the fact that their periods did not seem to change measurably contradicted the contraction idea, because pulsation periods should depend on density, and therefore as a giant star like a Cepheid contracts its period should change. In 1920, Eddington said, "If the contraction theory were proposed today as a novel hypothesis I do not think it would stand the smallest chance of acceptance."

The discovery of radioactivity at the close of the 19th century provided a possible way out. Some of the radioactive elements seemed capable of generating energy for several billion years, a much more satisfactory timescale. Since the naturally radioactive elements were the heavy species like radium and uranium, this produced the interesting concept of uranium stars. However, measurements of atomic masses, first by **Edward Morley** in 1895 (demonstrating that one oxygen atom weighed significantly less than 16.0 hydrogen atoms) and later by **F.W. Aston** (who invented the mass spectrometer in 1919), coupled with Einstein's $E = mc^2$ (1905) and Payne's determination of the hydrogenic composition of stars (1924), naturally led to the idea of hydrogen fusion. Eddington, in a paper of 1920, was an early champion of this energy source.

The problem with hydrogen fusion as initially considered was fairly simple. It didn't work. Thermodynamics and kinetic theory provide an estimate of $\sim 10^7$ K for the central temperature of a Sun-like star, which gives the protons a typical kinetic energy of $3/2\ kT = 2 \times 10^{-16}$ J. The potential energy of two protons separated by 10^{-15} m, as required for fusion, is $e^2/4\pi\varepsilon_0 r = 2 \times 10^{-13}$ J—about 1000 times greater. Therefore, at this temperature, the protons will not get close enough to fuse. Eddington knew about this problem but was sufficiently convinced of the reality of hydrogen fusion to ignore it: in *Stars and Atoms* he famously tells people who contend that the centres of stars are not hot enough for hydrogen fusion to "go and find a hotter place". However, rhetoric aside, this is clearly a real issue for the hydrogen fusion model.

Fortunately, quantum mechanics again came to the rescue. The Heisenberg uncertainty principle states that the uncertainties in momentum and position are coupled: $\Delta x \Delta p > \hbar \sim 10^{-34}$ J s. This means that it is not, even in principle, possible to know both the position and the momentum of a proton with complete precision: there is, therefore, a small but non-zero probability of finding proton 1 inside the Coulomb barrier of proton 2, even though the barrier is too high for it to surmount. This is known as **tunnelling**. The idea was first introduced by **Gamow**, to describe radioactive alpha decay, and was applied to the solar energy problem by **Atkinson** and **Houtermans** in 1929. The precise probability of hydrogen fusion can be calculated, for any given temperature, by solving the Schrodinger equation for the Coulomb potential.

The exact mechanisms by which hydrogen fusion takes place in stars were both worked out by Hans Bethe and collaborators just before WWII. [In fact, the first mechanism worked out was the CNO cycle, which powers stars more massive than the Sun; stars of the Sun's mass and lower are powered by the pp chain, which was discovered a little later.] At this point, the combination of thermodynamics, kinetic theory and quantum mechanics had finally produced a proper physical understanding of the interior of a main-sequence star. The subsequent stages in stellar evolution were worked out through the 1940s and 1950s, primarily—at least as regards the nuclear reactions—by Fred Hoyle and collaborators.

One of the most interesting features of this development concerns the triple-alpha process for helium burning. This goes via the extremely unstable nucleus ^{8}Be: two helium nuclei collide to produce ^{8}Be, and in the instant before it decays this nucleus is hit by another helium nucleus to produce stable ^{12}C. The problem with this is that there appears to be nothing to stop this being hit by another helium nucleus to produce equally stable ^{16}O (there is a slightly higher Coulomb barrier, but on the other hand the carbon is stable, so you can afford to wait for a higher-than-average energy helium nucleus). Therefore it would seem that we should be talking about the oxygen-producing quadruple-alpha process as the dominant fusion process in helium-burning stars, and carbon should be a rare element (and we should probably not exist). This is not true: carbon is only slightly less common than oxygen (they are respectively the fourth and third most common elements after H and He), and the oxygen seems to be made mostly in supernovae, not in He-burning stars. Hoyle realised that this discrepancy could be avoided if the carbon production was resonant (i.e., the total energy of the three helium nuclei matched an excited state of the carbon nucleus), while the oxygen production was non-resonant—this means that carbon production is much quicker than oxygen production, so carbon does build up in the star. Unfortunately the excited states of carbon had been mapped, and no such state had been found. Hoyle was sufficiently sure of his calculations that he managed to persuade Willy

Fowler to do some more nuclear physics experiments: the required excited state was duly found (its spin-parity state is such that it is easily missed in standard experiments).

Observational evidence for stellar nucleosynthesis was provided by **Paul Merrill** (1887–1961), who detected spectral lines of **technetium** in several red giant stars. As technetium is an unstable element whose most stable isotope lives for only 4.2 million years, this element could not have been present in the stars at their birth, but must have been formed during their lifetime and brought to the surface by convection.

This work culminated in the famous paper (*Rev. Mod. Phys.* **29** (1957) 547–650) by Geoffrey Burbidge, Margaret Burbidge, Willy Fowler, and Fred Hoyle (listed in alphabetical order!), now universally known as B^2FH. B^2FH sets out to explain not just fusion products, but the production of *all* atomic species by nuclear reactions within stars. Although the details are still being refined, and a few light nuclides have been moved from stellar interiors to the early universe (it should be noted that, despite his well-known opposition to the Big Bang, Hoyle also contributed to this: the key early paper is Wagoner, Fowler and Hoyle 1967), most of the conclusions of B^2FH have stood the test of time extremely well. (One should also credit the Canadian astrophysicist Alistair Cameron (1925–2005), who was working on nucleosynthesis at the same time as B^2FH; though less well known, his work was equally important in the development of the field.)

7.3 End stages of stellar evolution

19th century theories of stellar evolution all assumed that old stars simply cooled down into invisibility. The recognition that the end stages of stellar evolution are much more dramatic than this was rather gradual, and associated with both the increased understanding of stellar structure and evolution arising from the development of quantum mechanics and the improved observational data coming from the large reflecting telescopes of the 20^{th} century—particularly the highly influential 100" Hooker telescope on Mt Wilson.

Russell's diagram included one point in the lower left corner: o_2 Eridani A, a star which was hot but faint. Two years later, another such point was added: Sirius B, which had been discovered by Bessel in 1834 from the orbital motion of Sirius A, and first imaged by **Alvan Clark** in 1862, was found by Walter Adams to be a white A-class star. These stars are extremely faint, and therefore must be very small indeed given their high surface temperatures; they were therefore called **white dwarfs**.

Surprisingly, the analysis of the Sirius system showed that Sirius B had a mass similar to the Sun's. Adams also showed that the spectral lines of Sirius B showed the general relativistic gravitational redshift expected from this combination of large mass and small size. At the time, the extraordinarily high density implied by these measurements was difficult to understand in conventional physics. Once again, the solution came from quantum mechanics: this degenerate matter is supported by quantum mechanical effects (specifically, the consequences of the Pauli exclusion principle), and not by any conventional source of pressure. The necessary calculations were carried out in 1926 by **R.H. Fowler**; his student, **Subrahmanyan Chandrasekhar** (1910–95) later worked out the upper mass limit for such stars, which is now called the Chandrasekhar limit.

White dwarfs, as is common in astronomy, were observed first and understood later. Neutron stars, on the other hand, were predicted by **Fritz Zwicky** (1898–1974) and **Walter Baade** (1893–1960) in 1934, only a few years after Chadwick's discovery of the neutron. (Around the same time, Zwicky also published the first evidence for dark matter: the anomalously high velocity dispersion of the Coma cluster of galaxies.) It was, however, by no means obvious how such an object might be observed, since it would be tiny and hence extremely faint, and the paper was essentially forgotten for 30 years. Astronomers certainly did not institute programmes to search for neutron stars: in fact, they were discovered by pure accident in 1967, when Jocelyn Bell realised that "interference" in her radio telescope trace was in fact coming from a previously unknown type of celestial source. Once the "interference" was resolved into fast regular pulses, and the discovery of a second source in a different part of the sky ruled out extraterrestrial intelligence, neutron stars were rapidly fingered as the likely culprits (nothing else was small enough). This pattern of a prediction made, neglected, forgotten, and accidentally confirmed years later is not unique in the history of astronomy (something very similar happened with the cosmic microwave background); as far as I know it conforms to none of the standard philosophies of science!

Black holes, though seemingly more exotic than either white dwarfs or neutron stars, were in fact "predicted" earlier than either: in 1784 **John Michell** hypothesised that a sufficiently massive star might not allow light to escape from its surface; he also, prophetically, pointed out that such an object might nonetheless be discoverable if it was part of a binary system. He was, of course, thinking in terms of Newtonian gravity. The modern, relativistic, concept of a black hole was worked out as a solution to Einstein's equations in 1915 by **Karl Schwarzschild** (1873–1916) and named by Wheeler in 1967. Technically, the black hole concept has not yet been directly verified observationally: we have observed compact objects which are, according to our calculations, too massive to be neutron stars, and whose behaviour differs from neutron stars in ways which suggest that the compact object does not have an observable solid surface, but this falls short of complete confirmation. The situation might be compared to the heliocentric solar system before the discovery of aberration: the observations that we do have agree extremely well with our theory, the theory makes other predictions which can be and are confirmed by experiment, and there is no satisfactory alternative: therefore, even though there is no direct proof, the theory is firmly accepted and it would be a great shock if contrary evidence were found.

White dwarfs, neutron stars and even black holes are all less massive than the stars which produced them. The lost mass forms a gaseous shell around the compact object: a planetary nebula around a young white dwarf, and a supernova remnant around a neutron star. These are fairly conspicuous, and examples of both were discovered and studied long before their role in stellar evolution was properly understood.

Messier's catalogue of fuzzy objects includes several planetary nebulae. They were first described as a distinct type of object by William Herschel, who named them for their disc-like appearance, which resembles the telescopic image of a planet. Their gaseous nature was established by **William Huggins** in the 1860s: he observed the spectra of several planetary nebulae and found them to be dominated by two green emission lines (it was known from the 1859 work of Kirchhoff and Bunsen that emission line spectra were diagnostic of gas). The emission lines did not correspond to any laboratory measurements, and were initially assigned to the unknown element **nebulium** (much as similarly unidentified lines in the solar system were assigned to helium). Helium was a genuine element, subsequently discovered on Earth; nebulium, however, resisted identification. Eventually, with the understanding of atomic spectra in the 1920s, Russell realised that this must be an exotic transition in a known element; in 1928, the lines were identified by **Ira Bowen** as a forbidden transition of doubly ionised oxygen, [O III].

Some planetary nebulae were quickly found to contain blue or white central stars. The identification of these objects with very young white dwarfs came surprisingly late: although **Menzel** in 1926 recognised that they had similar characteristics of high temperature and low luminosity, he was misled by an erroneous calculation of their masses into thinking that they could not be similar objects. Bart Bok's 1958 diagram of the evolutionary path of a solar mass star (Hoskin p273) still shows the star moving directly from the end of the horizontal branch to the white dwarf cooling line: it was not until the 1960s that the formation of planetary nebulae as a result of mass loss during helium and hydrogen shell fusion on the asymptotic giant branch was understood.

Messier's catalogue also includes the Crab Nebula. In 1928, Edwin Hubble suggested that the Crab might be associated with the "guest star" recorded by the Chinese in 1054, which had occurred in the same region of the sky. This was confirmed in 1942, when measurements of the expansion of the nebula (made by comparing new and old images) showed that it could not be more than around 1000 years old. The recorded brightness of the 1054 "guest star" demonstrated that it must be a supernova, so the Crab was the first identified supernova remnant (it also contains one of the first identified neutron stars). The association of supernovae with their remnants was thus apparent from the very beginning: the remnants of Tycho's supernova of 1572 and Kepler's of 1604 have also been identified, as have several associated with earlier Chinese observations.

The concept of a supernova is itself a 20th century idea, as discussed in the previous chapter. In the 19th century, the existence of "temporary stars" or novae was well known, and spectroscopic studies by William Huggins demonstrated that they were associated with gaseous outbursts. Indeed, the "temporary star" S Andromedae, observed in 1885 in the Andromeda Nebula, helped to convince astronomers that nebulae were not external galaxies, since it seemed far too bright to be at great distance: it was generally interpreted as a flare-up of a normal star as it passed through the nebula. It was only in the period around 1917, that **Heber Curtis** (using the Crossley reflector) and **George Ritchey** (on Mt Wilson) began to observe ordinary novae in spiral nebulae and to realise that these were uniformly much fainter than S Andromedae. Baade and Zwicky coined the term "supernova" in 1934, in the paper immediately preceding that in which they (correctly!) suggested that supernovae might be caused by the collapse of a massive star into a neutron star.

7.4 Summary: the birth of astrophysics

The development of the science of astrophysics was a complex interplay between new theoretical ideas and improved observational data. Without quantum mechanics, the observational data could not be understood; without the observations, it would not be possible to confirm the nature of the phenomena. It is probably fair to date the start of astrophysics as we know it to the 1920s, although the *Astrophysical Journal* was founded in 1895 (as "an international review of spectroscopy and astronomical physics"), and the first use of the word *astrophysics* cited by the Oxford English Dictionary comes from 1870. Astrophysics between 1870 and the 1920s would have included qualitative spectroscopic studies and the early work on the thermodynamics of self-gravitating gaseous spheres by Lane, Emden and Ritter, but any real understanding of the inner workings of stars requires quantum mechanics.

In the 1920s, progress in astrophysics was driven largely by progress in quantum mechanics. If the Schrödinger equation, rather than classical mechanics, was used to describe the behaviour of protons in stellar interiors, then fusion between protons and light nuclei could take place at the temperatures deduced (using thermodynamics) for the central regions of stars. At the time, the only possible mechanism for fusion was something approximating to the CNO cycle: the key reaction of the pp chain, $p + p \rightarrow {}^2H + e^+ + \nu_e$, could not even be conceived of before the mid-1930s, since none of the particles on the right-hand side was known to exist (the neutrino was first proposed in 1930, deuterium was discovered in 1931, and the positron and the neutron were both discovered in 1932). Thus, although the *astronomical* data on stars do not change much during this period, more *physics* data were needed before the fusion process in the Sun could be worked out.

In the same way, quantitative interpretation of stellar spectra requires the Saha equation, which was derived in the early 1920s: the reason that the first quantitative calculation of elemental abundances in the solar atmosphere was produced by Cecilia Payne in 1924 was not that she had better spectra than previous workers, but that she had the necessary theoretical tools.

The 1920s also saw an improvement in the quality of observational data, with the new 100" Hooker telescope on Mt Wilson, coupled with steadily improving photographic techniques, leading to a much improved ability to image faint objects. Among other things, this resolved the long debate about the nature of the spiral nebulae, and kick-started the science of cosmology. In 1921, Michelson used interferometric techniques to determine the angular diameter of Betel-geuse, providing yet more data for astrophysical modelling. Later, this enhanced technology would lead to the discovery of stellar populations and the study of galactic chemical evolution.

In modern times, many branches of astrophysics are driven by technology of a different kind. It is not a coincidence that the first realistic stellar evolutionary paths date from the 1960s and the FORTRAN programming language from the late 1950s: the widespread availability of computer in universities and research labs made it possible to carry out calculations which up to then had been too complicated to consider. Computer simulations have essentially replaced analytical approximations as the method of choice for modelling complex systems, and computer capacity continues to limit certain branches of astrophysics such as the modelling of core-collapse super-nova explosions.