Theme 6: The Astronomical Zoo:

discovery and classification

6.1 Discovery

As discussed earlier, astronomy is atypical among the exact sciences in that discovery normally precedes understanding, sometimes by many years. Astronomical discovery is usually driven by availability of technology rather than by theoretical predictions or speculation. Figure 6.1, redrawn from Martin Harwit's *Cosmic Discovery* (MIT Press, 1984), show how the rate of discovery of "astronomical phenomena" (i.e. classes of object, rather than properties such as distance) has increased over time (the line is a rough fit to an exponential), and the age of the necessary technology at the time the discovery was made. (I have not attempted to update Harwit's data, because the decision as to what constitutes a new phenomenon is subjective to some degree; he puts in several items that I would not have regarded as significant, omits some I would have listed, and doesn't always assign the same date as I would.)

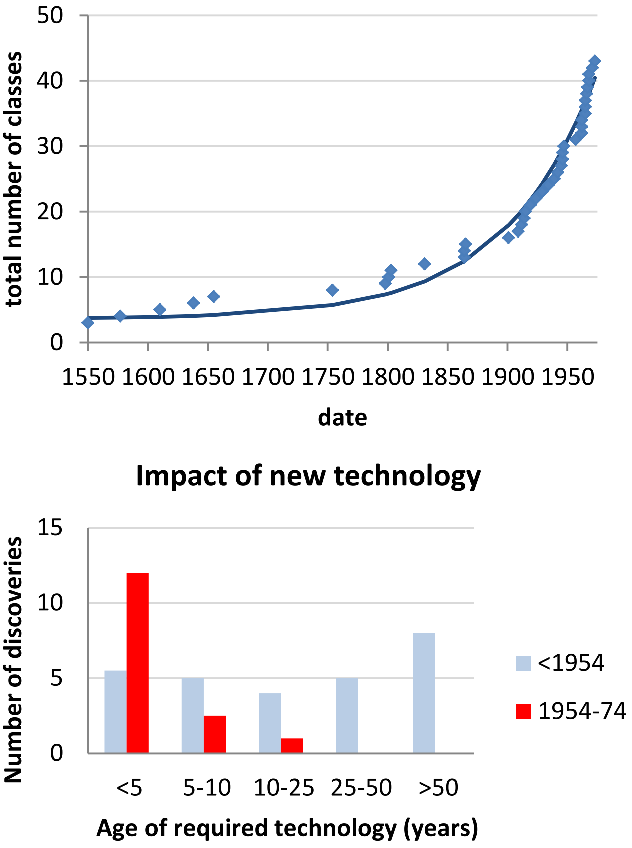

Figure 6.1: astronomical discoveries. Top, total number of recognised classes of astronomical object as a function of date; bottom, age of technology required to make the discovery, for the period before 1954 and for 1954-74. Redrawn from Harwit, Cosmic Discovery (1984).

Note that the pace of discoveries accelerates in the 20th century, and the time lag between the availability of the necessary technology and the actual discovery decreases. This is likely due to the greater number of working astronomers: if the technology to make the discovery exists, someone will make it. However, Harwit also makes the point that the discoveries made in the period 1954–74 were generally (~85%) made by people who were not initially trained in astronomy (mostly physicists, with a few engineers). This is because when radically new technologies are deployed in astronomy, they are often deployed first by experts in the technology, not experts in astronomy—for example, most of the radio astronomy community that emerged after WWII were physicists and electrical engineers from the wartime radar establishment, and many gamma-ray and cosmic-ray astronomers have a particle physics background. I suspect that if we were to update Harwit's plot, we might find that this trend has reversed: these days, facilities (e.g. spacecraft) are often constructed by engineers but used by astronomers. However, facilities that do not use electromagnetic radiation (neutrino telescopes and gravitational-wave observatories) are still built and run by non-astronomers.

6.2 Classification

Obviously, the discovery of unexplained phenomena stimulates both observational and theoretical work in the field: observers seek more examples of the phenomenon, while theoreticians seek to understand it. Historically, the usual sequence of events is that more examples are discovered before understanding is achieved. Therefore, the next stage in the process is **classifycation**: either grouping previously unconnected objects together as a class (for example, Carl Seyfert's recognition of Seyfert galaxies, or Lacaille's class of "nebulous clusters", i.e. open clusters), or dividing a previously monolithic category into subclasses, e.g. Hertzsprung's and Russell's distinction between "main sequence" stars and "red giant" stars, or Hubble's classification of galaxies by their appearance.

Classification is usually an important step towards understanding, since it enables observers and theorists to be more precise about what they are looking at; occasionally, however, a poorly chosen classification scheme may be positively unhelpful (early catalogues of "fuzzy objects which might be mistaken for comets" include unresolved open clusters, various types of gaseous nebulae within the Milky Way, and external galaxies: before any progress can be made here it is essential to grasp that these catalogues include a number of fundamentally different types of object).

Classification schemes are often updated as understanding progresses. For example, the simple class "star" (bright thing in sky which does not move) is increasingly subdivided as observations and understanding improve, whereas the several classes "radio galaxy", "Seyfert galaxy", "quasar", "BL Lac object", etc., are increasingly lumped together (as "active galaxy" or "AGN"). Occasionally, individual items may shift from one category to another, sometimes with important results. For example, reclassifying the Sun as a star immediately has the implications that the stars excluding the Sun must be very distant indeed (since they are so much fainter) and that there may be other solar systems (since the Sun has planets, if the Sun is a star, then it is plausible that other stars have planets). Similarly, Galileo's discovery of Jupiter's satellites introduced a category "satellite of planet" into which our Moon could be moved (from its Ptolemaic position as "planet"). This helps to validate the idea that the Moon behaves differently from true planets (in orbiting the Earth rather than the Sun).

Making discoveries that require the introduction of new classes is harder than discovering new members of existing classes, as discussed by Steven J. Dick in his book *Discovery and Classification in Astronomy* (Cambridge, 2013). For example, Uranus is actually just visible by the naked eye at opposition (its brightest apparent magnitude is 5.5) and had doubtless been seen many times before Herschel noticed that in his large (by 18th century standards) telescope it appeared to have a visible disc, and might therefore be a (comparatively) nearby solar system object rather than a star: he subsequently confirmed this by observing that it moved. Even then, his first reaction was to put it in the existing category "comets" (new comets are discovered all the time), rather than "planets" (no new planet, beyond the five known since antiquity, had ever been discovered). Only after its orbit had been determined as nearly circular was Uranus recognised as a new planet.

6.3 Case Study 1: transient phenomena

In Aristotelian philosophy, the heavens are perfect and unchanging: celestial objects either remain fixed (stars) or move in predictable cycles ("planets", here including the Sun and Moon). Objects which do not do this therefore cannot be celestial, but must belong to the upper air—hence the close similarity of the words "meteor" and "meteorology" (both deriving from the Greek *μετέωρα* "raised, lofty"; the OED says "Atmospheric phenomena were formerly often classed as *aerial* or *airy meteors* (winds), *aqueous* or *watery meteors* (rain, snow, hail, dew, etc.), *luminous meteors* (the aurora, rainbow, sun halo, etc.: see sense A. 2b), and *igneous* or *fiery meteors* (lightning, shooting stars, etc.)"). This has the interesting consequence that European and Arabic records of transient phenomena in antiquity are very much inferior to the Far Eastern (Chinese, Japanese and Korean) records: because these phenomena were seen as "weather", European astronomers did not pay them close attention.

The reclassification of transient phenomena from "atmospheric" to "astronomical" dates from Tycho Brahe, who showed in the late 16th century that both the supernova of 1572 and the bright comet of 1577 had no detectable diurnal parallax, and were therefore located well beyond the Moon. In 1596 David Fabricius identified Mira ("the wonderful", named by Hevelius) as a presumed "new star" or nova, but saw it reappear in the same place in 1609; it was identified as a long-period variable by Holwarda and Hevelius in 1638−9. (It is possible that an earlier maximum of Mira was the "new star" that inspired Hipparchos to create his star catalogue.) Algol was identified as variable in 1667 by Montanari, but had probably been seen to vary earlier, as its Arabic name translates as "the Ghoul" or "the Demon". These signs of change and impermanence in the heavens helped to reinforce the transition from Aristotelian ideas of the heavens as fundamentally different from the Earth to the Newtonian picture of universal physical laws.

The advent of photography greatly simplified the identification of variable and transient phenomena: comparison of photographic plates is much more straightforward and unambiguous than comparison of individual observers' drawings or notes. In the modern era, computerized automatic comparison of CCD images is used to make nearly real-time searches for variable behaviour, as in the gravitational lensing experiments such as MACHO and OGLE. The list of recognised types of variable and transient phenomena is now much too long to cite in its entirety: some examples include

- Solar system phenomena: comets, meteors, volcanic activity (e.g. on Io), break-up and impact events (e.g. Shoemaker-Levy 9), solar activity (spots, flares, etc.), radio storms on Jupiter, etc.;
- Eclipsing binary stars (and transiting extrasolar planets);
- Intrinsically variable stars, e.g. Cepheids, RR Lyrae stars, W Virginis stars, Luminous Blue Variables, semi-regular variables, Long Period Variables, ZZ Ceti stars;
- Stars with active atmospheres, e.g. T Tauri stars, Wolf-Rayet stars, FU Orionis variables, and stars with episodic mass loss, e.g. η Carinae;
- Cataclysmic variables (transient events in mass-transfer close binaries), e.g. classical novae, dwarf novae, nova-like variables, polars and intermediate polars, supersoft X-ray sources;
- Stellar explosions or collisions, e.g. core-collapse supernovae, Type Ia supernovae, gamma-ray bursts (long-soft bursts probably associated with supernovae, short-hard with coalescence of close compact binaries);
- Gravitational microlensing events;
- Variability associated with active galactic nuclei, e.g. BL Lac objects, OVV quasars.

6.4 Case Study 2: nebulae

The first catalogue of “fuzzy blobs” was produced in 1781 (partial lists appeared in 1774 and 1780) by **Charles Messier** (1730–1817), and contained just over 100 entries. Messier’s initial intent was to catalogue these rather comet-like objects to prevent confusion among comet hunters like himself; he does not appear to have done anything with his objects other than find, describe and catalogue them.

Messier’s catalogue contains a wide variety of objects: open clusters (e.g. Praesepe, M44—which doesn’t look in the least like a comet), globular clusters (e.g. M3), a supernova remnant (M1, the Crab Nebula), several planetary nebulae (e.g. M27, the Dumbbell Nebula), some diffuse nebulae (e.g. M20, the Trifid), numerous spiral and elliptical galaxies (e.g. M31, Andromeda, and M87, the central cD galaxy of the Virgo cluster) and a few random oddities (a couple of asterisms, M40 and M73, and M24, which appears to label a piece of the Milky Way in Sagittarius). The same is true of Herschel’s much more extensive catalogues, published in the *Philosophical Transactions* of 1786, 1789 and 1802, which later under the editorship of **John Dreyer** evolved into *the New General Catalogue of Nebulae and Clusters of Stars* (*Memoirs of the RAS*, 1888, and still very much in use).

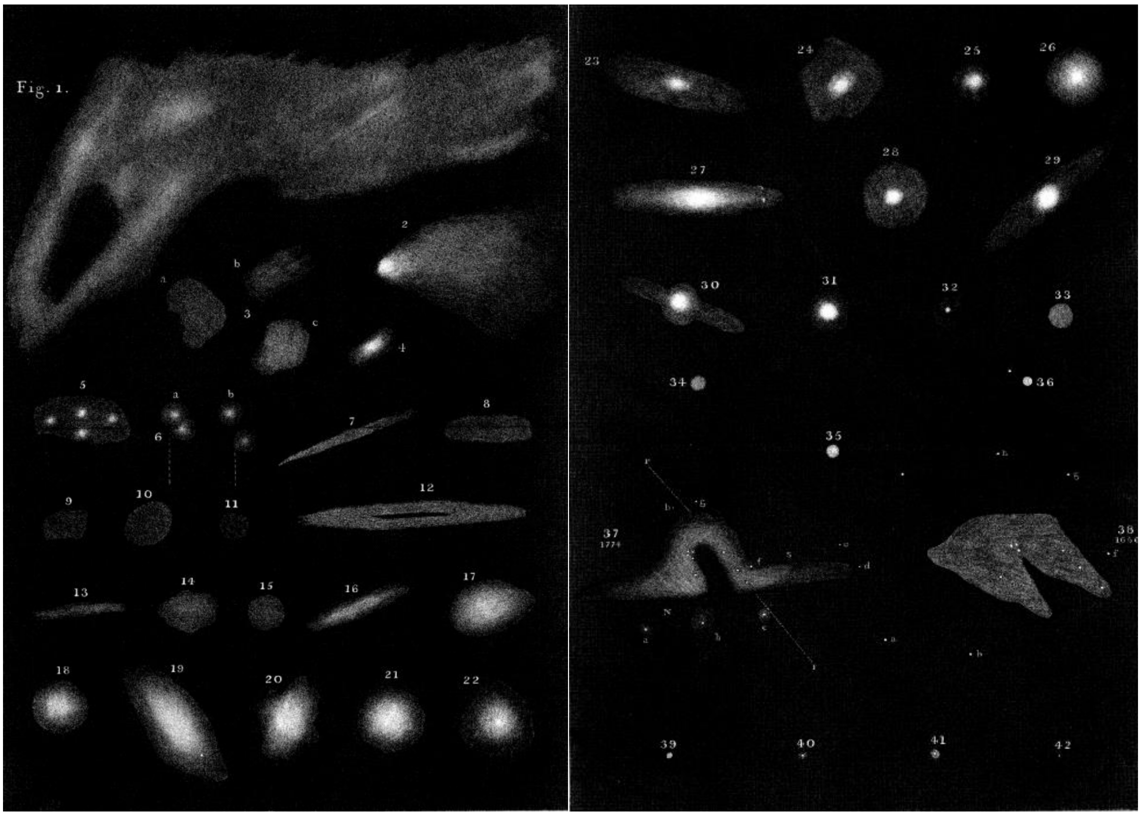

Figure 6.2: William Herschel’s drawings of nebulae, from Phil. Trans. ***101*** *(1811) 269–336. To modern eyes, some of these, e.g. 12, 19, 27, are clearly galaxies, and some, e.g. 1, 37, 38, clearly gaseous—but without spectroscopy, and with no preconceptions, it is not obvious that this gallery contains very different types of object.*

The trouble with listing disparate objects in the same catalogue is that it predisposes observers to consider them as related. This definitely happened with nebulae. Herschel had much the largest telescope then in existence, and believed he could resolve some of the larger, and hence presumably closer, “nebulae” in his list into stars (with hindsight, he was probably seeing larger entities such as H II regions); initially he reasonably (and correctly) regarded these as “no less than whole sidereal systems”, but in a later paper (*Phil. Trans*. (1811) 269-336) he says that “a longer experience and better acquaintance with the nature of nebulae, will not allow a general admission of such a principle" (i.e. "that nebulae properly speaking were clusters of stars"). In 1902, Agnes Mary Clerke goes so far as to say that "the conception of the nebulae as remote galaxies, which Lord Rosse's resolution of many into stellar points had appeared to support, began to withdraw into the region of discarded and half-forgotten speculations" (*A Popular History of Astronomy during the 19th Century*, 4th edition, p422). Among her arguments for this position is the observation (from spectroscopy) that "a considerable proportion of these perplexing objects are gaseous": that is, she is reasoning that because *some* nebulae are largely gaseous (established spectroscopically by **William Huggins** in 1864) and thus probably smallish objects within the Milky Way, it must follow that *all* nebulae are smallish objects within the Milky Way. Admittedly, this is oversimplifying her arguments a bit: she also makes the point that some of these gaseous nebulae are associated with stars—which is definitely true, cf. the Pleiades and the Orion Nebula—and therefore argues that there is a continuum of essentially similar objects stretching from entirely gaseous through gas-plus-stars to entirely stellar. Even with this more subtle logic, it is still clear with hindsight that she and her contemporaries were unduly influenced by the expectation that all the objects in the catalogue were essentially similar (she even comments on how small the stars in the resolved nebulae are, but does not therefore infer their very great distance).

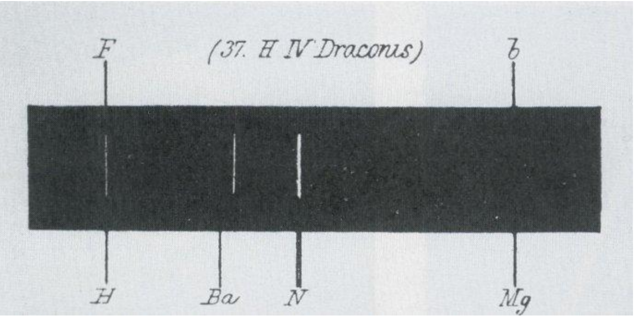

Figure 6.3: William Huggins' spectrum of the planetary nebula NGC 6543 (the Cat's Eye), showing emission lines and hence identifying the object as gaseous. The brightest lines, which Huggins initially identified as nitrogen, are in fact [O III] at 4959 and 5007 Å (the nitrogen line is at 4964 Å, so the misidentification is completely understandable). From Huggins, Phil. Trans. ***154*** *(1864) 437–444.*

This failure to recognise the essential *dissimilarity* of different types of catalogued "nebulae" persisted until the 1920s, when **Hubble** resolved Cepheids in M31 and deduced a distance of about 1 Mpc. This is a technology-driven advance: Hubble was using the then new 100-inch Hooker telescope on Mount Wilson, and could therefore resolve fainter objects.

In fairness, one should point out that the unfortunate occurrence of a supernova, S Andromedae, in the Andromeda Nebula in 1885 added to the confusion. Supernovae were not recognised as a separate class of object at the time (this was only realised by Baade and Zwicky in the 1930s), so S And was interpreted as an ordinary nova. As such, it was far too bright to allow M31 to be placed outside the Milky Way—it was generally concluded that S And was produced by an ordinary star flaring up as it entered the nebula. Hence, S And, which should have demonstrated the *large* distance of M31 (if it had been properly identified as a supernova) was actually thought to demonstrate its comparatively *small* distance. Note that this is *also* a classification problem: if there had been a supernova in our Galaxy during the 19th century, the astronomers of the day would have had the classification "supernova" available to them, and therefore would not have misidentified S And. But in fact the most recently observed local supernova was Kepler's, in 1604, just before the invention of the astronomical telescope, so nobody in the 19th century realised that Kepler's object was not simply a nearby case of the "novae" that they observed on a fairly regular basis. (By 1920, several genuine novae had been observed in M31 and other nearby spirals, using the new large reflecting telescopes: astronomers had therefore to make a decision about which they believed to be "typical" novae—S And or the faint photographic objects. As discussed later, the outcome was a classic example of "theory-laden data":

astronomers disposed to believe that spiral nebulae were galaxies assumed that the photographic nebulae were typical, and S And was an anomaly; those inclined to the opinion that spiral nebulae were small objects within the Milky Way assumed the opposite.)

6.5 Case Study 3: stellar classification

Through most of astronomy, the stars were simply the backdrop against which one measured the motions of the planets. Hipparchos in his star catalogue classified stars into six ranks according to their apparent brightness (thus, since the brightest stars were of the first rank, saddling us with the minus sign in the magnitude scale). Owing to the physiology of the human eye, colour is only perceived above a certain level of illumination, and therefore almost all stars are simply white dots to the naked eye. Ptolemy lists six stars as *hypokirros* ("somewhat yellow"), namely Aldebaran, Antares, Arcturus, Betelgeuse, Pollux and Sirius: these are all bright stars, with spectral classes K5 III (orange giant), M1.5 Iab (red supergiant), K1.5 III (orange giant), M2 Iab (red supergiant), K0 IIIb (orange giant) and A1 V (white main-sequence) respectively. The first five are genuinely reddish in colour and bright enough to excite the retinal cones. The inclusion of Sirius in the list has confused astronomers ever since, and has led to various fanciful theories involving the ejection of gas shells and so forth, but is almost certainly the result of the fact that from Alexandria Sirius is observed at fairly low altitude and can appear reddish due to atmospheric effects; Chinese astronomers at the same period record it as white.

Little further progress was made until the 19th century. Herschel's observations of double stars showed that stars have different *intrinsic brightnesses*, and Herschel also noted that stars had different *colours*. In 1798, Herschel even observed "by a prism applied to the eye-glasses of my reflectors" the spectra of a few bright stars. Some 19th-century astronomers, notably **Schjellurup** (1874) and **Espin** (1886) followed Ptolemy in making lists of red stars, but the real push toward classification came from the increasing data on stellar spectra.

Fraunhofer actually observed the spectra of some bright stars as well as the Sun, and developed the technique of using a cylindrical lens to broaden the spectrum for easier identification of lines, but (as a practising optical engineer rather than a physicist) did not engage in much study of the spectra he observed. However, after Kirchhoff and Bunsen's work in the 1850s clarified the chemical basis of spectroscopy, several astronomers began systematic studies of stellar spectra. **Father Angelo Secchi SJ** (1818–1878) was director of the Vatican Observatory from 1850 to his death. He made important contributions to solar astrophysics, but is primarily remembered for the first serious attempt at spectral classification. He observed over 4000 stellar spectra and by 1868 had classified them into four types:

- Type I (Sirian): white stars whose spectra contain four dark lines from hydrogen;
- Type II (solar): yellow stars whose spectra are distinguished by numerous fine lines;
- Type III (Antarean): orange and red stars whose spectra are divided into numerous bands which darken towards the violet end of the spectrum;
- Type IV ("small red stars"): resembles Type III, but the bands darken towards the red end.

The first three are clearly a somewhat coarser version of our (Harvard) system: Type I corresponds to classes B and A, Type II to F–K, and Type III to class M, with molecular bands. Type IV Father Secchi correctly identified as containing carbon compounds; they are now appropriately labelled C. A few stars, e.g. γ Cas, were also observed to display emission lines: Kaler in *Stars and their Spectra* refers to these as "Type V", but it is not clear that contemporaries did so

(Agnes Clerke describes only four types). In 1874, **Hermann Vogel** produced a very similar classification system, specifying the fine lines of Type II as due to metals and subdividing each class into two or three subclasses (for example, his Ia and Ib are our A and B respectively; his Ic has hydrogen in emission), but his version did not become as widely used as Secchi's. He defined only three main classes: he specifically—and quite wrongly—concluded that there was no real difference between Secchi's third and fourth types ("it seems to me advisable not to allow Secchi's 4th type to remain as an independent class, despite its being clearly distinguishable by eye from the 3rd type"—HC Vogel, *Astronomische Nachrichten* **84** (1874) 113–114, translated from the original German).

William Huggins also observed stellar spectra in the 1860s, but did not define a classification scheme. He did, however, correctly correlate the spectral types with temperature: Agnes Clerke records him as having attributed (in 1891) the lack of metallic lines in the spectra of white stars to "the high temperature of the vapours producing them" and stating that "a considerable cooling of the Sun would probably give rise to banded spectra due to compounds."

The modern classification scheme was developed at Harvard College Observatory under **Edward Pickering**, and is explicitly an extension of Secchi's classification. The original system is described in the Introduction to the Henry Draper catalogue of 1890 as follows:

Secchi	Harvard	Definition
I	A	hydrogen lines only
	B	hydrogen lines plus other lines (usually at 4026 and 4471 Å)
	C	as A but with doubled lines ("not certain if this represents a real condition")
	D	as A but with emission lines present
II	E	Fraunhofer F, H and K lines present, no others
	F	as E, but other hydrogen lines present
	G	as F, but with additional (non-hydrogen) lines
	H	as F, but with drop of intensity below 4310 Å
	I	as H, but with additional dark lines
	K	as H or I, but with dark bands present
	L	different varieties of K-like spectra
III	M	as Secchi
IV	N	as Secchi
(V)	O	emission lines (as Secchi)

In addition, planetary nebulae were designated class P and a class Q was defined for peculiar spectra which did not fit any of the defined classes.

The actual classification of spectra was carried out by a team of women ("Pickering's harem"), of whom the most important in classification were **Williamina Fleming**, **Antonia Maury** and **Annie Jump Cannon**. The original scheme as shown in the above table was due to Pickering and Fleming (1890): the alphabetical ordering corresponds approximately to decreasing order of strength of hydrogen lines.

Maury (1897) recognised several of the problems associated with this scheme, and devised a new system of 22 classes denoted by Roman numerals I-XXII. In her system, classes I-XX were ordered to produce a smooth variation in line strength for *all* lines, not just hydrogen: this corresponds to a temperature sequence, though Maury does not mention this and may not have known it. Although she placed class XXII (emission-line stars) at the end, she explicitly suggested that it might instead belong at the beginning; she also recognised that her type XXI (Secchi's type IV, Fleming's N) did not form part of the continuous sequence (as indeed it does not: the difference between class N, now called C, and classes GKM is one of composition, not temperature).

Maury also subclassified the main groups I-XX into three divisions *a*, *b* and *c* according to the *appearance* of the lines: division *a* stars had wide and well-defined lines, *b* had wide but hazy lines, and *c* had narrow, well-defined lines. This was a physically meaningful distinction (in fact, class *c* is supergiant stars—compare the MKK luminosity classes discussed below—and class *b* is stars in rapid rotation, with Doppler broadened lines), and influenced Hertzsprung in his later discovery of the HR diagram. (Hertzsprung was extremely cross when the *abc* classification was dropped in later versions of the system: "one hopes therefore," he says in 1909, "that the spectral analysts of the future will not turn away from the investigation of the *c*-properties, which seem to have great physical significance."[1])

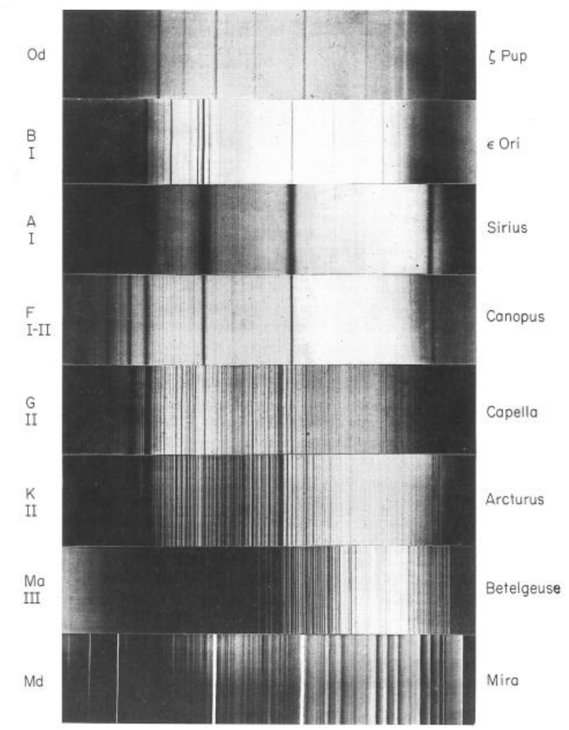

Figure 6.4: Harvard spectral classification sequence as adopted by Annie Cannon. The current classifications of these stars are: ζ Pup, O4 I(n)fp; ε Ori B0 Ia; Sirius A1 V; Canopus A9 II; Capella G1 III (+K0 III; Capella is a binary); Arcturus K0 III; Betelgeuse M1-M2 Ia-ab; Mira M5-9 IIIe (Mira is a long-period variable whose spectral type changes over its period).

*These are original Henry Draper catalogue spectra, taken from James Kaler, Stars and their Spectra (Cambridge 1989) figure 3.2. Kaler credits Annals of the Harvard College Observatory **23** (1901), but that seems to be wrong, as volume 23 dates from 1890. It appears to be a version of plate I of volume 28, but the published plate does not include ζ Pup or Mira, and labels all the stars by their Bayer designations not their names (e.g. Betelgeuse is given as α Orionis). Maybe Kaler had access to the original manuscript rather than the published article.*

Perhaps unfortunately, Maury's system did not catch on—with its 22 finely divided classes, it was too cumbersome to use in classifying (by hand) many thousands of stars—and the classification which survived was a variant of Fleming's original alphabetic list devised by Annie Jump Cannon (1901). Cannon scrapped C (which Pickering already doubted) and D (emission lines are now indicated by a suffix e), absorbed E and H into F, I into G, and L into K, introduced suffix numbers to subdivide the classes, and reordered them, as Maury had, so that all lines appeared

---

[1] E Hertzsprung, *Astronomische Nachrichten* **179** (1909) 373–380, translated from original German.

to strengthen and weaken in a smooth sequence. This gave the modern OBAFGKM ordering. Later, class R was introduced ahead of class N: R and N are carbon-rich analogues of GKM, and were subsequently replaced by class C (for carbon, using a letter that had been dropped from the original sequence); and a class S was added for red stars showing ZrO bands instead of the TiO bands characteristic of class M (S stars are intermediate between the oxygen-rich M stars and the oxygen-poor C stars).

This system was essentially in place by 1918, though the extreme classes (M and its relatives at the cool end, O and the Wolf-Rayets at the hot end) did not reach their final form until the 1930s: Cannon had originally used the numerical subdivisions only for classes B to K. In 1918, the sequence was known to be a sequence of decreasing temperature (this can be inferred from the colour of the stars; Wilhelm Wien published his displacement law, $\lambda_{\max}T$ = const, in 1893), although the full quantitative explanation was not available until Saha derived his famous equation in 1920, thus founding the discipline of astrophysics. The principal later modification, by **Morgan, Keenan** and **Kellman** in 1943, was the recognition that changes in the widths of particular lines were sensitive to the density, and therefore indirectly to the luminosity, of the star (giant stars are much lower density than main-sequence stars). They devised a system of Roman numerals running from Ia (bright supergiant) to V (main-sequence star) which is called the **luminosity class** and appended to the Harvard spectral type: thus the Sun is type G2 V. A system of suffix letters, introduced in the 1920s, is used to indicate unusual features: thus γ Cas is now classed as B0 IVpe (extreme hot end of class B, subgiant, with peculiar spectral features and emission lines).

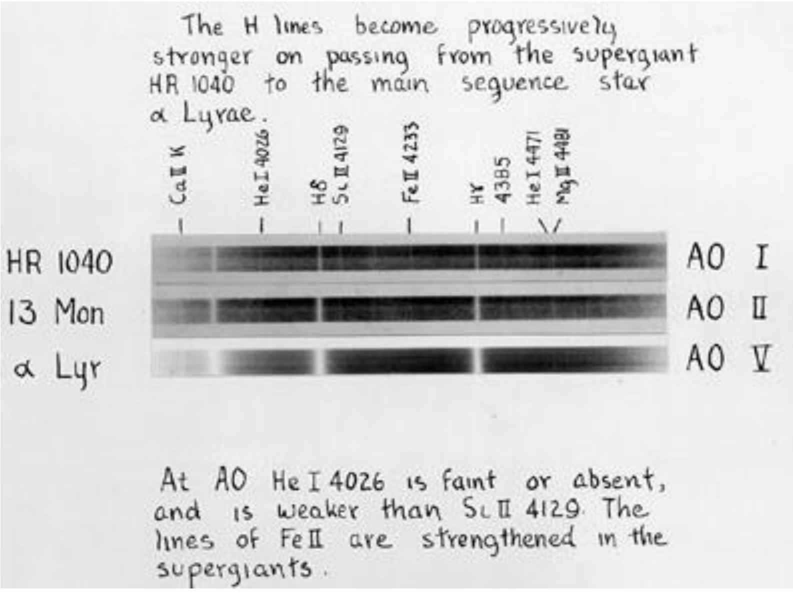

Figure 6.5: luminosity effects at class A0, from Morgan, Keenan and Kellman's 1943 Atlas of Stellar Spectra (available online at https://ned.ipac.caltech.edu/level5/ASS_Atlas/MK_contents.html).

Note that these are negatives, so absorption lines appear bright, unlike the positive prints in figure 6.4. The greater width in the hydrogen lines of main-sequence stars (Maury's class a) compared to supergiants (Maury's c) can also be seen comparing Sirius and Canopus in figure 6.4.

The stellar classification system was much more productive, scientifically speaking, than Herschel's earlier attempt to distinguish types of nebulae. This probably stems from two main facts:

- It is quite difficult to mistake a star for something else, or vice versa, so the objects catalogued in star catalogues *are* generally stars (not always: the globular cluster ω Centauri has a star name, so Bayer must have mistaken it for a star, and the active galaxy BL Lacertae has a variable-star name). In contrast, catalogues of nebulae always contained several distinct types of object.
- Differences in spectra are fundamentally related to differences in the physics or chemistry, whereas differences in appearance need not be: for example, quasars and classical double-lobed radio galaxies do not look very similar, but are probably the same objects viewed from

> different angles. If William Huggins had used his spectroscopic observations of nebulae to subdivide Herschel's catalogue clearly into pure gas, gas-plus-stars and purely stellar spectra, the recognition that these categories are very different objects might have come more quickly.

One might suspect from this that, for example, Hubble's classification of galaxies (which is morphological) probably doesn't help greatly in understanding galaxies, whereas the classification of gamma-ray bursts into short-hard and long-soft, which is spectral ("soft" meaning "lower energy") probably does. The latter conclusion seems sound (long bursts are associated with supernovae whereas short bursts are not; long bursts only occur in galaxies with ongoing star formation, whereas short bursts can occur in elliptical galaxies—these and other properties suggest that they really are different types of object); the former is debatable (there do seem to be differences, e.g. in star formation rate, which correlate with some aspects of the classification; overall, the subdivisions of the S and SB classes seem to be meaningful, whereas the shape-based subdivision of E galaxies is much less so).

6.6 Discovery, classification and understanding

Discovery in astronomy can be a complex process, especially when it requires recognition of a new class of objects or overturns long-held beliefs. As noted above, both Uranus and Neptune were observed multiple times before they were "discovered", but this does not mean that Herschel and Galle do not deserve credit for their discovery: the previous observers did not recognise that they had seen a planet rather than a star. (This happens so often in astronomy that looking for pre-discovery images of newly recognised solar system objects is standard practice: it even has a name, *precovery*[2], as distinguished from "recovery", obtaining new images after a recognised object has not been observed for some time, e.g. because it was too close to the Sun.) This type of discovery, where the key issue is not the observation but the interpretation, can be quite protracted: Herschel initially assumed Uranus was a comet, and thus nothing very remarkable (his 1781 paper, *Phil. Trans.* **71**, 492−501, is simply entitled "Account of a comet"), and only in 1783 (*Phil. Trans.* **73**, 1−3) did he acknowledge that "By the observations of the most eminent Astronomers in Europe it appears, that the new star, which I had the honour of pointing out to them in March, 1781, is a Primary Planet of our Solar System." On this basis, it could be argued that the "discovery" of the *planet* Uranus took around two years to complete.

Obviously, not all discoveries involve the recognition of a new class: the discovery of a new asteroid or comet is still a discovery, even though it only adds one more member to an already well-populated category. Conversely, not all recognitions of new classes are discoveries: supernovae were established as a new class by Walter Baade and Fritz Zwicky (*PNAS* **20** (1934) 254−259), but nobody would claim that Baade and Zwicky "discovered" supernovae.

Establishing that a discovery has actually taken place may be challenging. The theory of non-rotating black holes was worked out, as a solution to the field equations of general relativity, by **Karl Schwarzschild** in 1916, and the theory of rotating black holes by **Roy Kerr** in 1963, but black holes were not "discovered" until 1972 at the earliest, when the mass of the compact object in the X-ray binary Cygnus X-1 was estimated to be much higher than the maximum possible mass of a neutron star. This is a heavily theory-laden "discovery": it relies on the calculated maximum possible mass of a neutron star, which in turn relies on two theoretical constructions: the equation of state of neutron star matter (which determines the mass-radius relationship for

---

[2] https://en.wikipedia.org/wiki/Precovery

neutron stars) and general relativity—specifically the requirement that sound cannot travel faster than light. However, as pointed out by Jan van Paradijs[3], "The fact that compact objects with dynamical mass estimates exceeding $\sim 3\ M_\odot$ cannot be neutron stars is not equivalent to their being black holes, as defined by the particular space-time structure described by Schwarzschild and Kerr metrics", and it is in fact very difficult to demonstrate conclusively that a massive compact object actually is a black hole in this sense. The best that we have currently been able to do is to demonstrate that some massive compact objects behave in ways that are consistent with their having an event horizon rather than a solid surface: for example, the required accretion rate in the nearby active galaxy M87 would heat up any solid surface to the point where it would produce significant thermal emission, which is not observed[4], and X-ray binary systems with companions whose masses indicate that they should be black holes behave differently from those whose companions are neutron stars[5]. It would be fair to conclude that there is strong indirect evidence that the compact objects that "should" be black holes according to general relativity do indeed behave, where this can be checked, as though they have an event horizon, and there are no cases in which such objects have been definitively shown *not* to have an event horizon. Does this constitute "discovery"? If so, who are the "discoverers"—those who determined the masses of the relevant compact objects, or those who demonstrated that their properties indicated the presence of an event horizon?

In this case, understanding has preceded discovery, and indeed the understanding will be vital in establishing that the discovery has taken place: for example, a key area of investigation is the gravitational wave signals from coalescing black hole binaries, which probe the structure of spacetime around these systems. However, this is not typical of astronomy, where discovery often precedes understanding, sometimes by many years: gamma-ray bursts, discovered in the 1970s, are still not completely understood; novae and supernovae have been recorded, by the Chinese at least, since the second millennium BC (the oldest surviving record is said by Needham, in *Science and Civilisation in China*, to date from 1300 BC), but were only understood towards the end of the second millennium AD.

Classification—the recognition of categories—is a halfway house between discovery and understanding. As shown in the case studies above, it is entirely possible to develop a classification system before the objects being classified have been understood—this has happened often in astronomy, and is one of the root causes of the notoriously awful astronomical terminology (the other being a tendency to underestimate the number of examples of the class that will be discovered—see, for example, variable-star nomenclature[6]). Despite the counterexample of case study 2, classification is generally helpful in establishing understanding, because the definition of clear categories aids communication—different people can be sure that they are discussing the same type of object—and may make relationships easier to spot (the categories of the original Harvard classification were in the wrong order, but the fact that categories more finely divided than Secchi's *existed* doubtless helped Maury and Cannon to identify the correct ordering).

Although astronomers are highly conservative in their terminology, classification schemes do evolve with understanding. For example, Angelo Secchi's decision to separate the "banded" spectra into type III and type IV was vindicated by Maury and Cannon, both of whom recognised that Secchi's type IV (Maury's type XXI, Fleming and Cannon's class N) did not belong to the

---

[3] arXiv:astro-ph/9802177; contribution to *The Many Faces of Neutron Stars*, ed. R Buccheri, J van Paradijs, MA Alper (Springer, 1998)

[4] AE Broderick et al., *ApJ* **805** (2015) 179.

[5] See, for example, R Narayan and JE McClintock, *New Astr. Rev.* **51** (2008) 733–751.

[6] https://en.wikipedia.org/wiki/Variable_star_designation. Read it and weep.

smooth sequence of Cannon's classes O–M. Secchi had already identified the principal spectral features of these stars as due to carbon, and following the development of quantitative spectroscopy in the 1920s it was realised that these were stars with unusually carbon-rich atmospheres. In 1921, class R was introduced to extend the class N sequence to higher temperatures—N is roughly comparable to M in terms of surface temperature, and R to late G or K—and this was later replaced by a single class C (for carbon) with numerical subdivisions. The current C class has a different set of subdivisions, two of them, C-R and C-N, ironically returning to the old R and N classes. These changes have all come about as consequences of better understanding of the carbon stars. Likewise, class S, introduced in 1922 for stars with molecular bands of ZrO (instead of TiO for standard M stars) and now understood to represent stars with C:O ratios between those of the carbon stars and the standard sequence, has undergone several revisions of its subdivisions in an attempt to describe the complex interplay of surface temperature and chemical composition that determines the spectral features of S stars.

It is easy to dismiss classification as mere book-keeping, but this is unfair. Systematic scientific progress, as opposed to progress by means of unforeseen discoveries, often proceeds by investigating anomalies, such as the irregularities in Uranus' orbit that led to the discovery of Neptune. Good classification systems help us to recognise anomalies—it is easier to see differences when you have a standard to compare to—and can therefore facilitate progress. It is true that occasionally *bad* classification systems actively impede progress, but this is not the norm, and arguably just means that we need to invest enough time and thought into classification to minimise the risk of adopting unhelpful schemes.