Theme 2: Astronomy as a Science

2.1 Astronomy and Astrology

For most of their history, up to the 17th century, astronomy and astrology were completely intertwined: for example, astronomers devoted much time and effort to compiling tables of planetary positions, not only for navigation and time-keeping but also to aid in the casting of horoscopes. Nowadays, almost everyone would agree that astronomy is a science but astrology is not. Why is this? What do we mean by the statement "astronomy is a science" anyway? Are there any parts of astronomy which are not science (and, conversely, any parts of astrology which are science)?

This is the subject matter of philosophers of science, who aim to provide (1) a descriptive account of science (what scientists do), (2) a normative account (what scientists ought to do) and (3) a definition (what distinguishes science from other pursuits). It has to be said that most working scientists do not take philosophers of science seriously—and, indeed, some philosophers of science do rather invite this reaction because of the lack of any apparent connection between their descriptions and scientists' experiences. However, the serious issues around the definition of "science" which have arisen in recent years, particularly in the USA, strongly indicate that working scientists should try to develop a clear mental picture of the definition and practice of science, in order to be able to defend it against incursions by pseudoscience. [As far as I know, nobody has yet tried to demand "balanced treatment" for astrology in astronomy courses—but you never know!]

2.2 The Scientific Method

If pushed, most scientists would say that science is a process (a way of working and thinking) and not simply a body of knowledge—though the acquisition of a pre-existing body of knowledge is usually necessary in order to participate in the process of science. This is not well reflected in undergraduate courses, which tend to focus exclusively on delivering the said pre-existing body of knowledge as efficiently as possible. Most scientists only learn to *do* science in postgraduate courses or science-based employment.

The process of doing science is commonly referred to as the scientific method. The average scientist would probably characterise the scientific method approximately as in figure 1.

An example of this might be General Relativity:

1. the advance of the perihelion of Mercury is not explained by Newtonian gravity (so, a piece of unexplained data);
2. the idea that all observers should agree on the laws of physics (a principle);
3. the idea that an observer freely falling in a gravitational field is equivalent to an inertial observer, while an inertial observer in a gravitational field is equivalent to an accelerated observer in the absence of gravity (a thought experiment);
4. generalise special relativity to the case with accelerated motion and gravitational fields (a new/improved theory);

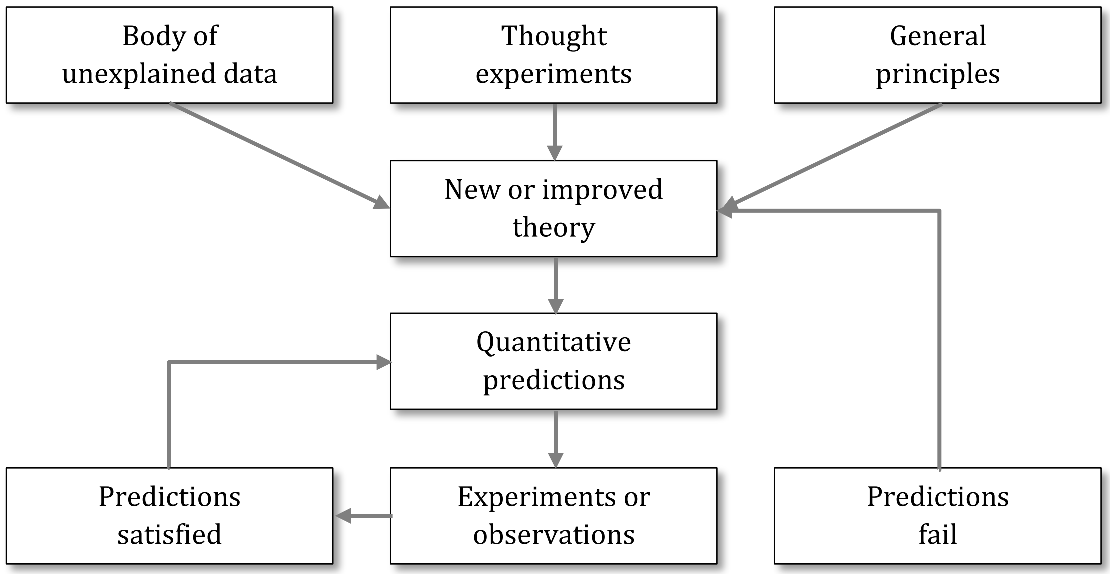

Figure 2.1: a naïve outline of the Scientific Method. Hypotheses are framed as a result of unexplained data, thought experiments or general principles and used to make quantitative predictions. The predictions are tested: if they work, the theory is used to make more predictions; if they fail, the theory is modified or discarded.

5. quantitative prediction of expected advance of Mercury's perihelion agrees with the observed (and previously unexplained) value (a first successful test);
6. quantitative prediction of gravitational bending of starlight agrees with Eddington et al's [admittedly not very accurate] measurements during the solar eclipse of 1919 (a second successful test).

One then continues to use the successful theory to make predictions until one encounters a situation where the predictions fail. At this point one attempts to look for a new or improved theory which will produce new predictions. Sometimes this theory is really just a modification of the initial conditions used in making the original prediction, as in the discovery of Neptune (Newton's theory of gravity fails to account for the orbital paths of Uranus and Saturn, but the same theory with the assumption of another gravitating object succeeds in accounting for the data and predicts where the assumed object is to be found). Sometimes, however, as with General Relativity, it is a completely new construction built on new physical principles.

The key features of this picture of science are:

- it is data-driven:
- theories must make successful predictions before they are accepted;
- theories which make unsuccessful predictions must be modified or replaced;
- it implicitly assumes a well-behaved underlying reality:
- general laws can be experimentally supported by looking at particular cases (inductivism);
- theories which have made many successful predictions are regarded as more secure than new theories which have made only a few successful predictions;
- it is assumed that the scientific community can and will agree on the experimental results and their significance (i.e. the theory's predictions unambiguously pass or fail experimental tests);
- the process of science consists of systematically testing the predictions of current theories.

According to this picture of science, a discipline which claims to be a science can be assessed based on these key features:

- Does it have theories which make quantitative, testable predictions?
- Are those predictions duly tested, and are the results taken seriously by practitioners (i.e. if they fail, is the theory modified or discarded)?

This seems like an entirely reasonable approach, and it can be successfully applied to many pseudosciences. For example, astrology makes many testable predictions, so it passes the first test. When carefully tested under controlled conditions, the predictions are repeatedly found to fail. However, practitioners of astrology continue to use the same theories despite the fact that their predictions are not successful. Therefore, astrology is not scientific, because it fails the second test.

The trouble with this picture is that, by this logic, many branches of physics are not (or have not been) science either! For example:

- The kinetic theory of gases fails the quantitative test of understanding the specific heats of diatomic molecules. This was known from the earliest days of kinetic theory: Maxwell wrote in 1860 that "a system of such particles could not possibly satisfy the known relation between the two specific heats of all gases". Despite this, Maxwell and others continued to develop kinetic theory with great success, whereas according to our recipe above they should have abandoned it (which would have been a disaster for physics). We now know that the specific heat problem is a quantum effect—but there is no realistic way that an abandonment of kinetic theory in the 1860s could have led to the development of quantum mechanics 50 years ahead of schedule!
- Many early formulations of theories which subsequently turn out to be highly successful are obviously incapable of describing the real world in their initial state. Examples include quantum field theory (which gave infinite answers to perfectly sensible calculations), the Bohr atom (which wasn't stable), the original "Hot Big Bang" paper by Alpher, Bethe and Gamow (which proposed a model for making the chemical elements which wasn't physically plausible and was found to disagree with observation) and cosmological inflation (which, in Guth's initial formulation, lacked a "graceful exit" and therefore could not produce an universe like ours).

According to the above recipe, all of these should have been strangled at birth. Clearly they weren't, and clearly this decision was correct. [You could argue that all of these were modified to remove their inconsistencies, which does satisfy the methodology; but the point is that they were taken seriously immediately despite their evident flaws.]

There are also philosophical issues with this model, particularly with the "key features" listed above.

- The principle of induction—generalising from a few particular cases to a universal law—is clearly not logically defensible, unlike deductive logic (for example, the deductive argument, "All persons named Susan are female; my name is Susan; therefore I am female," is clearly logically sound, whereas the inductive argument, "My name is Susan; I am English; therefore all persons named Susan are English," clearly is *not* logically sound). This worries many philosophers of science, especially when combined with the argument that theories which have made many successful predictions are more secure than those which have not (if there are an infinite number of possible cases, the philosophers argue, having verified 100 of them is not significantly more convincing than having verified 1).

- Many philosophers also worry about the assumed reliability of experimental data: they point out that it is very difficult to construct an experiment which is a pure test of the hypothesis in question (there are bound to be other questions, such as the calibration of the apparatus, the assessment of background levels, the reliability of approximations used in the calculation, etc.), and it is not obvious how the scientist decides whether an anomalous result is an indicator of problems with the theory or a consequence of one of these subsidiary issues. Looking at the history of astronomy, there is no doubt that this is a genuine problem. For example, in 1921 van Maanen published a paper indicating that the "spiral nebula" M51 had rotated significantly over a period of 10 years. This result was used by Shapley as evidence that the spiral nebulae were objects within our own Galaxy rather than (as we now know them to be) external galaxies—and indeed, if the result had been correct it would have been very strong evidence for that position. However, within a decade Hubble had shown, using Cepheids, that M51 and its fellow spirals *are* external galaxies. Although van Maanen was an acknowledged expert in the field, his results were simply wrong, for reasons which have never been conclusively explained but probably have something to do with observer prejudice (after Hubble obtained his results, several people independently remeasured van Maanen's plates and found no evidence for rotation).
- The last "key feature" is not, in fact, a realistic picture of how scientists work. Much scientific work does not consist of attempts to test theoretical predictions: for example, experiments designed to measure the numerical value of a parameter of the relevant theory (e.g. the value of Hubble's constant or the mass of the top quark) are not primarily aimed at testing a prediction (the parameter in question is an input to the theory, not a prediction from it). Indeed, a great deal of scientific work could be regarded as assuming the correctness of the relevant theory and investigating its consequences. Is this work not "good science"?

Another issue which concerns philosophers of science (but rarely worries working scientists) is the status of theory. Early philosophers of science (particularly the "logical positivists") felt strongly that the basic content of science was observational results (obtained "via the senses"), and that theories are not "real" but are patterns imposed on the data by scientists. Most scientists, in contrast, feel that theories reflect a basic structure or framework which is present in the "real world" even though it cannot be directly observed. In extreme cases, "anti-realist" philosophers deny the reality of entities (not merely theories) which are not directly observed—for example, electrons (this is the point at which most working scientists start shaking their heads and muttering).

2.3 Astronomy and Philosophy of Science

The history of astronomy is an interesting case study for philosophers of science because it includes many episodes which touch directly on issues of this kind. Indeed, one of the main schools of philosophy of science (Kuhn's theory of scientific revolutions) has its roots in a study of the "Copernican revolution" in the late 16th – early 17th centuries.

2.3.1 Popper and falsification

Karl Popper (1902-94) was one of the most influential 20th-century philosophers of science. His basic ideas can be summarised as follows:

- It is not possible to prove a scientific theory. It is, however, possible to *falsify* a scientific theory, by showing that it makes predictions which are not consistent with experiment. (This is essentially the difference between induction and deduction: you cannot use the facts that I am English and called Susan to prove that "all persons called Susan are English", but you can use them to *disprove* the assertion that "all persons called Susan are American".) Scientific theories are theories which can be falsified; theories which cannot be falsified are not scientific.

- Science consists of framing hypotheses and then attempting to falsify them. Confirmatory evidence is not important (you can never be "sure" that a theory is true). "Bold" hypotheses, which make many falsifiable predictions, are better than weak hypotheses which make only a few ("all planets orbit the Sun in ellipses" is a better hypothesis than "Mars orbits the Sun in an ellipse"). Falsified hypotheses should be rejected; not-yet-falsified hypotheses should be further tested.

This position is clearly closely related to the "naive" scientific method discussed above. The main difference is the lack of emphasis on confirmatory evidence: for Popper, all theories are only provisionally accepted, and the key role of experiments/observations is to falsify theories.

The Popperian view of science is generally liked by working scientists—which does suggest that it is getting something right. It is often invoked as a discriminator to distinguish science from pseudoscience, and works quite well in that role (for example, "God created the world in 4004 BC, but He created it with the appearance of being 13.7 billion years old" is clearly non-falsifiable, and therefore by Popper's standards not scientific; Popper's personal *bête noir*, Freudian psychology, is also given to non-falsifiable statements). However, it has the problems discussed above, namely that many successful episodes in the history of science rely on ignoring apparent falsifications, and that many scientists spend a great deal of time working on experiments that assume a theory rather than setting out to falsify it. (Some sciences also go through "data-gathering" phases which are not strongly reliant on an underlying theory: for example, the current searches for extrasolar planets really are not attempting to confirm or refute a particular theory of planet formation.) In addition, the lack of respect for confirmatory evidence is worrying: in general, confirmatory evidence is actually more highly valued than refutation, if anything (you get Nobel prizes for doing experiments that support new theories, not for those which refute them!), and well confirmed theories are more highly valued than weakly confirmed (but not falsified) theories. This can easily be seen by example: suppose you wish to design a bridge, and you need to calculate the load factors. Two programs are available to do this: one is heavily used and has been shown to work well, while the other is a new algorithm which has just emerged from beta-testing. Which one are you going to use?

2.3.2 Kuhn and scientific revolutions

Thomas Kuhn (1922-96), who trained as a physicist, was a historian of science as well as a philosopher of science, and therefore was well placed to recognise that real scientists do not appear to work in the way that Popper thinks they should. His description of the process of science can be summarised as follows:

- "Normal science" takes place in the context of a well-confirmed and trusted theoretical framework. Scientists agree on the theory and on an associated set of shared ideas (e.g. methodology, identification of important problems, etc.). Kuhn called this a **paradigm**. Within the paradigm, scientists work on well-defined problems, which may include the parameter-determining and data-gathering types of experiment as well as hypothesis-testing. The design of experiments and the interpretation of results presuppose the correctness of the paradigm.
- If results from "normal science" experiments are inconsistent with the paradigm and cannot be accommodated by "natural" extensions of the paradigm (e.g. postulating a

new planet to explain Uranus' orbit), they are **anomalies**. The initial response to anomalies is to ignore them in the hope that either they will go away or someone will come up with an explanation.
- Eventually, either because of a gradual build-up of individually minor anomalies or because a particular anomaly seems to be fundamentally inconsistent with the underlying framework of the paradigm, the field enters a **crisis**. The entire basis of the paradigm is now being questioned by the scientists, and so normal science, which relies on the essential correctness of the paradigm, cannot proceed.
- The crisis is resolved by the adoption of a new paradigm which resolves the anomalies—for example, the paradigm of classical electrodynamics threw up an anomaly in the results of the Michelson-Morley experiment (the Earth does not appear to be moving through the luminiferous aether); this is resolved in special relativity, in which the aether does not exist. This change of paradigm Kuhn called a **revolution**. Kuhn believed that paradigms were *incommensurable*—the change in the set of shared assumptions is so great that little communication is possible, and the standards applied to experimental results are not consistent. He also believed, in contrast to Popper, that rejection of a paradigm takes place only when there is both a crisis and a candidate new paradigm; if there is no new paradigm, the field will struggle on in extended crisis, using the existing paradigm despite its known problems, because there is nothing better.

Kuhn's approach does deal with many of the problems encountered by Popper's falsificationism. The concept of "normal science" is a better description of how scientists work than Popper's vision of everyone trying to refute hypotheses all the time, and the response of "normal science" to anomalies (be aware of them, but don't panic until they get too numerous or too serious) also accords with experience. The main problems with Kuhn's picture are

- He overstates the extent to which "normal science" unquestioningly accepts its paradigm. In reality, scientists do design experiments which test the fundamental basis of the underlying theory, while still working within the set of shared assumptions and practices that constitute the paradigm. They don't *expect* the experiments to produce a positive result, but they are open to the possibility. (See, for example, Hulse and Taylor's use of the binary pulsar to test General Relativity.)
- He is very determined that, at any given time, scientists in a single field have one paradigm. But this is very doubtful: for example, in the early 20th century astronomy clearly had two paradigms (the Big Galaxy of Shapley, in which the spiral nebulae were internal to the Galaxy and probably gaseous, and the Small Galaxy of Kapteyn, in which the spiral nebulae were external galaxies made up of unresolved stars). This is not Kuhn's expectation of what a crisis ought to look like: one should have a failing paradigm and no clear successor, not two competing paradigms.
- It is also the case that, historically, new paradigms have emerged following new discoveries in fields which do not seem, even in hindsight, to have been in crisis—for example, the discovery of the structure of DNA by Crick and Watson certainly ushered in a new paradigm in genetics, but there was no obvious crisis beforehand. People have argued that the problem here is a result of Kuhn's initial training in physics: his model is "tuned" to the physical sciences, and is less realistic when applied to other fields.
- The main problem, and the one which has resulted in the most undesirable consequences, is the insistence on incommensurability of paradigms. This makes it difficult to argue convincingly that replacement of one paradigm by another is an improvement—if you cannot compare results obtained under paradigm 1 with those obtained under paradigm 2, what can you say about their relative merits? This naïve interpretation of Kuhn therefore

has difficulty in modelling progress in science. Some later philosophers have indeed taken this to mean that paradigm replacement is not necessarily an improvement, and that progress in science is an illusion. Kuhn definitely did not believe this—Wikipedia quotes him (citing Freeman Dyson as the source) as saying "I am not a Kuhnian" in relation to this interpretation. But his description of the effects of paradigm change makes it hard to avoid this problem.

Kuhn has been enormously influential, as can be seen from the extent to which the phrase "paradigm shift" has entered common usage. His picture does have merit, in that the idea of "normal science" seems to be a better match to reality than Popper's view of good science as only attempts to falsify hypotheses. However, it seems that the overall idea of single incommensurable paradigms, succeeding each other by means of crisis and revolution, is oversimplified and misleading.

In contrast to Popper, Kuhn is not enormously helpful in distinguishing science from pseudoscience: the natural response of a faith healer to Richard Dawkins would seem to be "I am working in a different paradigm which is incommensurable with yours." [None of them *does* say that, but that just tells you that faith healers aren't very good at rational argument.] Again, this is not what Kuhn intended: his writings indicate that he did see science as uniquely effective at understanding the world, precisely because of the complex interplay between normal science (which makes steady but gradual progress by ignoring anomalies) and revolution (which makes discontinuous progress by replacing a failed paradigm with an improved one; in some of his writing he defines "improved" in terms of "better problem-solving ability").

2.3.3 Experiment and Observation

One of the main issues which divides philosophers of science (and distances them from working scientists) is the status of experimental or observational knowledge. Early philosophers of science, particularly the logical positivists, regarded experimental knowledge as "pure", unassailable, and fundamental in a way that the theories built on that knowledge were not. However, this view is not borne out by history: observations are fallible and revisable in the light of new knowledge. For example, the absence of observable stellar parallax was originally accepted as an observational fact, and used as evidence against a heliocentric cosmology; now we recognise that the actual stellar parallaxes were simply too small for naked-eye detection. Similarly, the Galilean "principle of relativity", which holds that velocities simply add (i.e. an object travelling at velocity $\mathbf{v}$ in a train travelling at velocity $\mathbf{u}$ is travelling at velocity $\mathbf{u} + \mathbf{v}$ as measured by an observer on the station platform) was repeatedly "verified" experimentally in the 300 years between Galileo and Einstein, but is now known to be incorrect. The experimental "verifications" reflect the fact that the true relativistic law for addition of velocities is experimentally indistinguishable from the Galilean law as long as all speeds are small compared to c.

The history of science shows that it is often necessary to interpret experimental results in the context of a specific theoretical framework. The framework allows the experimenter to design the experiment (for example, Michelson and Morley knew how precise their interferometer needed to be because they knew the size of the predicted aether drift) and identify possible systematic errors (Hertz was aware that some of his early measurements of the properties of radio waves were likely to be wrong because his laboratory was too small to neglect stray reflections; but if he had not known, from Maxwell's theory, how radio waves should behave he would not have recognised this).

Some modern philosophers of science, especially those better described as sociologists of science, leap from this dependence of interpretation on theory to the conclusion that experimental results as published are entirely dependent on the experimenter's theoretical framework, and therefore that changes is scientific theory are to be accounted for by changes in the social background rather than the effects of experiment. All working scientists will of course recognise that this view is not tenable, not least because experiments and observations have a nasty habit of *not* fulfilling the experimenter's preconceived expectations! To take one of our earlier examples, the Michelson- Morley experiment was carefully designed to measure that Earth's motion through the "luminiferous aether" as (thought to be) required by Maxwell's electrodynamics. Maxwell's equations had been experimentally confirmed to high precision, for example by Hertz's experiments, and so this motion was confidently expected, and the experimenters were primed to interpret their results appropriately. Nonetheless, the experiment stubbornly refused to conform to expectations, and Michelson and Morley were forced to conclude that the Earth did not seem to be moving relative to the aether. Subsequently, Einstein was able to reinterpret Maxwell's equations in a way that did not require the existence of an aether, and therefore explained Michelson and Morley's negative result; this reinterpretation was in fact the special theory of relativity (this explains why the special relativity paper of 1905 is actually called *On the Electrodynamics of Moving Bodies*).

Experimental results are therefore influenced by both the external world (with whose laws they must obviously be consistent) and the experimenter's theoretical framework (which will determine how the experiment is designed, which variables it is sensitive to and to what precision, and how both positive and negative results are interpreted). Although scientists instinctively feel that carefully-controlled and well-designed experiments are scientifically preferable to observations of naturally occurring phenomena, it can be argued that the latter are actually less subject to bias introduced by the prevailing theoretical framework: experiments are specifically designed not to be sensitive to effects other that the particular one being investigated, and it is therefore difficult to spot the truly unexpected. (The history of physics is littered with people who failed to make important discoveries which, in hindsight, are clearly visible in their experimental results!)

Observational campaigns, on the other hand, are often designed around what one can observe, rather than any more theoretical motivation, and any unexpected results are more likely to be noticed (for example, the first campaign of any instrument which opens up a new wavelength range is normally an all-sky survey; historically, these have almost always turned up unexpected, unpredicted and initially unexplained classes of sources). This does not mean that observations are entirely independent of the theoretical framework; the framework will still influence which instruments are built, which objects are singled out for observation, and how the results are interpreted. (A particularly acrimonious example of the latter is the debate about radio source counts in the early 1960s, with the supporters of the Steady State cosmology seeking reasons to discount an apparently contradictory observation.)

Philosophy of science is a large and complex field. This course is not the place to investigate its ideas further: the books by Chalmers and Godfrey-Smith are sensible starting points if you are interested.

2.4 Astronomy as a Science

Astronomy has provided philosophers of science with two key case studies: the Copernican revolution of the late 16th and early 17th centuries, and the replacement of Newtonian by Einsteinian mechanics in the early 20th century. These have been extensively used as test-beds for the ideas of various philosophical schools, with varying degrees of success. Philosophers of science therefore have no doubt that astronomy *is* a science—otherwise it would not make sense to use its history in this way. However, some aspects of astronomy, particularly the more modern variants of astrophysics and cosmology, do present problems.

The difference between astronomy and the "ideal" definition of a science—perhaps physics—is that astronomy is very rarely an *experimental* science. Most of our knowledge about astronomical objects is obtained indirectly, through techniques such as spectroscopy, analysis of binary star orbits, redshift surveys, etc. This raises the question of realism and anti-realism that we mentioned above: to what extent are these indirectly determined quantities "real"? Working astronomers of course are entirely confident that they are, but in some cases this confidence requires justification (see Exercises). Even if we accept them as real, they are clearly "theory-laden" —i.e. theoretical assumptions are required in order to convert the observations into physically relevant quantities. A current example of this is the interpretation of galactic rotation curves: everyone agrees on the data (curves of velocity against distance from the centre of the galaxy), but, while the majority of astronomers interpret these curves as demonstrating the presence of large quantities of dark matter, a minority prefer to see them as demonstrating a deviation from Newtonian gravity at small gravitational accelerations (Modified Newtonian Dynamics, or MOND). The two camps therefore deduce entirely *different* physical properties from the *same* empirical data. In physics, this sort of ambiguity can usually be resolved by designing an appropriate experiment; in astronomy, however, it is not in general possible to do this (the threshold gravitational acceleration relevant to MOND is far smaller than anything accessible in the solar system).

How do working scientists interpret the scientific method as applied to an observational science such as astronomy? (Other observational sciences include palaeontology and geology.) The key feature of a scientific theory, according to both the naive scientific method and Popper, is that it should make *testable predictions*. The definition of a "prediction" is obviously that it refers to a measurement which has not yet been made—explanations of existing unexplained effects are less highly regarded, because the framer(s) of the theory knew about these beforehand and may have "tuned" the theory to explain them. This does not, of course, mean that explanations of existing results have no value at all: if they are quantitatively precise and come "naturally" out of the theory with no obvious evidence of fine-tuning, they can be very persuasive (an example is the explanation by General Relativity of the excess 43"/century in the advance of the perihelion of Mercury: this was persuasive because (1) it was quantitatively correct and emerged without fine-tuning and (2) the previous Newtonian theory was extremely well understood and definitely did *not* have a good explanation).

In an experimental science, predictions can be novel because they refer to, and indeed suggest, an experiment which has not yet been done—for example, Maxwell's electrodynamics implied that oscillating electric and magnetic fields should propagate as a wave travelling at the speed of light; Hertz's production and detection of radio waves was inspired by and strongly confirmed this novel prediction. In contrast, astronomers cannot create new phenomena to test a new astronomical theory: they must rely on objects which are already "out there". However, theories might suggest particular kinds of observations, which have not been made before because nobody thought that they might be important: for example, the MOND theory discussed above makes specific predictions for the velocity dispersions of stars in globular clusters which differ from the predictions of standard Newtonian gravity (General Relativity is not relevant here because the gravitational fields are not strong and the velocities not close to the speed of light). These measurements had not been made at the time that the difference in predictions was recognised. Since then, a few globular cluster velocity dispersions have been measured, and the results turn out to support the existing Newtonian theory in preference to MOND. This is a valid test of the MOND hypothesis (which it turns out to fail—by Popper's criteria it should now be abandoned, although Kuhn would allow MONDians to keep working in the MOND paradigm while recognising the globular cluster results as potentially worrying anomalies).

Historically, observation has often outstripped theory in astronomy, and therefore the "natural explanation" class of support has more influence in astronomy than it does in most experimental sciences. Examples include the problem of stellar energy generation, which remained a serious difficulty for nearly a century (it became acute on the publication of the Origin of Species in 1859 and was finally solved convincingly by Hans Bethe in 1939), the nature and energy source of quasars, the origin of the chemical elements, and the cause of the different morphologies of spiral and elliptical galaxies (still not completely worked out). Occasionally theories come first, and wait for a long time before they are experimentally tested (the theory of neutron stars was first studied by Baade and Zwicky only a couple of years after the discovery of the neutron, but neutron stars were only observed (as pulsars) over 30 years later). The classic picture of new theory → prediction → test is actually fairly uncommon in astronomy (Eddington et al's test of Einstein's prediction of gravitational bending of starlight is perhaps the most widely known example). This is a clear difference between physics, for example, and astronomy, and accounts for a good deal of the notoriously inconsistent and illogical nomenclature of astronomy (if the theory of stellar spectra had preceded the Harvard stellar catalogue, we would not be stuck with OBAFGKM!).

Despite these differences in methodology imposed by the nature of an observational science, astronomy clearly does satisfy both Popper's and the "scientific method" criteria for science: it consists of a body of observational data interpreted and explained by a framework of theoretical ideas, all of which have been or are intended to be tested by observation—either by providing a natural, quantitative explanation of a previously unexplained body of data or by making predictions about the results of new observations. Theories which make predictions which are not borne out by observation are abandoned, e.g. the Steady State cosmology—albeit, as Kuhn would expect though Popper would not, rather more slowly and hesitantly than seems appropriate in hindsight. New theories generally have greater explanatory power than those that they replace: in Popper's terms, they are bolder; in Kuhn's, they have greater problem-solving potential. Both of these features are characteristic of science, and both help us to differentiate between astronomy and astrology: astrologers do not discard (even reluctantly!) theories which have been shown to make inaccurate predictions, and consequently astrological theory has not made the progress that we see in astronomy.

Note, incidentally, that the requirement that a theory should make testable predictions is not equivalent to a demand that *all* predictions made by a theory should be testable. For example, a cosmological theory which predicts multiple universes can still be scientific even though this particular prediction is not testable, provided that the theory makes additional predictions which *are* testable.

As we concluded above, astrology fails the science test because, although it makes falsifiable predictions, its practitioners do not accept the falsifications. Other astronomy-related pseudosciences fail for other reasons: for example, proponents of UFOs tend to make unfalsifiable/untestable statements (take an archive of sightings of "unidentified things in the sky", remove all those which do have clear explanations, e.g. Venus, aircraft, weather balloons, and then assert that the ones which do not have verified explanations are alien spacecraft—this clearly is not a testable hypothesis, since any observations in this category which are subsequently explained are simply reclassified in the other category!).

2.5 Philosophy of science and the history of astronomy

Both Popper's model of science and the naïve scientific method expect science to advance largely through the formulation and testing of hypotheses—though the hypotheses may in turn spring from the existence of a body of unexplained data. Kuhn's model of science predicts slow but steady progress during periods of "normal science", generally *not* through falsification but rather through deeper understanding of an existing theoretical framework or paradigm, punctuated by brief bursts of dramatic change when a new paradigm is established.

The history of astronomy provides examples of all of these models of scientific advance: the rise and fall of the Steady State model of cosmology is a case where a bold hypothesis (the Steady State is a strongly constrained theory which makes lots of testable predictions) is proposed, its predictions are tested, and the failure of the predictions leads to the abandonment of the hypothesis; the increasing precision of our knowledge of cosmological parameters since the 1990s is an example of normal science at work; the introduction of Kepler's laws of planetary motion, with their use of elliptical orbits rather than combinations of circles, represents a change of paradigm (though one could argue about exactly when the paradigm change takes place: the period from the mid-16th century to the end of the 17th century sees a wholesale shift in our view of planetary motions, but it takes place in several stages: Copernicus, Kepler, Newton). However, there are also examples of progress which are not so obviously consistent with any of the models we have discussed, most obviously the many cases where technological advances have led to new and initially unexplained discoveries, with the theoretical framework required to interpret them coming years or even decades later. Such events do not represent the failure of a prediction, and usually they do not really provoke a full-blown crisis of confidence in the current paradigm, but neither are they ignored as anomalies—instead, they typically open up a new field of investigation, which initially has no clear theoretical framework at all. Neither Popper nor Kuhn describes this situation well, and indeed many subsequent philosophers of science have introduced more complicated theories, such as Imre Lakatos' concept of "research programmes" to address issues such as these. However, some sort of interplay between theory and observation is certainly essential for the practice of science in general, and astronomy in particular. As we investigate the history of astronomy in this course, it is worth thinking about the events that we discuss in the context of philosophy of science and the meaning of the scientific method, not just as regards their effect on our understanding of astronomy.