

TAP 704- 8: The ladder of astronomical distances

Few measurements in physics have proved as difficult to make reliably, or have caused as much argument and dispute, as the seemingly simple question, 'How far is it to the stars and galaxies', or, in modern terms, 'How big is the Universe?'. Here you can find out a little about some of those arguments, and what they were about.

Challenge:

It's half as big as you say

The prestigious meeting of the International Astronomical Union in 1976 was startled to be told that the Universe is only half as big as the astronomers present all thought, and therefore only half as old. The challenger was the French-American astronomer Gerard de Vaucouleurs; the leader of the challenged orthodoxy was Allan Sandage, who had inherited the mantle of the American astronomer Edwin Hubble, who had in the 1920s first assembled evidence for the expansion of the Universe from red-shifts of galaxies. Sandage's main collaborator was Gustav Tammann.

De Vaucouleurs mounted a sarcastic attack on Sandage and Tammann's arguments for the distances of galaxies, accusing him and his collaborators of confusion, circular argument and observational errors. Both sides admitted uncertainties in their estimates, but the ranges of estimates of error did not even overlap. The dispute lasted for 20 years: today – with some irony – the generally agreed values seem to have settled down half way between the two. But for those 20 years astronomers had either to decide for one value or the other or in despair to split the difference.

How could such a strong difference of opinion about the facts come about? It arose because establishing distances to stars and galaxies does involve a long chain of reasoning, and an error at one step propagates through all the others. There was also a matter of belief involved. The size and age of the Universe which Sandage and Tammann preferred was large and long enough for models of the evolution of the Universe in which he believed to be right; the smaller size and shorter age which de Vaucouleurs advocated made serious trouble for these models. The Universe according to de Vaucouleurs came out dangerously close to being younger than some of the oldest objects, such as globular star clusters, to be found in it.

This was by no means the first time that changes to astronomical distance scales had been suggested. Successive revisions from the 1920s had in fact lengthened the scale, with Sandage himself having at least quadrupled the scale over that time. You may wonder how there can be such uncertainty in a measurement. The reason is that distances have to be worked out by making links in a long chain, each link introducing uncertainties and doubts of its own.

It all hangs by a hair

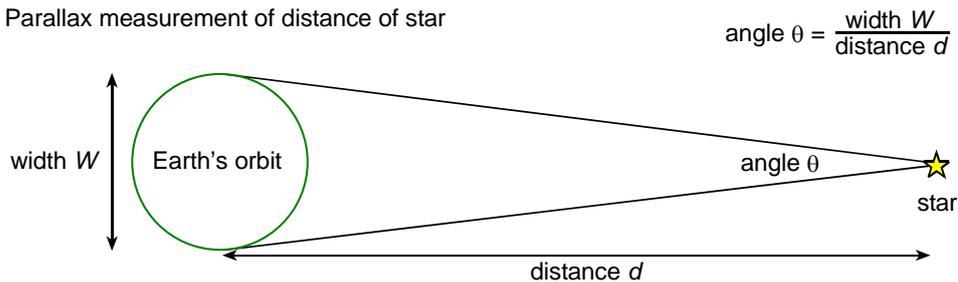
There are two ways to estimate the distances to the nearby stars, one of which is simple but bad and the other of which is difficult but good. The simple bad method is to assume that stars are all much of a muchness, and that all are about as bright as the Sun. Then comparing how bright they look with how bright the Sun looks gives you the ratio of the two distances, using the inverse square law. This turns out to be a bad method because stars vary in brightness by a factor of a million, though the astronomer Herschel used it to guess at the scale of the Milky Way, for lack of anything better. But bad as it is, this method illustrates several important features of astronomical distance measurement:

- If you have a class of objects of the same or known brightness, you can get distances just from the inverse square law of light intensity. Astronomers call such objects 'standard candles'. It just turns out that stars are bad standard candles.

- The result you get from this kind of method is a ratio of distances, not the distance itself. One of the distances has to be linked back somehow to a known distance, often through a series of other estimated distance ratios.
- Questionable assumptions have to be made: for example that the light is not dimmed by absorption along the way.
- There is a built-in bias (called Malmquist bias) because at great distances one only sees the brightest stars and compares them with average stars nearby. Thus large distances consistently get underestimated. It is one reason why the astronomical distance scale has usually tended to enlarge when corrected.

The ultimate distance, found directly by radar measurement, is the scale of the solar system, and in particular the diameter of the Earth's orbit.

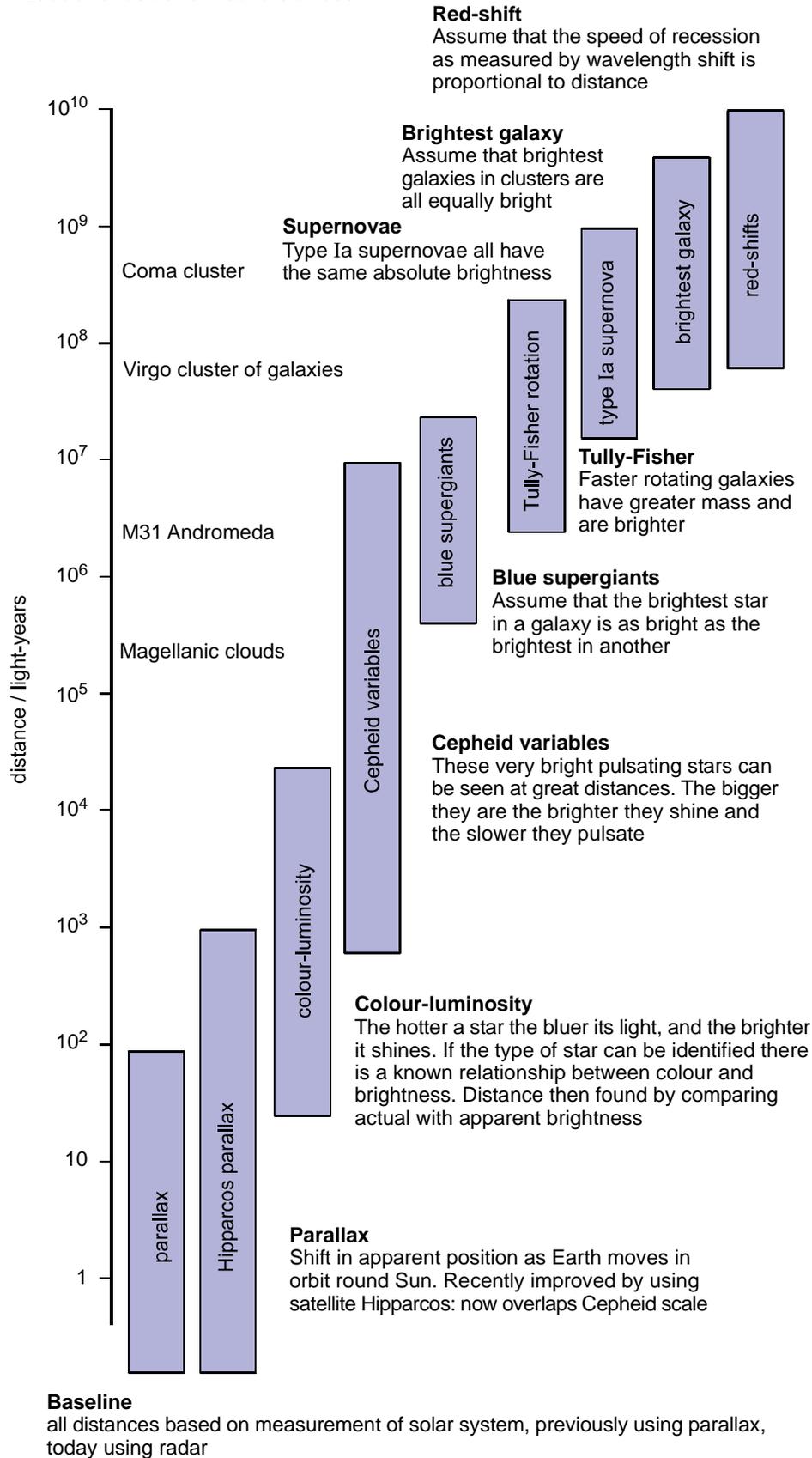
Parallax measurement of distance of star



The difficult but good method of finding the distances of nearby stars is to look for a shift in their direction, against the background of much more distant stars, as the Earth goes round its orbit. The effect is called parallax. You use it unconsciously yourself to judge distances of nearby things, using the baseline between your two eyes, which see things in slightly different directions. The trouble is that even the nearest star, Alpha Centauri, is more than 200 000 times further away than the diameter of the Earth's orbit. The shift in angle is less than 1 second of arc, less than the thickness of a hair seen across a large room. It was not until the mid nineteenth century that the first such parallax shift was measured. The distance at which a star would shift by 1 second of arc to and fro as the Earth moves round its orbit is called 1 parsec, equal to roughly 3 light-years.

Difficult as measurements of parallax are, they are not the source of substantial uncertainties in the astronomical distance scale. They are the place where, in the end, other estimates of ratios of distances link to the known diameter of the Earth's orbit. In recent years, good parallax measurements, previously limited to a few thousand stars, have been extended to greater distances by the scientific satellite Hipparcos. Crucially, Hipparcos provided parallax measurements of a number of the bright variable stars called Cepheid variables, which play an important part in the story of the astronomical distance scale.

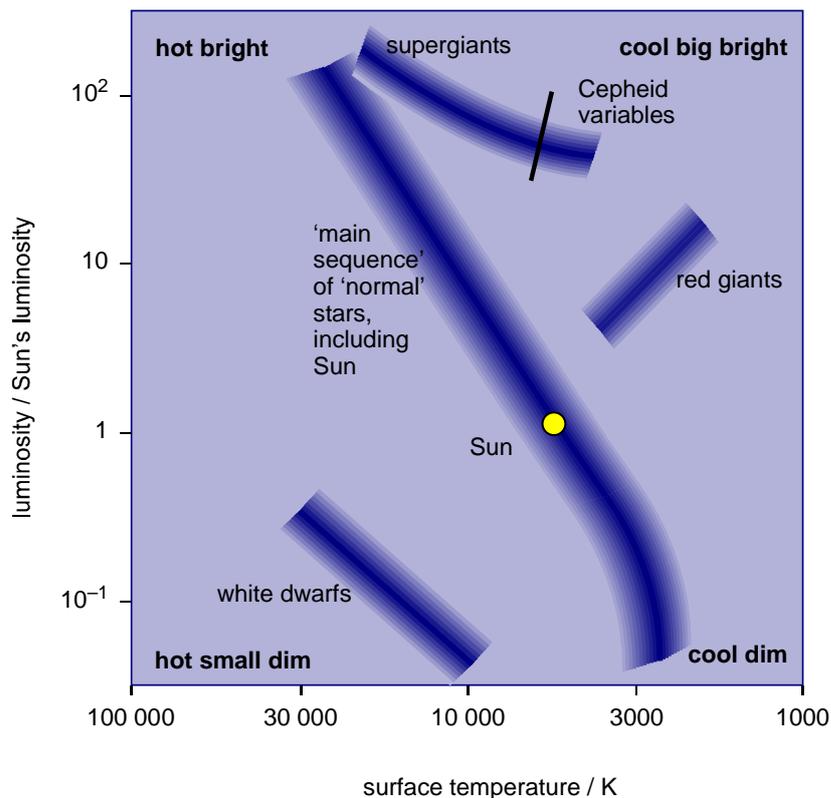
Ladder of astronomical distances



Colour and brightness

The first step beyond parallax in building a distance scale is a development of the simple but bad method of judging the distance of stars by their brightness. As stars were surveyed, it became clear that a large number of 'normal' stars ran on a scale from being very massive, very hot and very bright to being less massive than the Sun, less hot than the Sun and not so bright. The temperature of a star can be measured from its spectrum, intense in the blue if very hot and intense towards the red if relatively cool. So although it was a mistake to suppose that all stars were more or less equally bright, how bright many of them truly were could be found by looking at the spectrum of their light. Knowing the ratios of brightness of two stars, and again using the inverse square law, the ratio of their distances can be found. Again, by linking to the Sun, a link to the measured size of the Earth's orbit could be made.

Brightness and hotness of stars:
Hertzsprung-Russell diagram



Henrietta Leavitt measures Cepheid variables

In 1912 women of talent found it hard to get university positions to do science. But they were often welcome doing work such as making calculations for their male colleagues. The American Henrietta Leavitt found herself such a role at Harvard. In 1912, photography was beginning to be used not just to make pictures through telescopes, but to make careful measurements. Over time, Miss Leavitt became a world authority on the new science of measuring the optical density of images of stars, so as to measure how bright they appear to be. She made a study of a special kind of variable stars which are so bright that they could be detected outside the Milky Way Galaxy, in a 'satellite' star cluster called the Large Magellanic Cloud, visible to the naked eye in the Southern Hemisphere. These Cepheid variable stars vary in brightness over periods from 2 to 200 days. Detecting them calls for painstaking examination and measurement of repeated images of thousands of stars, to detect those which are varying, and how rapidly.

Henrietta Leavitt found that there is a simple relationship between the brightness of the Cepheid variables and their period of oscillation: the slower the brighter. She could assume that differences in the apparent brightness of the Cepheid stars in the Magellanic Cloud were

due to real differences in brightness, since they were all more or less at the same distance. She could not know how bright the stars were, or how far away they were. But she provided a way for astronomers to compare true brightness just by clocking the time of variation, and so to get ratios of distances from ratios of apparent brightness.

Since then, Cepheid variables have been an essential link in the astronomical distance scale. With parallax measurements to nearby Cepheids from the Hipparcos satellite, and with the Hubble telescope just recently able to detect Cepheid variables as far away as the Virgo cluster (tens of millions of light-years), it turns out that Henrietta Leavitt identified one of the strongest and longest links in the chain of constructing an astronomical distance scale. The value of her work was recognised also, back in the 1920s, when Edwin Hubble observed Cepheid variables in the nearest large galaxy, M31 in Andromeda, establishing it as an 'island universe' outside and on the same scale as the Milky Way. So Henrietta Leavitt contributed also to the first solid evidence that the building bricks of the Universe are the galaxies, each with perhaps 10¹¹ stars. But, just to illustrate the difficulties, you should know that Hubble's 1920s estimate of the distance of M31 was half as big as the current estimate, because of the indirect way he had to use to estimate the distance of the Magellanic Clouds.

More brightness guesswork

Beyond the range of the Cepheid variables, only the very brightest of stars could be picked out in distant galaxies. These are the blue supergiant stars, very big, bright and hot. The distance scale could be pushed out further if one assumed that there is some limit to how bright and big a star can get. If so, the very brightest stars in a galaxy would all be equally bright, and could be used as standard candles as far out as they could be seen. You will see for yourself that this is a rather dangerous assumption to make. Could there not be other perhaps unknown kinds of even brighter object? (And in fact there are). But for a time it was the only way available to get distances of distant galaxies from the brightness of stars in them.

A lucky break:

Brighter galaxies spin faster

Beyond the distance where individual stars could be seen in galaxies, the only hope was to make further dangerous assumptions, for example that galaxies of the same type are equally bright, or equal in size. Neither method is helped by the fact that galaxies are seen at all angles to the line of sight, head-on, sideways, and in-between.

In 1972 two young American graduate students, Brent Tully and Richard Fisher, had the idea of using radio astronomy to measure how fast distant galaxies were rotating. The idea was that the radio emission at 21 cm from hydrogen in a galaxy (then recently used to map the arms of the Milky Way) would show through Doppler shifts in its frequency the difference in speeds of the gas towards or away from us as the galaxy rotates. The point of finding out how fast a galaxy rotates is that more massive galaxies should rotate faster, and more massive galaxies should be brighter. Tully and Fisher were after a clock which could measure galaxy brightness, much as Henrietta Leavitt had found for Cepheid variables.

People were not slow to point out the dangerous assumptions involved. Most serious was that most of the brightness of a galaxy comes from its hot blue stars, while most of its mass is in the bulk of cooler redder ones and indeed in dust and gas clouds. So the empirical relation they could show for galaxies whose distances had been estimated by other means, between brightness and rotation speed, was a relationship between not really connected things. Luckily for Tully and Fisher, in more recent years infrared telescopes on satellites have confirmed that the infrared brightness is also related to rotation speed. The Tully-Fisher step on the distance ladder has proved very useful.

Standard outbursts

One very special and very rare kind of stellar object can be seen out to nearly the largest distances telescopes can reach. They are the stellar explosions called supernovae, in which a cataclysm in a dying star briefly makes it shine as brightly as a whole galaxy. There are two types, Type 1 and Type 2. Type 2 explosions are of giant stars, and vary in brightness depending on the mass of the star. But Type 1 are remarkable standard candles, each as bright as another. This is because they come from a collapsed dwarf star which collects matter from a neighbouring star. It happens that such a collapsed star cannot ever be more than 1.4 times the mass of the Sun. Beyond that mass, it cannot support its own gravitational attraction, and collapses yet further. These second collapses all start from the same point, the point when the dwarf has just collected the maximum mass possible, and all work in the same way. This makes every such explosion equally big. But they are also rare, perhaps 20 or 30 a year in a given galaxy. So to use them for distance measurement means looking at thousands of distant galaxies in the hope of detecting a few such outbursts.

Practical advice

Students may ask how the distance to different objects was established. The text is intended to provide background information and aid a discussion of how the distance scale was established

The scale is subject to constant revision. The story is a good example of substantial uncertainty in physics, and it remains true that the distance scale and so the Hubble constant are amongst the least accurately known facts in physics.

The story provides a good opportunity to discuss the calibration of measurements, for example the calibration curve needed between Cepheid time periods and brightness. It also illustrates well how errors can propagate along a chain of argument.

With less able students it may be best to build a class discussion around the diagram showing some of the methods on the astronomical distance ladder. It would be a mistake to go into too much detail, however. Better students will be able to take the reading as it stands, and perhaps describe the main ideas to others.

A compact account at the right level for a teacher needing some updating on astronomical distance measurement can be found in M Rowan-Robinson (1996) *Cosmology* 3rd edn (Oxford University Press).

Alternative approaches

You may prefer to use one or more of the excellent popular books that discuss the measurement of astronomical distances. One which gives a lot of personal detail and historical background is Dennis Overbye (1991) *Lonely Hearts of the Cosmos* (Macmillan). The story is told briefly in M Rowan-Robinson (1993) *Ripples in the Cosmos* (Freeman), and in technical detail in M Rowan-Robinson (1985) *The Cosmological Distance Ladder* (Freeman).

Social and human context

The distance scale is a dramatic and very real example both of uncertainty in physics, and of the bitter arguments that sometimes arise. The clash between Sandage and de Vaucouleurs was fierce and long, and harsh words were spoken. An initial enmity may have arisen when de Vaucouleurs had, in Sandage's opinion, used some of Sandage's unpublished data without permission.

External reference

This activity is taken from *Advancing Physics*, chapter 12, 20T