The Broken Cosmic Distance Ladder
Professor Roberto Trotta
31st January 2022

Measuring distances to astronomical objects outside our Galaxy is a surprisingly hard challenge: it wasn't until 1923 that Edwin Hubble obtained proof that Andromeda is indeed a galaxy in its own right. Today, astronomers extend distance measurements in the cosmos to the edge of the visible Universe, building up a 'cosmic distance ladder' made of several rungs.

This talk will explore a major conundrum of contemporary astronomy: as observations have become more precise, the distance ladder appears today to be broken.

A Rabbit on the Moon

At times the Moon appears so close that it seems one could touch it simply by climbing on a tall enough ladder. The outline of a rabbit, etched in a darker shade on its surface, seems to be leaping out at us - or in some cultures it is a frog, or a toad, or a man carrying a bundle of wood on his back (I must confess, however, that I always have the hardest time to make out the face of the man that's also allegedly there!). The stars, on the other hand, appear always aloof, and forever out of reach. As W.H. Auden put it in his poem, *The More Loving One:*

> *How should we like it were stars to burn*
> *With a passion for us we could not return?*
> *If equal affection cannot be,*
> *Let the more loving one be me.*
>
> *Admirer as I think I am*
> *Of stars that do not give a damn,*
> *I cannot, now I see them, say*
> *I missed one terribly all day.*

But when we happen to catch the sight of a "shooting star", it seems as if for one brief instant one of them has briefly entered our Earthly domain (in reality, all we have witnessed is a speck of cosmic rock, a meteorite, burning up in the atmosphere).

How far away is the Moon? And the Sun? And the thousands of stars that bejewel the night? And what about the myriad nebulae, fuzzy glowing patches of light that only the big telescopes of the 20th century could clearly reveal? How distant are they? Can our human eye combined with our powers of imagination reach the end of the visible universe, and extend a cosmic measuring tape all the way to there? These are the questions we will address today, casting our mind to the endeavour of building a "cosmic distance ladder", a method for measuring distances that rung after rung, century after century, has taken us all the way to the end of the observable cosmos, and therefore all the way back in time to almost the very beginning. Mysteriously, that ladder appears today to be broken - and how to fix it is one of the most pressing questions of modern astronomy.

A Jumping Finger: The First Rung

The earliest method to determine distances to heavenly objects remains very much useful today, and it is literally at your fingertips: stretch your arm in front of you, make a fist and extend your index finger pointing upwards. Close your right eye and observe the position of your index finger against a background landmark; now close your left eye instead and notice how the finger has jumped to a different place with respect to the distant background. This effect is due to the different vantage point of each of your eyes, which are about 10 cm apart and thus have a slightly different line of sight to your finger, in the foreground, with respect to the background. Elementary trigonometry allows us to compute the distance to the finger if we know the separation between the eyes, and we have measured the angle between the two apparent positions of the finger. This phenomenon is called "parallax", from a Greek word meaning "a change". If you now move your finger closer to the tip of your nose and repeat the experiment, you will see that the jump in its apparent position as you alternate the eye that's open is far greater.

Parallax is a great tool to obtain distances of relatively nearby objects, provided we have a more distant background of reference points to judge their apparent position by. The experiment with the finger also shows that the more distant the finger, the smaller its parallax. We can anticipate that this method of determining distance to heavenly bodies will take us only so far in distance. There comes a point when the parallax becomes too small to be measurable.

The first person to use this method to try and measure the distance to the Sun and the Moon is reputed to be the ancient Greek astronomer and mathematician Hipparchus of Nicaea (in modern-day Turkey), who lived in the second century BC. Hipparchus is the inventor of trigonometry and is considered "the father of astronomy" for his many discoveries, which include the precession of the equinoxes, the fact that the seasons are of unequal length, a method to predict more accurately the occurrence of eclipses and the creation of the first astronomical catalogue encompassing both the position and brightness of 850 stars.

Hipparchus used the fact that a solar eclipse, presumably in 190 BC, appeared total in his birthplace Nicaea but only partial as seen from the city of Alexandria, on the same meridian but 9 degrees further South. By attributing the difference in aspect to the parallax of the Moon, and by noting that the Sun was sufficiently far away to not exhibit any noticeable parallax (at least not for the observers at that time), he managed to estimate the distance of the Moon in units of the Earth diameter. He eventually arrived at a figure of 63 times the Earth's radius for the Moon mean distance – remarkably close to today's accepted value of about 60 times (or approximately 384,000 km).

Erathostenes had already provided a measurement of the size of the Earth 50 years earlier. A traveller told him that on the summer equinox, the Sun's rays at noon illuminated the bottom of a well in the city Syene, in modern-day Egypt, without casting any shadows, which meant that the Sun was directly overhead. He then measured the length of the shadow cast by a stick in Alexandria on the same day, thus finding the difference in latitude between the two cities to be 1/50th of the circumference of a circle. All that was needed now to determine the circumference of the Earth was the physical distance between Alexandria and Syene: he hired professional surveyors who walked with equally spaced steps between the two cities, reporting a distance of 5000 stadia. He concluded that the circumference of our planet was 250,000 stadia, a number that is difficult to translate in modern units for the ancient unit of length of a stadium was not of uniform dimensions. Taking 10 stadia to the mile, Earthostene's estimate would be 25,000 miles, putting the Earth radius (assuming a perfect sphere) to 3,980 miles or 6,400 km, very close to the actual value of 6,378 km.

Be as it might, Hipparchus' value of 63 Earth radii did put the Moon out of reach of even the tallest of ladders on top of the highest of mountains. But it did establish the lowest rung in what, in time, would become to be called "the cosmic distance ladder".

Horrocks' Splendid Sight: The Second Rung

While Hipparchus' value for the mean Earth-Moon distance was quite accurate, his estimate for the solar distance came out hopelessly low: he put it at 490 Earth radii, compared to the actual value of 23,400 times (or 150 million km). A better method was required.

By the early 17th century, the three laws of planetary motion announced by Kepler (which Newton would later explain from his more fundamental theory of gravitation) put the planets on elliptical orbits around the Sun, which is sitting in one of the foci of the ellipses. Kepler's third law said that the square of the orbital period of a planet divided by the cube of the semi-major axis of its elliptical orbit was a constant. This provided a way to obtain the relative distances from the Sun to all the planets in the Solar system, given knowledge of the time they took to complete a revolution around the Sun. If the Earth's distance to the Sun was assigned a value of 1 "astronomical unit" or AU, it followed from Kepler's third law that Mercury was 0.39 AU from the Sun, Venus 0.72 AU, Mars 1.52 AU, Jupiter 5.20 and Saturn 9.5 AU (the outer planets would only be discovered later). In order to obtain the absolute distances of all the planets, and thus determine the actual size of the Solar System, one needed a way to measure the AU.

That's where Edmond Halley came in. In 1716, the famous astronomer issued a call to arms to all "diligent searchers of the Heavens" to take advantage of the rare astronomical phenomenon of Venus transiting in front of the disk of the Sun to at last measure the Sun-Earth distance. Venus being the second planet in the Solar System, its orbit around the Sun is smaller than the Earth's, and thus its period shorter: the Sun, Venus and the Earth line up in a straight line approximately every 1.5 Earth years. However, because of the relative inclination of the Venusian and Terrestrial orbits, Venus is seen transiting across the face of the Sun only in the rare occasions when the alignment matches the intersection of the planes of the planet's orbits – a rare phenomenon, with a period of 243 years. For an observer on the surface of the Earth, Venus will pass in front of the Sun (a so-called "transit") twice 8 years apart, then nothing for 105 years, then another pair of passages 8 years apart, and nothing again for the next 121 years, after which the cycle recommences. Halley, writing in 1716, rallied "natural scientists" to mount a world-wide expedition for the pair of transits expected in 1761 and 1769. Venus, as seen from different locations on the Earth, would cut a different path through the Sun's disc, and timing the duration of the transit would give a mean of measuring the solar parallax, and with it the absolute distance between the Earth and the Sun (Figure 1).

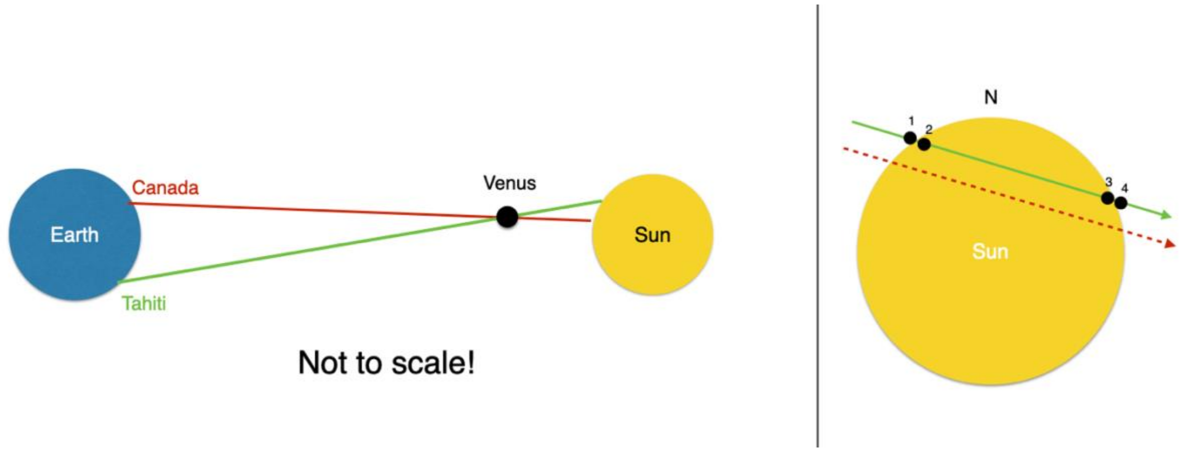

Figure 1: During a transit of Venus, observers in different locations will see the planet take a different path across the face of the Sun due to parallax. Timing the transit allows to determine the distance to Venus, and hence the size of the Solar System.

The English astronomer Jeremiah Horrocks had been one of the only two people in recorded history to observe the previous transit in 1639, which Horrocks himself had predicted against Kepler's forecast of a near miss. A memorial tablets commemorates him inside Westminster Abbey thus: "Having in so short a life detected the long inequality in the mean motion of Jupiter and Saturn; discovered the orbit of the moon to be an ellipse; determined the motion of the lunar apse; suggested the physical cause of its revolution; and predicted from his own observations the transit of Venus, which was seen by himself and his friend, William Crabtree, on Sunday the 24th of November 1639; this tablet, facing the monument of Newton, was raised after the lapse of more than two centuries, December 9th 1874."

Horrocks was so elated at the spectacle that he described it in these lyrical terms:

> *" ...Thy return*
>
> *Posterity shall witness; years must roll*
>
> *Away, but then at length the splendid sight*
>
> *Again shall greet our distant children's eyes."*

Sadly, Horrock's destiny was not to have any children who could witness the "splendid sight" in the distant future. He died two years later, and the tragically young age of 22.

When the 1761 transit came, over 120 missions in 62 countries were set to observe it, in the grandest international scientific effort ever attempted. The Royal Astronomer Neville Maskelyne, for example, sailed all the way to St Helena, only to have his observations scuppered by clouds. Halley's good wishes expressed in 1716 ("And I wish them luck and pray above all that they are not robbed of the hoped-for spectacle by the untimely gloom of a cloudy sky") hadn't been enough. Despite the many measurements made during the 1761 transit, the quality of the data was too variable to arrive at a satisfactory agreement about the solar parallax, and hence the Earth-Sun distance. A second chance would present itself in 1769, then nothing for over a century. This time, failure was not an option.

In England, the Royal Society mounted an ambitious scientific expedition: they convinced the Admiralty to buy and outfit a sturdy merchant ship, renamed *The Endeavour*, and to give its command to the ablest navigator and cartographer among his Majesty's seamen, Lieutenant James Cook. They petitioned the King, George III, to bankroll the mission with a grant worth over a million pounds in today's money; and they equipped the Endeavour with the finest telescopes and clocks that money could buy, before sending it to the other side of the planet, to the small tropical island of Tahiti, which had conveniently just been discovered by Capt. Wallis. When the transit occurred on June 3rd, 1769, it found captain Cook, the astronomer Charles Green and the botanist Daniel Solander ready to observe it, and blessed with perfect conditions, as the captain reported in his logbook:

"This day proved as favourable to our purpose as we could wish. Not a Cloud was to be seen the whole day, and the Air was perfectly Clear, so that we had every advantage we could desire in observing the whole of the Passage of the planet Venus over the Sun's Disk."

The data Cook brought back to England allowed Thomas Horsnby, the Savillian Professor of Astronomy at Oxford, to derive a value of the AU equal to 93,726,900 English miles, or an error of less than 1% with respect to the actual value!

The Sounding Line Touches Bottom: the Third Rung

By the end of the 18th century, the size of the solar system had been reliably established thanks to the transit of Venus observations: Uranus, discovered in 1781, stood at an average distance from the Sun of 2.8 billion km; when Neptune was discovered in 1846, that already enormous expanse of empty space would further grow to almost 5 billion km. But the stars hadn't been pinned to a cosmic ruler yet. Everybody agreed they had to be almost unconceivably far away, but exactly how far, nobody could say. To paraphrase John Herschel, the sounding line in the universe of stars had not yet touched bottom.

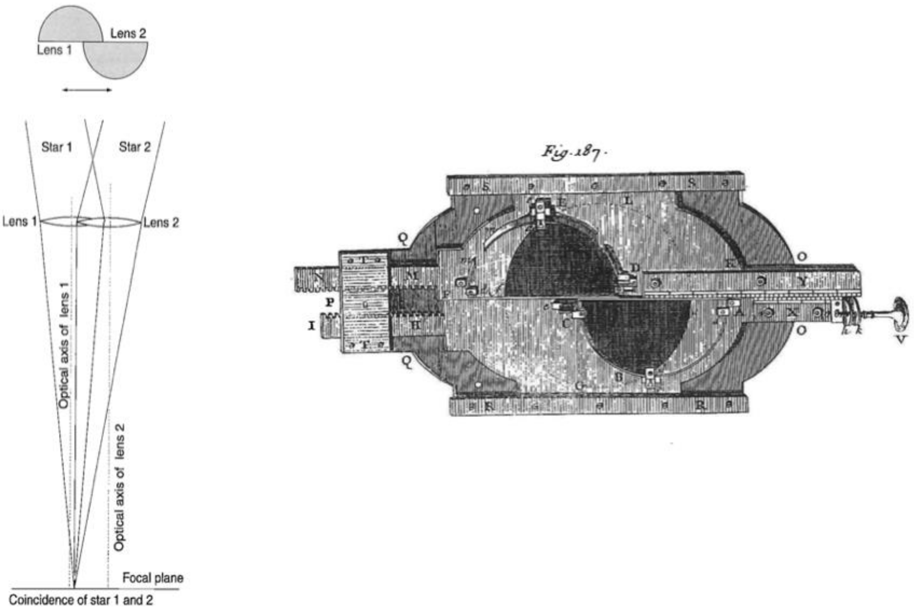

Figure 2: The heliometer, a lens sliced in two parts which can be shifted with respect to each other by means of a thumbscrew, allows the measurement of tiny angles between stars. Source: Willach (2004).

A parallax measurement six months apart could in principle do it, by exploiting the change of vantage point between opposite ends of the Earth's orbit around the Sun, a massive 300 million km. Even so, the immensity of the distance to the stars could make the parallax angle so small as to become potentially impossible to measure. To acquire a sense of scale for angles, consider that the angle subtended by the full Moon is about half a degree. A degree is further subdivided in 60 minutes of arc (or "arcminutes"); each minute of arc is, in turn, subdivided in 60 seconds of arc (or "arcseconds"; the reasons for the choice of 60 subdivisions are fascinating, and go back to Babylonian times, over 5,000 years ago!). So the full Moon's diameter is about 30 arcminutes, though this value changes as the Moon's distance to the Earth is not constant. The unaided human eye can discern angles as small as 1 arcminute: this corresponds to the angular size of a 1-pound coin seen from 80 meters away. An arcsecond is 60 times smaller: the angular size of a pound coin observed from 4.8 km away. Current space-based observatories, with the advantage of flying above the trembling of the atmosphere, have achieved accuracies of a millionth of an arcsecond – this is the angle subtended by a pound coin on Neptune, as seen from the Earth.

The parallax angle is so fundamental to astronomy that it was used to define, in 1922, a basic unit of measure: the "parsec". A contraction of "parallax" and "second", one parsec is the distance at which a star would show a parallax angle of 1 arcsecond. It works out at 3.26 light years – just over thirty thousand billion km! The nearest star, Proxima Centauri, is a little over a parsec away, so its parallax is a little smaller than 1 arcsecond. In order to measure the distance to the stars, 19th century astronomers needed to build instruments capable of spotting a coin 5 km away. The German mathematician and astronomer Friedrich Bessel, the director of the Königsberg Observatory (in today's Kaliningrad, on the Russian Baltic), did it in 1838.

To do so, he used an heliometer (Figure 2), a telescope designed to measure the diameter of the sun, in which the lens had been sliced in half, and the relative position of the two semi-lenses could be adjusted by turning a screw – in so doing, the double image of a pair of stars produced by the sliced lens could be made to coincide, and their angular difference determined from the setting of the screw. Bessel carefully monitored the position of the star 61 Cygni over the course of a year with respect to background stars with the heliometer, finally announcing a parallax of 0.314 arcseconds: this pinned 61 Cygni to about 3 parsecs away, or 10 light years.

A new rung had been added to the cosmic distance ladder, one that was described as “the greatest and most glorious triumph which practical astronomy has ever witnessed” by Royal Astronomical Society president John Herschel as he awarded to Bessel the Society's Gold Medal in 1841.

Miss Leavitt's Relationship: The Fourth Rung

But of course, the sounding line of distance hadn't touched bottom yet. The question was merely displaced from the province of the stars, whose distances could now be routinely measured up to thousands of light years away, to the mysterious realm of the nebulae. Those faint puffs of bluish haze, sometimes showing the hint of twirling lanes and arms, remained unknown in their nature, unmeasured in their distances. If astronomers could only somehow find their distance, this would give them a clue as to their nature: nearby vaporous masses of hydrogen, or mighty galaxies in their own right, reduced to faintness by their unimaginable distance to us?

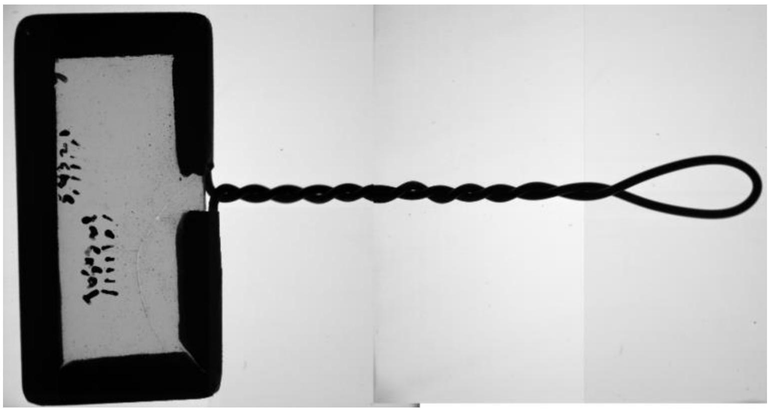

Figure 3: The "fly spanker" which Henrietta Leavitt used to gauge the brightness of variable stars. Picture credit: Harvard College Observatory.

With the turn of the 20th century, a new, powerful tool came in aid to the astronomers: the photographic plate. One of the pioneers of the technique had been Henry Draper, a medical doctor and keen amateur astronomer who, together with his wife Mary Anna Draper, was among the first to capture stellar spectra on photographic plates. His untimely death in 1882 spurred his widower wife to support generously over several decades the production of the Henry Draper Catalogue of stellar spectra at Harvard Observatory – an effort that would eventually produce over half a million photographic plates. The ever-growing collection of plates required an army of "computers" – women who would scrutinise each plate to determine position, apparent magnitude and spectrum of each object. Figure 3 shows the so-called "fly spanker", a glass surface with calibrated magnitudes used to estimate visually the magnitude of stars (so called as it was too tiny to be a fly swatter).

The most gifted and prolific of the lady astronomers was undoubtedly Miss Annie Jump Cannon, who over forty years of work at the Observatory personally classified over 200,000 stars! But she was by no means the only one whose contributions transformed the way astronomy was done. A contemporary of Miss Cannon, Miss Henrietta Swan Leavitt had excelled in her studies at Radcliffe College, one of the few women's colleges at the time. She was particularly gifted in mathematics, algebra and calculus. When she landed an unpaid assistantship at the Harvard Observatory in 1895, the director, Edward Pickering, charged her with assessing magnitudes of stars from photographic plates. When, after a stint in Europe and then Wisconsin, she returned in 1903, she was offered 30 cents an hour to first identify variable stars in plates of the Magellanic Clouds, and then measure the changes in their brightness as a function of time -- a painstakingly difficult job. By 1908, she had amassed 1,777 variable stars, and noticed a strange pattern among 16 of them in the Small Magellanic Cloud. "It is worthy of notice –she wrote– that the brighter variables have the longer periods".

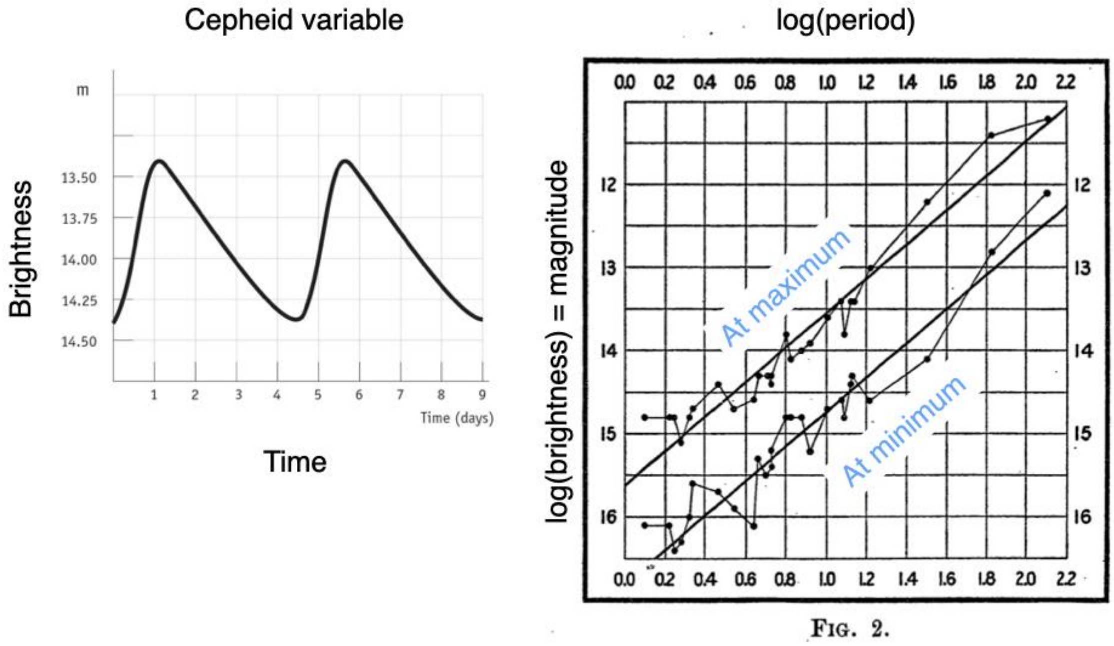

Figure 4: A Cepheid variable light curver (left) and the luminosity-period relationship (represented by the two diagonal lines) discovered by Henrietta Leavitt in 19212. Sources: ESA/ESO; Leavitt and Pickering (1912).

By 1912, she had found another 9 variables exhibiting the same pattern: the period of variability of those remarkable stars appeared to reflect their brightness (Figure 4). The brighter the star, the longer the period. Since all the stars were in the Small Magellanic Cloud, it was reasonable to assume that they were all, more or less, at the same distance from Earth. Therefore, their different apparent brightness reflected a difference in their true (or intrinsic) brightness, not distance. This was a ground-breaking result: once calibrated, Miss Leavitt's law gave astronomer the means to establish the absolute brightness of a variable star by observing its period; from the observed brightness they could then find the distance, as apparent brightness drops as the square of the distance. We today call Miss Leavitt's variable stars "Cepheid variables", after the prototype star of this kind in the constellation of Cepheus. This kind of stars, several times more massive than the Sun, trap some of their radiation in the outer layers, thus expanding and eventually cooling. The cooling allows the radiation to escape: the star brightens and contracts, and the cycle begins anew. It is as if the star was breathing, letting light out at regular intervals.

Guided by Miss Leavitt's discovery, Edwin Hubble spent hundreds of nights at the 100-inch Hooker telescope at Mount Wilson observatory in California, before he finally tracked down two such variable stars in the Andromeda Nebula. His historic photographic plate still bears his handwritten, exclamatory annotation next to the first one he found: *VAR!* Thanks to Miss Leavitt's relationship, Hubble was able to estimate the distance to Andromeda from the period of the variable, obtaining a million light years: less than half the actual distance, but still enough to deal a fatal blow to those who argued that Andromeda wasn't a galaxy in its own right, but just a relatively nearby gas cloud. Among the proponents of the latter theory was the newly appointed director of Harvard Observatory, Harlow Shapley – who ironically had perfected the calibration of Leavitt's law. So it was only fitting that when the letter from Hubble arrived, triumphantly announcing his definitive discovery ("Dear Shapley, you will be interested to hear that I have found a Cepheid variable in the Andromeda Nebula"), Miss Cecilia Payne, a Cambridge graduate who had taken over Miss Leavitt's old desk, found herself witnessing Shapley's reaction to it: "This is the letter that has destroyed my universe", he declared.

Hubble's discovery, thanks to Miss Leavitt's law, added a fourth rung to our cosmic distance ladder – its top end now reaching into the millions of light years. The distance ladder began to seriously stretch.

Catching Star Explosions: The Fifth Rung

Cepheid variables opened up a much grander vista, delivering reliable measurements of distances far beyond the realm where parallax methods are applicable. As telescopes grew bigger, and photographic plates more sensitive and eventually replaced by digital sensors, fainter and fainter galaxies became observable – so distant that there was no hope to see any individual star in them, much less to measure the variability of a single Cepheid. The distance ladder needed a new, farther rung.

The light from distance galaxies is reddened by the expansion of the universe, as it loses energy on its way to us – a discovery that is usually attributed to Edwin Hubble in 1929, but should really be credited to George Lemaître, who demonstrated it two years earlier (see my Gresham lecture, "Einstein's Blunder"). This uniform reddening of light, called "redshift", tells us how much the universe has expanded since the light was emitted, but says nothing about the distance to the galaxy that emitted it. When multiplied by the speed of light, the redshift can also be interpreted as the "recession velocity" of the galaxy that emitted the light – it is not a velocity in space, but rather caused by the expanding universe (as an aside, this recession velocity can and does exceed the speed of light for sufficiently distant galaxies, in an apparent –but only so– contradiction to special relativity!).

Einstein's theory of General Relativity describes the relationship between distance and redshift by a function containing a number of unknown quantities. For small distances (where "small" here is a few hundred million light years!), the relationship is a straight line, meaning that the redshift (or recession velocity) is proportional to distance, with the constant of proportionality called the *Hubble-Lemaître constant* (denoted by the symbol H_0, pronounced h-not). This constant tells us how fast the universe is expanding at the present day: it is measured in units of km/s/Mpc (where Mpc stands for a million parsecs, or 3.26 million light years), so that, for example, a value of 70 km/s/Mpc means that the recession velocity of a galaxy increases by 70 km/s for every Mpc of distance from us. So, a galaxy 100 million light years away (or about 30 Mpc) is moving away at a speed of 2,100 km/s.

The Hubble-Lemaître constant is thus one of the key quantities in cosmology. Hubble and Lemaître before him, working out of a handful of distances to nearby galaxies, massively overestimated its value, putting it in the region of 500 to 600 km/s/Mpc. The value has been going down ever since, as better and better distance measurements became available. Its exact value has sparked one of the hottest controversies in cosmology today.

At even greater distances, the distance-redshift relation begins to deviate from a straight line, in a way that is characteristic of the matter and energy contents of the universe – this is what controls the expansion rate of the cosmos. Therefore, by measuring the distance to very far-away galaxies and comparing it with their redshift, we can determine what the universe is made of. The story of how two teams, one led by astronomers, the other by particle physicists, engaged in a madcap race to be the first to measure the distance-redshift relationship for very distant galaxies has become one of the classic epics of the field. Their surprising discovery, in 1998, was that the new rung of the distance ladder was much farther away than anybody expected: this meant that the expansion of the universe had begun to accelerate again 6 billion years ago, after having been decelerating under the influence of gravity since the Big Bang, 13.8 billion years ago (the implications of this discovery are further discussed in my Gresham lecture "Einstein's blunder").

To peer even further out in the universe and measure distances of billions of light years, it had been necessary to resort to a new kind of light beacon: a type of stellar explosion called "supernova type Ia". These explosions are very bright, for a few days shining as bright as 10 billion Suns and can thus be detected from halfway across the universe. They are caused by the thermonuclear runaway reaction of a compact, dense star called "a white dwarf", which gains mass from a companion star. The additional mass increases the temperature of the core of the white dwarf, and this triggers the explosion. Because of the physical threshold of temperature that ignites the star, type Ia supernovae explosions are all similar, and their similarity to one another can be further improved by accounting for the time it takes for them to fade away: the brighter explosions take longer to fade. The time from maximum brightness to disappearance can thus be used to determine how intrinsically bright the explosion was, and by observing the apparent brightness one can determine distance – a similar trick to Miss Leavitt's law.

The Ladder Dangling from the End of the Universe

While astronomers were busily adding new rungs to their ladder, cosmologists built their own contraption -- and instead of planting it firmly on the ground and going up, like their colleagues had been doing for centuries, they decided to do it in exactly the opposite way.

In their quest to understand the fundamental properties of the universe, cosmologists shot for their holy grail, the leftover light from the Big Bang. This cosmic radiation is the oldest light in the universe, and also the most distant light: it comes from the very early times of the cosmos, a mere 380,000 years after the beginning. At this point in time, galaxies, stars and planets did not exist yet -- the universe was filled with light, neutrinos, dark matter, hydrogen and helium (plus a sprinkling of beryllium and boron). Detailed measurements of this relic radiation allowed cosmologist to map out the baby universe, and thereby reconstruct many of its characteristic properties with very high precision – a process I describe in more detail in my Gresham lecture "Weighing the Universe".

Among the quantities that the relic radiation can measure is the Hubble-Lemaître constant. This is not a direct measurement, as the Hubble-Lemaître constant expresses the present-day speed of expansion of the universe, and the relic radiation comes from a long time ago, 13.8 billion years in the past – when the light was emitted, the expansion speed was different than it is today. However, as the Hubble-Lemaître constant enters the equations that are used to analyze the relic radiation, its value can be indirectly inferred from the data of this radiation from a long time ago. This is subject to an important caveat, though, namely the assumption that the universe is spatially flat – a well-motivated hypothesis, supported by the theory of cosmic inflation (see my Gresham Lecture "Weighing the Universe"). We know the redshift of the relic radiation very precisely, but its distance once again depends on the value of the Hubble-Lemaître constant: by dialling the Hubble-Lemaître constant up and down we can, as it were, bring the relic radiation into sharper focus by changing its distance. The distance at which the data snap into focus corresponds to the best choice of the Hubble-Lemaître constant. In this sense, the primordial relic radiation is the ultimate distance ladder rung – it sits at the farthest possible distance from us. Before that moment in time, the universe was rendered opaque to light by its high temperature; it is physically impossible to see past the relic radiation (unless one uses neutrinos or gravitational waves to do so).

The astronomers' distance ladder, painstakingly built up from parallax measurements, Cepheid variables and ultimately supernova explosions, reaches out billion of light years into space; at the same time, data on the primordial relic radiation, coming from 13.8 billion years ago, only snap into precise focus when one chooses the right value of the Hubble-Lemaître constant. One would hope that these two measurements of one of the key quantities describing our universe would match – bringing the nearby universe in perfect alignment with the most distant observable parts of the cosmos.

The cosmology community was stunned to learn that they don't.

The SH0ES That Broke the Ladder

The colourfully named SH0ES programme (for "Supernovae and H0 for the Equation of State of dark energy") started in 2005 as an effort to use the Hubble Space Telescope to refine the measurement of its namesake constant, H_0 (The Hubble-Lemaître constant was known until 2018 as the "Hubble constant"; the name of Lemaître was added by the International Astronomical Union in recognition of his pioneering work on the scientific theory of the expansion of the universe, that until then had been somewhat underrecognized). Orbiting Earth above the atmosphere, the Hubble Space Telescope has a uniquely clear view of the cosmos; importantly, its ability to measure accurately the brightness of its targets remains constant over time, ensuring that no errors are introduced in the measurements due to changes in observing conditions. Furthermore, it can also capture light with longer wavelength than red (i.e., light that is invisible to the human eye and that is absorbed by water vapour in the atmosphere): this "near infrared" light is less subject to being absorbed by interstellar dust, and therefore provides a more precise determination of the objects' brightness than optical light can. For all of these reasons, the Hubble Space Telescope is an ideal instrument for the task of building a precise distance ladder, and since its inception in 2005 the SH0ES programme has been awarded over 1000 orbits of precious (and highly sought-after) Hubble Space Telescope time. When the astronomers of SH0ES reported their latest results in December 2021, their new cosmic distance ladder was the most accurate ever built – and yet it decisively departed from the inverse ladder constructed from the relic radiation at the end of the visible universe.

Over the last four decades, measurements of the Hubble-Lemaître constant have been improving in accuracy, spanning a range between 50 and 90 km/s/Mpc in the early 1980s, to a much narrower spread of plausible values, in the region between 64 and 72 km/s/Mpc in the 2010s. The inverse ladder coming from the relic radiation gave in 2020 a very precise value of 67.4 km/s/Mpc, with a margin of error of only 0.5 km/s/Mpc. But the SH0ES value disagreed in a way that could not be explained by random error.

To obtain their new, highly precise value for the Hubble-Lemaître constant, the SH0ES astronomers built an improved cosmic distance ladder, paying great attention to avoid introducing any source of error when jumping from one rung to the next. For the first rung, they used parallax measurements to determine the distance to stars within the Milky Way, to the Large Magellanic Cloud and to the Andromeda Galaxy; they also measured Cepheid variables in all those regions (including 33 Cepheids in Andromeda, thus retracing Edwin Hubble's steps), and verified that distances obtained with the two methods agreed. For the second rung, they tracked down Cepheids variables in galaxies that had spawned a supernova type Ia in the past 40 years, finding hundreds of Cepheids spread among 37 galaxies that had hosted a supernova. Here, they checked that Cepheids distances matched those from the supernovae. Finally, the third rung took them to larger distances still, where only supernovae could be seen, and observing them with the same telescope and method as those closer by ensured that no errors would be introduced that could offset the more distant rung from the previous one.

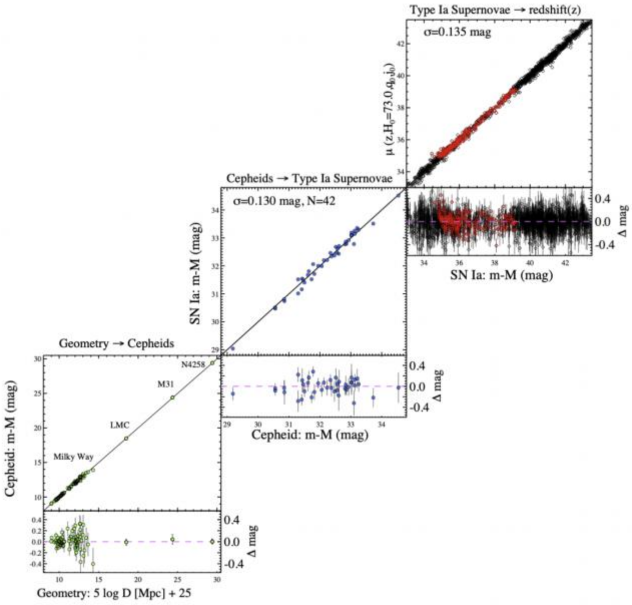

Figure 5: The SH0ES distance ladder combines three rungs (from bottom left to top right): parallax, Cepheids and supernovae type Ia. Source: Riess et al (2021).

Overall, the SH0ES distance ladder reached out to 2 billion light years, and thanks to the careful observations made with the Hubble Space Telescope the margin of uncertainty was greatly reduced. The final result for the Hubble-Lemaître constant came out as 73.04 km/s/Mpc, with a margin of error of 1.04 km/s/Mpc. Now you may think that this value of approximately 73 km/s/Mpc is not a great deal different from the one flowing from the relic radiation, 67 km/s/Mpc – a 10% discrepancy is not the end of the world, considering how arduous the measurements are in the first place. But the greatest efforts went into quantifying not just the most probable value for the Hubble-Lemaître constant, but also its margin of error, about 1 km/s/Mpc for the SH0ES result. This means that[1] there is about 68% probability that the true value of the Hubble-Lemaître constant is within 1 km/s/Mpc either side of 67, so in the range 66 to 68; and there is a probability of about 95% that it is between 65 and 69; a probability of 99.7% that it is between 64 and 70; and about 1 in 3.5 million probability that the true value lies between 62 and 72 km/s/Mpc. And here's the snag: the relic radiation value stands at 73 km/s/Mpc, a very improbable value indeed according to the SH0ES result. What was shocking for cosmologists wasn't the disagreement between the two numbers: rather, the fact that the margin of uncertainty had been reduced so much that it became very improbable that they could both be right.

The conclusion seemed inescapable: either one of the two measurements was in error, or the universe was up to something.

A Cosmic Distance Conundrum

There are only three ways to resolve what has become to be known as the “H-not tension”: either the astronomical distance ladder is somehow in error – despite the extreme care with which it has been built and cross-checked by the SH0ES team; or the relic light measurement is mistaken – again, very difficult to believe, as we have built a detailed understanding of the early universe from where the light comes from. Indeed, in many ways the primordial universe was a much simpler place than the contemporary cosmos, and subject to processes that are much easier to model and predict statistically.

So, if both measurements are right, there remains only one last possibility: our physical model for the expansion of the universe must somehow be in error. It’s conceivable that we are missing some crucial ingredient, for example in the very early times of the life of the universe, that would make the relic light snap into focus even with the larger value of the Hubble-Lemaître constant measured using the SH0ES distance ladder. There are dozens of different theories as to what the missing ingredient might be, but none so far has clearly emerged as a solution.

In the near future, the hope is that a third distance ladder might be built, using a different method than the two we have now, to cross-check which one if any is right. Gravitational waves are one of the exciting possibilities to do so – at long last, we might be able to stretch an unbroken distance ladder, whose first rung was laid out when Hipparchus wondered how far away is the Moon over two millennia ago, all the way to the end of the visible cosmos.

References and Further Reading

Sobel, Dava, The Glass Universe: How the Ladies of the Harvard Observatory Took the Measure of the Stars, Penguin (2016)

Kirshner, Robert P. (2004). Hubble's diagram and cosmic expansion. *Proceedings of the National Academy of Sciences, 101*(1), 8-13. doi:10.1073/pnas.2536799100

Perryman, Michael, The history of astrometry, Eur. Phys. J. H 37 (5) 745-792 (2012)

Teets, Donald A., Transits of Venus and the Astronomical Unit, Mathematics Magazine, Vol 76, 5, Dec. 2003, 335-348

---

[1] A technical point for the sake of accuracy: I am taking here a Bayesian view of probability, interpreting the SH0ES result as a posterior distribution for the Hubble-Lemaître constant. Taking a classical statistics interpretation of the measurements would make the interpretation of the result more contorted, thus needlessly clouding the issue. The essential point stands from either perspective.

Sources of figures:

Leavitt, H. S., & Pickering, E. C. (1912). Periods of 25 Variable Stars in the Small Magellanic Cloud. Harvard College Observatory Circular, 173, 1-3. Retrieved from https://ui.adsabs.harvard.edu/abs/1912HarCi.173....1L

Riess, A. G., Yuan, W., Macri, L. M., Scolnic, D., Brout, D., Casertano, S., . . . Zheng, W. (2021). A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km/s/Mpc Uncertainty from the Hubble Space Telescope and the SH0ES Team. arXiv:2112.04510. Retrieved from https://ui.adsabs.harvard.edu/abs/2021arXiv211204510R

The ESA/ESO astronomy exercice series 2, retrieved from: https://sci.esa.int/documents/34439/36575/1567254558082-4-exercise2-high.pdf

Willach, R. (2004). The Heliometer: Instrument for Gauging Distances in Space. Journal of the Antique Telescope Society, 26, 5-16. Retrieved from https://ui.adsabs.harvard.edu/abs/2004JATSo..26....5W