Cosmic Distance Ladder videos with Grant Sanderson (3blue1brown): commentary and corrections

13 February, 2025 in expository, media | Tags: 3blue1brown, cosmic distance ladder, Grant Sanderson, Tanya Klowden | by Terence Tao

Grant Sanderson (who runs, and creates most of the content for, the website and Youtube channel 3blue1brown) has been collaborating with myself and others (including my coauthor Tanya Klowden) on producing a two-part video giving an account of some of the history of the cosmic distance ladder, building upon a previous public lecture I gave on this topic, and also relating to a forthcoming popular book with Tanya on this topic. The first part of this video is available here; the second part is available here.

The videos were based on a somewhat unscripted interview that Grant conducted with me some months ago, and as such contained some minor inaccuracies and omissions (including some made for editing reasons to keep the overall narrative coherent and within a reasonable length). They also generated many good questions from the viewers of the Youtube video. I am therefore compiling here a "FAQ" of various clarifications and corrections to the videos; this was originally placed as a series of comments on the Youtube channel, but the blog post format here will be easier to maintain going forward. Some related content will also be posted on the Instagram page for the forthcoming book with Tanya.

Questions on the two main videos are marked with an appropriate timestamp to the video.

Comments on part 1 of the video

4:26 **Did Eratosthenes really check a local well in Alexandria?**

This was a narrative embellishment on my part. Eratosthenes's original work is lost to us. The most detailed contemperaneous account, by Cleomedes, gives a simplified version of the method, and makes reference only to sundials (gnomons) rather than wells. However, a secondary account of Pliny states (using this English translation), "Similarly it is reported that at the town of Syene, 5000 stades South of Alexandria, at noon in midsummer no shadow is cast, and that in a well made for the sake of testing this the light reaches to the bottom, clearly showing that the sun is vertically above that place at the time". However, no mention is made of any well in Alexandria in either account.

4:50 **How did Eratosthenes know that the Sun was so far away that its light rays were close to parallel**?

This was not made so clear in our discussions or in the video (other than a brief glimpse of the timeline at 18:27), but Eratosthenes's work actually came after Aristarchus, so it is very likely that Eratosthenes was aware of Aristarchus's conclusions about how distant the Sun was from the Earth. Even if Aristarchus's heliocentric model was disputed by the other Greeks, at least some of his other conclusions appear to have attracted some support. Also, after Eratosthenes's time, there was further work by Greek, Indian, and Islamic astronomers (such as Hipparchus, Ptolemy, Aryabhata, and Al-Battani) to measure the same distances that Aristarchus did, although these subsequent measurements for the Sun also were somewhat far from modern accepted values.

5:17 **Is it completely accurate to say that on the summer solstice, the Earth's axis of rotation is tilted "directly towards the Sun"**?

Strictly speaking, "in the direction towards the Sun" is more accurate than "directly towards the Sun"; it tilts at about 23.5 degrees towards the Sun, but it is not a total 90-degree tilt towards the Sun.

5:39 **Wait, aren't there two tropics? The tropic of Cancer and the tropic of Capricorn?**

Yes! This corresponds to the two summers Earth experiences, one in the Northern hemisphere and one in the

Southern hemisphere. The tropic of Cancer, at a latitude of about 23 degrees north, is where the Sun is directly overhead at noon during the Northern summer solstice (around June 21); the tropic of Capricorn, at a latitude of about 23 degrees south, is where the Sun is directly overhead at noon during the Southern summer solstice (around December 21). But Alexandria and Syene were both in the Northern Hemisphere, so it is the tropic of Cancer that is relevant to Eratosthenes' calculations.

5:41 **Isn't it kind of a massive coincidence that Syene was on the tropic of Cancer?**

Actually, Syene (now known as Aswan) was about half a degree of latitude away from the tropic of Cancer, which was one of the sources of inaccuracy in Eratosthenes' calculations. But one should take the "look-elsewhere effect" into account: because the Nile cuts across the tropic of Cancer, it was quite likely to happen that the Nile would intersect the tropic near *some* inhabited town. It might not necessarily have been Syene, but that would just mean that Syene would have been substituted by this other town in Eratosthenes's account.

On the other hand, it was fortunate that the Nile ran from South to North, so that distances between towns were a good proxy for the differences in latitude. Apparently, Eratosthenes actually had a more complicated argument that would also work if the two towns in question were not necessarily oriented along the North-South direction, and if neither town was on the tropic of Cancer; but unfortunately the original writings of Eratosthenes are lost to us, and we do not know the details of this more general argument. (But some variants of the method can be found in later work of Posidonius, Aryabhata, and others.)

Nowadays, the "Eratosthenes experiment" is run every year on the March equinox, in which schools at the same longitude are paired up to measure the elevation of the Sun at the same point in time, in order to obtain a measurement of the circumference of the Earth. (The equinox is more convenient than the solstice when neither location is on a tropic, due to the simple motion of the Sun at that date.) With modern timekeeping, communications, surveying, and navigation, this is a far easier task to accomplish today than it was in Eratosthenes' time.

6:30 **I thought the Earth wasn't a perfect sphere. Does this affect this calculation?**

Yes, but only by a small amount. The centrifugal forces caused by the Earth's rotation along its axis cause an equatorial bulge and a polar flattening so that the radius of the Earth fluctuates by about 20 kilometers from pole to equator. This sounds like a lot, but it is only about 0.3% of the mean Earth radius of 6371 km and is not the primary source of error in Eratosthenes' calculations.

7:27 **Are the riverboat merchants and the "grad student" the leading theories for how Eratosthenes measured the distance from Alexandria to Syene?**

There is some recent research that suggests that Eratosthenes may have drawn on the work of professional bematists (step measurers – a precursor to the modern profession of surveyor) for this calculation. This somewhat ruins the "grad student" joke, but perhaps should be disclosed for the sake of completeness.

8:51 **How long is a "lunar month" in this context? Is it really 28 days?**

In this context the correct notion of a lunar month is a "synodic month" – the length of a lunar cycle relative to the Sun – which is actually about 29 days and 12 hours. It differs from the "sidereal month" – the length of a lunar cycle relative to the fixed stars – which is about 27 days and 8 hours – due to the motion of the Earth around the Sun (or the Sun around the Earth, in the geocentric model). [A similar correction needs to be made around 14:59, using the synodic month of 29 days and 12 hours rather than the "English lunar month" of 28 days (4 weeks).]

10:47 **Is the time taken for the Moon to complete an observed rotation around the Earth slightly less than 24 hours as claimed?**

Actually, I made a sign error: the lunar day (also known as a tidal day) is actually 24 hours and 50 minutes, because the Moon rotates in the same direction as the spinning of Earth around its axis. The animation therefore is also moving in the wrong direction as well (related to this, the line of sight is covering up the Moon in the wrong direction to the Moon rising at around 10:38).

11:32 **Is this really just a coincidence that the Moon and Sun have almost the same angular width?**

I believe so. First of all, the agreement is not *that* good: due to the non-circular nature of the orbit of the Moon around the Earth, and Earth around the Sun, the angular width of the Moon actually fluctuates to be as much as 10% larger or smaller than the Sun at various times (cf. the "supermoon" phenomenon). All other known planets with known moons do not exhibit this sort of agreement, so there does not appear to be any universal law of nature that would enforce this coincidence. (This is in contrast with the empirical fact that the Moon always presents the same side to the Earth, which occurs in all other known large moons (as well as Pluto), and is well explained by the physical phenomenon of tidal locking.)

On the other hand, as the video hopefully demonstrates, the existence of the Moon was extremely helpful in allowing the ancients to understand the basic nature of the solar system. Without the Moon, their task would have been significantly more difficult; but in this hypothetical alternate universe, it is likely that modern cosmology would have still become possible once advanced technology such as telescopes, spaceflight, and computers became available, especially when combined with the modern mathematics of data science. Without giving away too many spoilers, a scenario similar to this was explored in the classic short story and novel "Nightfall" by Isaac Asimov.

12:58 **Isn't the illuminated portion of the Moon, as well as the visible portion of the Moon, slightly smaller than half of the entire Moon, because the Earth and Sun are not an infinite distance away from the Moon?**

Technically yes (and this is actually for a very similar reason to why half Moons don't quite occur halfway between the new Moon and the full Moon); but this fact turns out to have only a very small effect on the calculations, and is not the major source of error. In reality, the Sun turns out to be about 86,000 Moon radii away from the Moon, so asserting that half of the Moon is illuminated by the Sun is actually a very good first approximation. (The Earth is "only" about 220 Moon radii away, so the visible portion of the Moon is a bit more noticeably less than half; but this doesn't actually affect Aristarchus's arguments much.)

The angular diameter of the Sun also creates an additional thin band between the fully illuminated and fully non-illuminated portions of the Moon, in which the Sun is intersecting the lunar horizon and so only illuminates the Moon with a portion of its light, but this is also a relatively minor effect (and the midpoints of this band can still be used to define the terminator between illuminated and non-illuminated for the purposes of Aristarchus's arguments).

13:27 **What is the difference between a half Moon and a quarter Moon?**

If one divides the lunar month, starting and ending at a new Moon, into quarters (weeks), then half moons occur both near the end of the first quarter (a week after the new Moon, and a week before the full Moon), and near the end of the third quarter (a week after the full Moon, and a week before the new Moon). So, somewhat confusingly, half Moons come in two types, known as "first quarter Moons" and "third quarter Moons".

14:49 **I thought the sine function was introduced well after the ancient Greeks.**

It's true that the modern sine function only dates back to the Indian and Islamic mathematical traditions in the first millennium CE, several centuries after Aristarchus. However, he still had Euclidean geometry at his disposal, which provided tools such as similar triangles that could be used to reach basically the same conclusions, albeit with significantly more effort than would be needed if one could use modern trigonometry.

On the other hand, Aristarchus was somewhat hampered by not knowing an accurate value for π, which is also known as Archimedes' constant: the fundamental work of Archimedes on this constant actually took place a few decades after that of Aristarchus!

15:17 **I plugged in the modern values for the distances to the Sun and Moon and got 18 minutes for the discrepancy, instead of half an hour.**

Yes; I quoted the wrong number here. In 1630, Godfried Wendelen replicated Aristarchus's experiment. With improved timekeeping and the then-recent invention of the telescope, Wendelen obtained a measurement of half an hour for the discrepancy, which is significantly better than Aristarchus's calculation of six hours, but still a little bit off from the true value of 18 minutes. (As such, Wendelinus's estimate for the distance to the Sun was 60% of the true value.)

15:27 **Wouldn't Aristarchus also have access to other timekeeping devices than sundials?**

Yes, for instance clepsydrae (water clocks) were available by that time; but they were of limited accuracy. It is also possible that Aristarchus could have used measurements of star elevations to also estimate time; it is not clear whether the astrolabe or the armillary sphere was available to him, but he would have had some other more primitive astronomical instruments such as the dioptra at his disposal. But again, the accuracy and calibration of these timekeeping tools would have been poor.

However, most likely the more important limiting factor was the ability to determine the precise moment at which a perfect half Moon (or new Moon, or full Moon) occurs; this is extremely difficult to do with the naked eye. (The telescope would not be invented for almost two more millennia.)

17:37 **Could the parallax problem be solved by assuming that the stars are not distributed in a three-dimensional space, but instead on a celestial sphere?**

Putting all the stars on a fixed sphere would make the parallax effects less visible, as the stars in a given portion of the sky would now all move together at the same apparent velocity – but there would still be visible large-scale distortions in the shape of the constellations because the Earth would be closer to some portions of the celestial sphere than others; there would also be variability in the brightness of the stars, and (if they were very close) the apparent angular diameter of the stars. (These problems would be solved if the celestial sphere was somehow centered around the moving Earth rather than the fixed Sun, but then this basically becomes the geocentric model with extra steps.)

18:29 **Did nothing of note happen in astronomy between Eratosthenes and Copernicus?**

Not at all! There were significant mathematical, technological, theoretical, and observational advances by astronomers from many cultures (Greek, Islamic, Indian, Chinese, European, and others) during this time, for instance improving some of the previous measurements on the distance ladder, a better understanding of eclipses, axial tilt, and even axial precession, more sophisticated trigonometry, and the development of new astronomical tools such as the astrolabe. See for instance this "deleted scene" from the video, as well as the FAQ entry for 14:49 for this video and 24:54 for the second video, or this instagram post. But in order to make the overall story of the cosmic distance ladder fit into a two-part video, we chose to focus primarily on the first time each rung of the ladder was climbed.

18:30 **Is that really Kepler's portrait?**

We have since learned that this portrait was most likely painted in the 19th century, and may have been based more on Kepler's mentor, Michael Mästlin. A more commonly accepted portrait of Kepler may be found at his current Wikipedia page.

19:07 **Isn't it tautological to say that the Earth takes one year to perform a full orbit around the Sun?**

Technically yes, but this is an illustration of the philosophical concept of "referential opacity": the content of a sentence can change when substituting one term for another (e.g., "1 year" and "365 days"), even when both terms refer to the same object. Amusingly, the classic illustration of this, known as Frege's puzzles, also comes from astronomy: it is an informative statement that Hesperus (the evening star) and Phosphorus (the morning star, also known as Lucifer) are the same object (which nowadays we call Venus), but it is a mere tautology that Hesperus and Hesperus are the same object: changing the reference from Phosphorus to Hesperus changes the meaning.

19:10 **How did Copernicus figure out the crucial fact that Mars takes 687 days to go around the Sun? Was it directly drawn from Babylonian data?**

Technically, Copernicus drew from tables by European astronomers that were largely based on earlier tables from the Islamic golden age, which in turn drew from earlier tables by Indian and Greek astronomers, the latter of which also incorporated data from the ancient Babylonians, so it is more accurate to say that Copernicus relied on centuries of data, at least some of which went all the way back to the Babylonians. Among all of this data was the times when Mars was in opposition to the Sun; if one imagines the Earth and Mars as being like runners going around a race track circling the Sun, with Earth on an inner track and Mars on an outer track, oppositions are analogous to when the Earth runner "laps" the Mars runner. From the centuries of observational data, such "laps" were known to occur about once every 780 days (this is known as the synodic period of Mars). Because the Earth takes roughly 365 days to perform a "lap", it is possible to do a little math and conclude that Mars must therefore complete its own "lap" in 687 days (this is known as the sidereal period of Mars). (See also this post on the cosmic distance ladder Instagram for some further elaboration.)

20:52 **Did Kepler really steal data from Brahe?**

The situation is complex. When Kepler served as Brahe's assistant, Brahe only provided Kepler with a limited amount of data, primarily involving Mars, in order to confirm Brahe's own geo-heliocentric model. After Brahe's death, the data was inherited by Brahe's son-in-law and other relatives, who intended to publish Brahe's work separately; however, Kepler, who was appointed as Imperial Mathematician to succeed Brahe, had at least some partial access to the data, and many historians believe he secretly copied portions of this data to aid his own research before finally securing complete access to the data from Brahe's heirs after several years of disputes. On the other hand, as intellectual property rights laws were not well developed at this time, Kepler's actions were technically legal, if ethically questionable.

21:39 **What is that funny loop in the orbit of Mars?**

This is known as retrograde motion. This arises because the orbital velocity of Earth (about 30 km/sec) is a little bit larger than that of Mars (about 24 km/sec). So, in opposition (when Mars is in the opposite position in the sky than the Sun), Earth will briefly overtake Mars, causing its observed position to move westward rather than eastward. But in most other times, the motion of Earth and Mars are at a sufficient angle that Mars will continue its apparent eastward motion despite the slightly faster speed of the Earth.

21:59 **Couldn't one also work out the direction to other celestial objects in addition to the Sun and Mars, such as the stars, the Moon, or the other planets? Would that have helped?**

Actually, the directions to the fixed stars were implicitly used in all of these observations to determine how the celestial sphere was positioned, and all the other directions were taken relative to that celestial sphere. (Otherwise, all the calculations would be taken on a rotating frame of reference in which the unknown orbits of the planets were themselves rotating, which would have been an even more complex task.) But the stars are too far away to be useful as one of the two landmarks to triangulate from, as they generate almost no parallax and so cannot distinguish one location from another.

Measuring the direction to the Moon would tell you which portion of the lunar cycle one was in, and would determine the phase of the Moon, but this information would not help one triangulate, because the Moon's position in the heliocentric model varies over time in a somewhat complicated fashion, and is too tied to the motion of the Earth to be a useful "landmark" to one to determine the Earth's orbit around the Sun.

In principle, using the measurements to all the planets at once could allow for some multidimensional analysis that would be more accurate than analyzing each of the planets separately, but this would require some sophisticated statistical analysis and modeling, as well as non-trivial amounts of compute – neither of which were available in Kepler's time.

22:57 **Can you elaborate on how we know that the planets all move on a plane?**

The Earth's orbit lies in a plane known as the ecliptic (it is where the lunar and solar eclipses occur). Different cultures have divided up the ecliptic in various ways; in Western astrology, for instance, the twelve main constellations that cross the ecliptic are known as the Zodiac. The planets can be observed to only wander along the Zodiac, but not other constellations: for instance, Mars can be observed to be in Cancer or Libra, but never in Orion or Ursa Major. From this, one can conclude (as a first approximation, at least), that the planets all lie on the ecliptic.

However, this isn't perfectly true, and the planets will deviate from the ecliptic by a small angle known as the ecliptic latitude. Tycho Brahe's observations on these latitudes for Mars were an additional useful piece of data that helped Kepler complete his calculations (basically by suggesting how to join together the different "jigsaw pieces"), but the math here gets somewhat complicated, so the story here has been somewhat simplified to convey the main ideas.

23:04 **What are the other universal problem solving tips?**

Grant Sanderson has a list (in a somewhat different order) in this previous video.

23:28 **Can one work out the position of Earth from fixed locations of the Sun and Mars when the Sun and Mars are in conjunction (the same location in the sky) or opposition (opposite locations in the sky)?**

Technically, these are two times when the technique of triangulation fails to be accurate; and also in the former case it is extremely difficult to observe Mars due to the proximity to the Sun. But again, following the Universal Problem Solving Tip from 23:07, one should initially ignore these difficulties to locate a viable method, and correct for these issues later. This video series by Welch Labs goes into Kepler's methods in more detail.

24:04 **So Kepler used Copernicus's calculation of 687 days for the period of Mars. But didn't Kepler discard Copernicus's theory of circular orbits?**

Good question! It turns out that Copernicus's calculations of orbital periods are quite robust (especially with centuries of data), and continue to work even when the orbits are not perfectly circular. But even if the calculations did depend on the circular orbit hypothesis, it would have been possible to use the Copernican model as a first approximation for the period, in order to get a better, but still approximate, description of the orbits of the planets. This in turn can be fed back into the Copernican calculations to give a second approximation to the period, which can then give a further refinement of the orbits. Thanks to the branch of mathematics known as perturbation theory, one can often make this type of iterative process converge to an exact answer, with the error in each successive approximation being smaller than the previous one. (But performing such an iteration would probably have been beyond the computational resources available in Kepler's time; also, the foundations of perturbation theory require calculus, which only was developed several decades after Kepler.)

24:21 **Did Brahe have exactly 10 years of data on Mars's positions?**

Actually, it was more like 17 years, but with many gaps, due both to inclement weather, as well as Brahe turning his attention to other astronomical objects than Mars in some years; also, in times of conjunction, Mars might only be visible in the daytime sky instead of the night sky, again complicating measurements. So the "jigsaw puzzle pieces" in 25:26 are in fact more complicated than always just five locations equally spaced in time; there are gaps and also observational errors to grapple with. But to understand the method one should ignore these complications; again, see "Universal Problem Solving Tip #1". Even with his "idea of true genius", it took many years of further painstaking calculation for Kepler to tease out his laws of planetary motion from Brahe's messy and incomplete observational data.

26:44 **Shouldn't the Earth's orbit be spread out at perihelion and clustered closer together at aphelion, to be consistent with Kepler's laws?**

Yes, you are right; there was a coding error here.

26:53 **What is the reference for Einstein's "idea of pure genius"?**

Actually, the precise quote was "an idea of true genius", and can be found in the introduction to Carola Baumgardt's "Life of Kepler".

Comments on the "deleted scene" on Al-Biruni

Was Al-Biruni really of Arab origin?

Strictly speaking; no; his writings are all in Arabic, and he was nominally a subject of the Abbasid Caliphate whose rulers were Arab; but he was born in Khwarazm (in modern day Uzbekistan), and would have been a subject of either the Samanid empire or the Khrawazmian empire, both of which were largely self-governed and primarily Persian in culture and ethnic makeup, despite being technically vassals of the Caliphate. So he would have been part of what is sometimes called "Greater Persia" or "Greater Iran".

Another minor correction: while Al-Biruni was born in the tenth century, his work on the measurement of the Earth was published in the early eleventh century.

Is θ really called the angle of declination?

This was a misnomer on my part; this angle is more commonly called the dip angle.

But the height of the mountain would be so small compared to the radius of the Earth! How could this method work?

Using the Taylor approximation $\cos\theta \approx 1 - \theta^2/2$, one can approximately write the relationship $R = \frac{h\cos\theta}{1-\cos\theta}$ between the mountain height h, the Earth radius R, and the dip angle θ (in radians) as $R \approx 2h/\theta^2$. The key point here is the inverse quadratic dependence on θ, which allows for even relatively small values of h to still be realistically useful for computing R. Al-Biruni's measurement of the dip angle θ was about 0.01 radians, leading to an estimate of R that is about four orders of magnitude larger than h, which is within ballpark at least of a typical height of a mountain (on the order of a kilometer) and the radius of the Earth (6400 kilometers).

Was the method really accurate to within a percentage point?

This is disputed, somewhat similarly to the previous calculations of Eratosthenes. Al-Biruni's measurements were in cubits, but there were multiple incompatible types of cubit in use at the time. It has also been pointed out that atmospheric refraction effects would have created noticeable changes in the observed dip angle θ. It is thus likely that the true accuracy of Al-Biruni's method was poorer than 1%, but that this was somehow compensated for by choosing a favorable conversion between cubits and modern units.

Comments on the second part of the video

1:13 **Did Captain Cook set out to discover Australia?**

One of the objectives of Cook's first voyage was to discover the hypothetical continent of Terra Australis. This was considered to be distinct from Australia, which at the time was known as New Holland. As this name might suggest, prior to Cook's voyage, the northwest coastline of New Holland had been explored by the Dutch; Cook instead explored the eastern coastline, naming this portion New South Wales. The entire continent was later renamed to Australia by the British government, following a suggestion of Matthew Flinders; and the concept of Terra Australis was abandoned.

4:40 **The relative position of the Northern and Southern hemisphere observations is reversed from those earlier in the video.**

Yes, this was a slight error in the animation; the labels here should be swapped for consistency of orientation.

7:06 **So, when did they finally manage to measure the transit of Venus, and use this to compute the astronomical unit?**

While Le Gentil had the misfortune to not be able to measure either the 1761 or 1769 transits, other expeditions of astronomers (led by Dixon-Mason, Chappe d'Auteroche, and Cook) did take measurements of one or both of these transits with varying degrees of success, with the measurements of Cook's team of the 1769 transit in Tahiti being of particularly high quality. All of this data was assembled later by Lalande in 1771, leading to the most accurate measurement of the astronomical unit at the time (within 2.3% of modern values, which was about three times more accurate than any previous measurement).

8:53 **What does it mean for the transit of Io to be "twenty minutes ahead of schedule" when Jupiter is in opposition (Jupiter is opposite to the Sun when viewed from the Earth)?**

Actually, it should be halved to "ten minutes ahead of schedule", with the transit being "ten minutes behind schedule" when Jupiter is in conjunction, with the net discrepancy being twenty minutes (or actually closer to 16 minutes when measured with modern technology). Both transits are being compared against an idealized periodic schedule in which the transits are occuring at a perfectly regular rate (about 42 hours), where the period is chosen to be the best fit to the actual data. This discrepancy is only noticeable after carefully comparing transit times over a period of months; at any given position of Jupiter, the Doppler effects of Earth moving towards or away from Jupiter would only affect shift each transit by just a few seconds compared to the previous transit, with the delays or accelerations only becoming cumulatively noticeable after many such transits.

Also, the presentation here is oversimplified: at times of conjunction, Jupiter and Io are too close to the Sun for observation of the transit. Rømer actually observed the transits at other times than conjunction, and Huygens used more complicated trigonometry than what was presented here to infer a measurement for the speed of light in terms of the astronomical unit (which they had begun to measure a bit more accurately than in Aristarchus's time; see the FAQ entry for 15:17 in the first video).

10:05 **Are the astrological signs for Earth and Venus swapped here?**

Yes, this was a small mistake in the animation.

10:34 **Shouldn't one have to account for the elliptical orbit of the Earth, as well as the proper motion of the star being observed, or the effects of general relativity?**

Yes; the presentation given here is a simplified one to convey the idea of the method, but in the most advanced parallax measurements, such as the ones taken by the Hipparcos and Gaia spacecraft, these factors are taken into account, basically by taking as many measurements (not just two) as possible of a single star, and locating the best fit of that data to a multi-parameter model that incorporates the (known) orbit of the Earth with the (unknown) distance and motion of the star, as well as additional gravitational effects from other celestial bodies, such as the Sun and other planets.

14:53 **The formula I was taught for apparent magnitude of stars looks a bit different from the one here.**

This is because astronomers use a logarithmic scale to measure both apparent magnitude m and absolute magnitude M. If one takes the logarithm of the inverse square law in the video, and performs the normalizations used by astronomers to define magnitude, one arrives at the standard relation $M = m - 5\log_{10} d_{pc} + 5$ between absolute and apparent magnitude.

But this is an oversimplification, most notably due to neglect of the effects of extinction effects caused by interstellar dust. This is not a major issue for the relatively short distances observable via parallax, but causes problems at larger scales of the ladder (see for instance the FAQ entry here for 18:08). To compensate for this, one can work in multiple frequencies of the spectrum (visible, x-ray, radio, etc.), as some frequencies are less susceptible to extinction than others. From the discrepancies between these frequencies one can infer the amount of extinction, leading to "dust maps" that can then be used to facilitate such corrections for subsequent measurements in the same area of the universe. (More generally, the trend in modern astronomy is towards "multi-messenger astronomy" in which one combines together very different types of measurements of the same object to obtain a more accurate understanding of that object and its surroundings.)

18:08 **Can we really measure the entire Milky Way with this method?**

Strictly speaking, there is a "zone of avoidance" on the far side of the Milky way that is very difficult to measure in the visible portion of the spectrum, due to the large amount of intervening stars, dust, and even a supermassive black hole in the galactic center. However, in recent years it has become possible to explore this zone to some extent using the radio, infrared, and x-ray portions of the spectrum, which are less affected by these factors.

18:19 **How did astronomers know that the Milky Way was only a small portion of the entire universe?**

This issue was the topic of the "Great Debate" in the early twentieth century. It was only with the work of Hubble using Leavitt's law to measure distances to Magellanic clouds and "spiral nebulae" (that we now know to be other galaxies), building on earlier work of Leavitt and Hertzsprung, that it was conclusively established that these clouds and nebulae in fact were at much greater distances than the diameter of the Milky Way.

18:45 **How can one compensate for light blending effects when measuring the apparent magnitude of Cepheids?**

This is a non-trivial task, especially if one demands a high level of accuracy. Using the highest resolution telescopes available (such as HST or JWST) is of course helpful, as is switching to other frequencies, such as near-infrared, where Cepheids are even brighter relative to nearby non-Cepheid stars. One can also apply sophisticated statistical methods to fit to models of the point spread of light from unwanted sources, and use nearby measurements of the same galaxy without the Cepheid as a reference to help calibrate those models. Improving the accuracy of the Cepheid portion of the distance ladder is an ongoing research activity in modern astronomy.

18:54 **What is the mechanism that causes Cepheids to oscillate?**

For most stars, there is an equilibrium size: if the star's radius collapses, then the reduced potential energy is converted to heat, creating pressure to pushing the star outward again; and conversely, if the star expands, then it cools, causing a reduction in pressure that no longer counteracts gravitational forces. But for Cepheids, there is an additional mechanism called the kappa mechanism: the increased temperature caused by contraction increases ionization of helium, which drains energy from the star and accelerates the contraction; conversely, the cooling caused by expansion causes the ionized helium to recombine, with the energy released accelerating the expansion. If the parameters of the Cepheid are in a certain "instability strip", then the interaction of the kappa mechanism with the other mechanisms of stellar dynamics create a periodic oscillation in the Cepheid's radius, which increases with the mass and brightness of the Cepheid.

For a recent re-analysis of Leavitt's original Cepheid data, see this paper.

19:10 **Did Leavitt mainly study the Cepheids in our own galaxy?**

This was an inaccuracy in the presentation. Leavitt's original breakthrough paper studied Cepheids in the Small Magellanic Cloud. At the time, the distance to this cloud was not known; indeed, it was a matter of debate whether this cloud was in the Milky Way, or some distance away from it. However, Leavitt (correctly) assumed that all the Cepheids in this cloud were roughly the same distance away from our solar system, so that the apparent brightness was proportional to the absolute brightness. This gave an uncalibrated form of Leavitt's law between absolute brightness and period, subject to the (then unknown) distance to the Small Magellanic Cloud. After Leavitt's work, there were several efforts (by Hertzsprung, Russell, and Shapley) to calibrate the law by using the few Cepheids for which other distance methods were available, such as parallax. (Main sequence fitting to the Hertzsprung-Russell diagram was not directly usable, as Cepheids did not lie on the main sequence; but in some cases one could indirectly use this method if the Cepheid was in the same stellar cluster as a main sequence star.) Once the law was calibrated, it could be used to measure distances to other Cepheids, and in particular to compute distances to extragalactic objects such as the Magellanic clouds.

19:15 **Was Leavitt's law really a linear law between period and luminosity?**

Strictly speaking, the period-luminosity relation commonly known as Leavitt's law was a linear relation between the absolute magnitude of the Cepheid and the *logarithm* of the period; undoing the logarithms, this becomes a power law between the luminosity and the period.

20:26 **Was Hubble the one to discover the redshift of galaxies?**

This was an error on my part; Hubble was using earlier work of Vesto Slipher on these redshifts, and combining it with his own measurements of distances using Leavitt's law to arrive at the law that now bears his name; he was also assisted in his observations by Milton Humason. It should also be noted that Georges Lemaître had also independently arrived at essentially the same law a few years prior, but his work was published in a somewhat obscure journal and did not receive broad recognition until some time later.

20:37 **Hubble's original graph doesn't look like a very good fit to a linear law.**

Hubble's original data was somewhat noisy and inaccurate by modern standards, and the redshifts were affected by the peculiar velocities of individual galaxies in addition to the expanding nature of the universe. However, as the data was extended to more galaxies, it became increasingly possible to compensate for these effects and obtain a much tighter fit, particularly at larger scales where the effects of peculiar velocity are less significant. See for instance this article from 2015 where Hubble's original graph is compared with a more modern graph. This more recent graph also reveals a slight nonlinear correction to Hubble's law at very large scales that has led to the remarkable discovery that the expansion of the universe is in fact accelerating over time, a phenomenon that is attributed to a positive cosmological constant (or perhaps a more complex form of dark energy in the universe). On the other hand, even with this nonlinear correction, there continues to be a roughly 10% discrepancy of this law with predictions based primarily on the cosmic microwave background radiation; see the FAQ entry for 23:49.

20:46 **Does general relativity alone predict an uniformly expanding universe?**

This was an oversimplification. Einstein's equations of general relativity contain a parameter Λ, known as the cosmological constant, which currently is only computable indirectly from fitting to experimental data. But even with this constant fixed, there are multiple solutions to these equations (basically because there are multiple possible initial conditions for the universe). For the purposes of cosmology, a particularly successful family of solutions are the solutions given by the Lambda-CDM model. This family of solutions contains additional parameters, such as the density of dark matter in the universe. Depending on the precise values of these parameters, the universe could be expanding or contracting, with the rate of expansion or contraction either increasing, decreasing, or staying roughly constant. But if one fits this model to all available data (including not just red shift measurements, but also measurements on the cosmic microwave background radiation and the spatial distribution of galaxies), one deduces a version of Hubble's law which is nearly linear, but with an additional correction at very large scales; see the next item of this FAQ.

21:07 **Is Hubble's original law sufficiently accurate to allow for good measurements of distances at the scale of the observable universe?**

Not really; as mentioned in the end of the video, there were additional efforts to cross-check and calibrate Hubble's law at intermediate scales between the range of Cepheid methods (about 100 million light years) and observable universe scales (about 100 billion light years) by using further "standard candles" than Cepheids, most notably Type Ia supernovae (which are bright enough and predictable enough to be usable out to about 10 billion light years), the Tully-Fisher relation between the luminosity of a galaxy and its rotational speed, and gamma ray bursts. It turns out that due to the accelerating nature of the universe's expansion, Hubble's law is not completely linear at these large scales; this important correction cannot be discerned purely from Cepheid data, but also requires the other standard candles, as well as fitting that data (as well as other observational data, such as the cosmic microwave background radiation) to the cosmological models provided by general relativity (with the best fitting models to date being some version of the Lambda-CDM model).

On the other hand, a naive linear extrapolation of Hubble's original law to all larger scales does provide a very rough picture of the observable universe which, while too inaccurate for cutting edge research in astronomy, does give some general idea of its large-scale structure.

21:15 **Where did this guess of the observable universe being about 20% of the full universe come from?**

There are some ways to get a lower bound on the size of the entire universe that go beyond the edge of the observable universe. One is through analysis of the cosmic microwave background radiation (CMB), that has been carefully mapped out by several satellite observatories, most notably WMAP and Planck. Roughly speaking, a universe that was less than twice the size of the observable universe would create certain periodicities in the CMB data; such periodicities are not observed, so this provides a lower bound (see for instance this paper for an example of such a calculation). The 20% number was a guess based on my vague recollection of these works, but there is no consensus currently on what the ratio truly is; there are some proposals that the entire universe is in fact several orders of magnitude larger than the observable one.

The situation is somewhat analogous to Aristarchus's measurement of the distance to the Sun, which was very sensitive to a small angle (the half-moon discrepancy). Here, the predicted size of the universe under the standard cosmological model is similarly dependent in a highly sensitive fashion on a measure Ω_k of the flatness of the universe which, for reasons still not fully understood (but likely caused by some sort of inflation mechanism), happens to be extremely close to zero. As such, predictions for the size of the universe remain highly volatile at the current level of measurement accuracy.

23:44 **Was it a black hole collision that allowed for an independent measurement of Hubble's law?**

This was a slight error in the presentation. While the first gravitational wave observation by LIGO in 2015 was of a black hole collision, it did not come with an electromagnetic counterpart that allowed for a redshift calculation that would yield a Hubble's law measurement. However, a later collision of neutron stars, observed in 2017, did come with an associated kilonova in which a redshift was calculated, and led to a Hubble measurement which was independent of most of the rungs of the distance ladder.

23:49 **Where can I learn more about this 10% discrepancy in Hubble's law?**

This is known as the Hubble tension (or, in more sensational media, the "crisis in cosmology"): roughly speaking, the various measurements of Hubble's constant (either from climbing the cosmic distance ladder, or by fitting various observational data to standard cosmological models) tend to arrive at one of two values, that are about 10% apart from each other. The values based on gravitational wave observations are currently consistent with both values, due to significant error bars in this extremely sensitive method; but other more mature methods are now of sufficient accuracy that they are basically only consistent with one of the two values. Currently there is no consensus on the origin of this tension: possibilities include systemic biases in the observational data, subtle statistical issues with the methodology used to interpret the data, a correction to the standard cosmological model, the influence of some previously undiscovered law of physics, or some partial breakdown of the Copernican principle.

For an accessible recent summary of the situation, see this video by Becky Smethurst ("Dr. Becky").

24:49 **So, what is a Type Ia supernova and why is it so useful in the distance ladder?**

A Type Ia supernova occurs when a white dwarf in a binary system draws more and more mass from its companion star, until it reaches the Chandrasekhar limit, at which point its gravitational forces are strong enough to cause a collapse that increases the pressure to the point where a supernova is triggered via a process known as carbon detonation. Because of the universal nature of the Chandrasekhar limit, all such supernovae have (as a first approximation) the same absolute brightness and can thus be used as standard candles in a similar fashion to Cepheids (but without the need to first measure any auxiliary observable, such as a period). But these supernovae are also far brighter than Cepheids, and can so this method can be used at significantly larger distances than the Cepheid method (roughly speaking it can handle distances of ~100 billion light years, whereas Cepheids are reliable out to ~10 billion light years). Among other things, the supernovae measurements were the key to detecting an important nonlinear correction to Hubble's law at these scales, leading to the remarkable conclusion that the expansion of the universe is in fact accelerating over time, which in the Lambda-CDM model corresponds to a positive cosmological constant, though there are more complex "dark energy" models that are also proposed to explain this acceleration.

24:54 **Besides Type Ia supernovae, I felt that a lot of other topics relevant to the modern distance ladder (e.g., the cosmic microwave background radiation, the Lambda CDM model, dark matter, dark energy, inflation, multi-messenger astronomy, etc.) were omitted.**

This is partly due to time constraints, and the need for editing to tighten the narrative, but was also a conscious decision on my part. Advanced classes on the distance ladder will naturally focus on the most modern, sophisticated, and precise ways to measure distances, backed up by the latest mathematics, physics, technology, observational data, and cosmological models. However, the focus in this video series was rather different; we sought to portray the cosmic distance ladder as evolving in a fully synergestic way, across many historical eras, with the evolution of mathematics, science, and technology, as opposed to being a mere byproduct of the current state of these other disciplines. As one specific consequence of this change of focus, we emphasized the *first* time any rung of the distance ladder was achieved, at the expense of more accurate and sophisticated later measurements at that rung. For instance, refinements in the measurement of the radius of the Earth since Eratosthenes, improvements in the measurement of the astronomical unit between Aristarchus and Cook, or the refinements of Hubble's law and the cosmological model of the universe in the twentieth and twenty-first centuries, were largely omitted (though some of the answers in this FAQ are intended to address these omissions).

Many of the topics not covered here (or only given a simplified treatment) are discussed in depth in other expositions, including other Youtube videos. I would welcome suggestions from readers for links to such resources in the comments to this post. Here is a partial list:

"Eratosthenes" – Cosmos (Carl Sagan), video posted Apr 24, 2009 (originally released Oct 1, 1980, as part of the episode "The Shores of the Cosmic Ocean").

"How Far Away Is It" – David Butler, a multi-part series beginning Aug 16 2013.

"How the Bizarre Path of Mars Reshaped Astronomy [Kepler's Laws Part 1]", Welch Labs, May 8, 2024. See also Part 2.

"An ASTROPHYSICIST'S TOP 5 space news stories of 2024", Becky Smethurst (Dr. Becky), Dec 26, 2024 – covers the Hubble tension as one of the stories.

"Measuring the Earth… from a vacation photo", George Lowther (Almost sure), Feb 22 2025.

"How Did This Ancient Genius Measure The Sun?", Ben Syversen, Feb 28 2025.

90 comments

Comments feed for this article

13 February, 2025 at 1:55 pm
Anonymous

I really enjoyed the video and I found that the animations are very helpful. I am eager to watch the second part : thank you !

30 3 Rate This

Reply

13 February, 2025 at 4:14 pm
Anonymous

What is remarkable in the parallel rays thought is the embedding of corpuscular theory of light in ancient Greek thought. Without an ultimate philosophical underpinning of the nature of light parallel rays theory of light can be assumed. Do you know the philosophical reasonings in Greek thought behind the assumption of parallel rays theory of light without any physics experiments? It seems to be different from other ancient civilizations view. It might be argued that the origin of the quest to identify distances themselves were a way to justify the corpuscular philosophy. Is there any evidence for this?

Also other civilizations had period of orbits of planets in their astrological beliefs long before the middle ages without knowledge of distances. However the video states that period of Mars was known only in the modern European era post 16th century. Do you know other civilizations had guessed period of orbits of planets integrated into their astrological works with or without an heliocentric view?

5 2 Rate This

Reply

13 February, 2025 at 5:43 pm
Terence Tao

Euclid's second best known book *Optics* postulates that light travels in straight lines, which is the most important axiom needed for this analysis, but does get other fundamental facts about light wrong, in particular subscribing to the emission theory of light rather than an intromission theory, with the debate not being conclusively settled until the Islamic golden age. Still, the idea of starting with an axiomatic approach at all, however flawed, seemed to have been somewhat unique to the Greek school in this era.

It is the sidereal period of Mars (687 days) that was first computed correctly by Copernicus, as even the very concept of this period for any planet other than the Earth basically requires the heliocentric model to begin with. But the synodic period of Mars (780 days) was significantly easier to discern from observational data, and seems to have been known to reasonable accuracy even to the ancient Babylonians.

13 2 Rate This

Reply

14 February, 2025 at 9:26 am
Anonymous

I think you might be wrong with Copernicus being the first. Check https://en.wikipedia.org/wiki/Surya_Siddhanta and https://en.wikipedia.org/wiki/Aryabhata. If there were Greek influences, 1) they should be after 350BC and 2) there should be transfer of knowledge of astronomical distances (which does not seem to be the case) and lesser errors in the Surya Siddhanta for measurement of sizes of astronomical objects (which I am not sure is the case). Tilts are important in astrology and it is not clear how the measurements were achieved.

5 3 Rate This

Reply

---

14 February, 2025 at 11:23 am

Terence Tao

Interesting! It appears that these works did contain several calculations that are closely related to Copernicus's computation of sidereal periods of planets in a heliocentric model (such as the computation of extremely long periods ("yuga") for the solar system as a whole, on the order of millions of years), but they were working instead in a geocentric model (with Ptolemaic-style epicycles), and it takes a certain amount of mathematical conversion between the models to extract, for instance, the 687 day sidereal period of Mars from those works, though it can be done. (Most likely some of the Islamic-era tables that Copernicus was drawing from to perform the same calculations included some input from these Indian sources, in addition to the older Greek and Babylonian sources.)

In principle, the sidereal period of (say) Mars could even be extracted purely from Babylonian data (albeit with lower accuracy) once one had the heliocentric model in which all planets moved in periodic orbits around the Sun, but I have not found a source prior to Copernicus where this extraction was explicitly performed within that source (as opposed to in a modern re-interpretation of that source). These sidereal periods were "hidden in plain sight" in the data for millennia, but it required the perspective of the heliocentric model to bring them to prominence.

12 1 Rate This

Reply

---

14 February, 2025 at 12:58 pm

Anonymous

Also note wiki says Ptolemy (Millennium before Copernicus) knew (https://en.wikipedia.org/wiki/Surya_Siddhanta#Importance_in_history_of_science) the sidereal period of Mars before Hindu-Arabic decimals were provided and before sophisticated Trigonometry was available. But Ptolemy lived after Alexander left what was India then for 100s of years. So it is possible Surya Siddhanta is independent of Ptolemy. But it is not clear how two independent teams arrived at similar mechanisms of mathematical techniques in the pre-internet era. We do not know if Aryabhatta himself (unlikely) made all the calculations of orbits etc that were known to Ptolemy (likely built on the shoulder of Greek giants) but were derived from even earlier (possibly) independent calculations. We do know Greek planetary God names and Hindu God names had common roots. But the position of the Gods in their relative hierarchy changes in Hinduism as it evolved (For example post 1000AD we cannot find temples of Indra (who can be compared with Zeus in Greek culture) but temples in Thailand continued to be maintained.. although it is now sensible to compare Zeus with Brihaspati (Jupiter) who is significant in Hindu astrology).

2 2 Rate This

---

14 February, 2025 at 3:23 pm

Terence Tao

The numbers in the table of implied sidereal periods given in the Wikipedia article come from the 1935 work of Burgess (see pages 27-28 of this link) in which he translates the models of Ptolemy and the Siddhanta into a modern, post-Copernican form, for instance correcting for Ptolemy's incorrect value of the motion of the equinox to infer a mean sidereal angular motion for each planet, and in the case of the Siddhanta, deducing the sidereal period of the planets by fitting the Copernican heliocentric model to the data actually provided in the Siddhanta, namely the length of a yuga and the number of conjunctions of each planet within that yuga (see page 17). However, these sidereal periods are not explicit in either reference; they are only extracted in a post-Copernican analysis from the modern era. To quote from page 17 of the Burgess reference: "But the apparent motions of the planets are greatly complicated by the fact, unknown to the Greek and the Hindu, that they are revolving around a center about which the earth is also revolving."

As mentioned in my previous comment, the raw data and computation necessary to compute the sidereal period of Mars and the other planets (with a low level of accuracy) was already available in principle even to the ancient Babylonians, who possessed enough mathematical skill to verify for instance that $1/365 - 1/780$ was very close to $1/689$ (though of course they would not have used Hindu-Arabic numerals for this calculation). What was missing though was the heliocentric perspective that elevated the fundamental importance of these sidereal periods above more observationally evident quantities such as the synodic periods of the planets (or the number of conjunctions in a fixed period of time such as a yuga).

9 1 Rate This

---

14 February, 2025 at 3:31 pm
Anonymous

"As quoted on page 17 of the Burgess reference: "But the apparent motions of the planets are greatly complicated by the fact, unknown to the Greek and the Hindu, that they are revolving around a center about which the earth is also revolving."

I am not sure of this. Because all the temples which has an abode for the Navagrahas (Nine planets) have the abode constructed with Sun as the center. I am unable to find when the worship started and how old it is. But it is reasonable to say the Cholas knew of it (https://en.wikipedia.org/wiki/Navagraha_temples#Navagraha_temples_in_Tamil_Nadu) and Cholas peaked well before European renaissance (https://en.wikipedia.org/wiki/Chola_Empire). All one needs for evidence is a picture from a temple from the Chola dynasty of such a Navagraha worship abode but I am unable to find this on the internet.

0 4 Rate This

---

14 February, 2025 at 3:57 pm
Terence Tao

The Navagrahas omits the Earth – an absolutely crucial component of any heliocentric model – and also includes two "shadow planets" to generate eclipses, which are not present in the heliocentric model. In particular, the most important insight of the heliocentric model – namely, that the Earth has much the same status as the other planets in the Solar System – is not indicated at all from these displays; if anything, the absence of the Earth only serves to reinforce a geocentric perspective. Given that the major astronomical texts in this period such as the Siddhanta explicitly adopt a geocentric model, I would not view these religious arrangements as strong evidence that geocentrism was anything other than the norm in the Indian astronomy of the period.

15 1 Rate This

---

14 February, 2025 at 3:36 pm
Anonymous

This should suffice https://en.wikipedia.org/wiki/Kapaleeshwarar_Temple as there is a Navagraha worship abode but I cannot find pictures on internet.

0 6 Rate This

14 February, 2025 at 6:16 pm
Anonymous

Thank you that clarifies a lot. https://www.quora.com/Who-proposed-heliocentrism-first-Aryabhatta-or-Copernicus clarifies influence of Greek astronomers on Aryabhatta too. Still it would be nice if we had access to documents such as those in the library of Alexandria or if the ancient Hindu temples were not plundered. And Sun in the center of the Navagrahas at least indicates that the Hindus were aware of the size of the Sun. I do not know what additional things it could indicate.. as to whether it indicates if they already knew before the 1st millennia that the other grahas had apparent orbit around the Sun and whether such tweaks were consistent with geocentrism and whether Hindus were ready to accept inconsistencies into the religious frameworks (and whether they were ready to accept consistencies from the Greek knowledge such as relative size of the Sun (unlike the Greeks themselves)).

0 5 Rate This

14 February, 2025 at 9:36 am
Anonymous

Also arguably Hinduism as practiced (but not as philosophized) is about worship of deities of planets and planets themselves directly. There is an age old tradition for worshipping Navagrahas (nine planets) by circling their representative statues in a square (in south Indian temples) with 8 statues representing 8 planets (Moon and two lunar nodes are considered planets) and the Sun in the center. It would be interesting to know how old this tradition of heliocentric model has been worshipped.

2 5 Rate This
Reply

15 February, 2025 at 3:24 am
Πώς μετράμε το σύμπαν; – physicsgg

[…] ιστότοπο του Terence Tao (ΕΔΩ) περιέχονται περαιτέρω διευκρινίσεις και απαντήσεις […]

0 2 Rate This
Reply

16 February, 2025 at 1:32 pm
Anonymous

I was curious about how Kepler pieced together the orbit of Earth from those disjoint pieces he obtained from the orbit of Mars. To me, this seems like a superhuman effort.

However, it is rather frustrating to find sources confirming exactly how this was done. From what I gathered, Kepler never actually figured out that the orbit of the earth was an ellipse! He only assumed this after he had figured out that the orbit of Mars is an ellipse. He figured out that the orbit of Mars caved in slightly from the circular orbit; his idea of measuring the angles every 683 days was instrumental in his disproof that Mars had a circular orbit. But from what I gathered (and I couldn't find a source explicitly confirming this), he just worked with the model that Earth had a circular orbit which obeyed his second (area) law and used that with various trial and error with mathematical models to deduce that the orbit of Mars is an ellipse.

Perhaps you have more insight on this?

3 2 Rate This
Reply

17 February, 2025 at 8:17 am
Terence Tao

Kepler worked for many years on this problem, with several incomplete results and dead ends – I believe at an early stage he had even tried fitting the data to an ellipse, but had discarded this approach due to an unfortunate numerical miscalculation.

The fact that Mars's orbit was not perfectly in the ecliptic plane actually helped Kepler fix at least some pieces of this "jigsaw puzzle", despite apparently adding more complexity to the problem. In particular, Brahe's measurements of ecliptic latitude of Mars at times of opposition would allow Kepler to work out the location to

Mars at those times in terms of Mars's orbital plane with respect to the ecliptic plane. Some further data fitting would have needed to be done to determine that orbital plane, but this is a significantly simpler problem than determining the full orbit, as there are much fewer degrees of freedom.

Another fact that would have been helpful in piecing together the "jigsaw puzzle" was that the Earth would have had to return to the same position every 365 days, which allowed some of these "jigsaw pieces" to connect to each other.

Each "jigsaw piece" also would have provided evidence of Kepler's second law even without fitting those individual pieces into a larger orbit, so once he had grasped the universality and significance of that law, that would have provided further clues to gluing these pieces together. But it is unlikely that he proceeded in a purely deductive fashion from such axioms; as you say, most likely he kept trying different hypotheses that slowly revealed more and more of the structure of the solution, until he found one that fit all the partial information and laws that he had already worked out to astonishing accuracy.

10 1 Rate This
Reply

---

17 February, 2025 at 4:04 pm
Anonymous

Was there any sort of inverse square law available when Kepler was around or was it brought up by Newton? Would it have helped Kepler if there was any idea of inverse square law floating around?

1 1 Rate This
Reply

---

18 February, 2025 at 6:08 pm
Terence Tao

Given that neither Newtonian mechanics nor differential calculus was available in anything resembling modern form at Kepler's time, I doubt that even such an explicit hint would have been usable by Kepler. Historically, the arrow of causality went the other way: Kepler's laws helped guide Newton to his universal law of gravitation based on an inverse square force.

10 1 Rate This
Reply

---

1 March, 2025 at 7:10 am
Anonymous

Given the limited and error-prone data, Kepler got really lucky that his task was the orbit of Mars. It has 6 times the eccentricity as Earth, and more than all the other reasonable options.

Thank you kindly.

0 0 Rate This
Reply

---

18 February, 2025 at 5:14 pm
Nick Klein-Baer

I was glad to see that the question of how Eratosthenes knew that the sun's rays are (effectively) parallel is addressed here, and just wanted to share this excellent physics stack exchange post I came across years ago when wondering about that very question. It's particularly apt here as it provides a nice *visual* explanation of how this can be deduced from simple celestial observations.

When I first came across that post I don't recall having the impression that it was saying "this is literally how they knew", so much as "this is one way they might have known". But as I revisit it now I see that the poster does indicate that this was in fact the line of reasoning presented by Aristarchus in On the Sizes and Distances (of the Sun and Moon).

Thanks so much for the excellent video and blog post!

6 1 Rate This

Reply

19 February, 2025 at 9:47 am

jmdpark

Nice popularized expansion on the old cosmic distance ladder lecture, thanks for more details.

2 1 Rate This

Reply

23 February, 2025 at 8:44 am

Anonymous

Too bad there wasn't enough time to dive deeper into standard sirens, especially multi-messenger astronomy and GW170817—the only gravitational wave source we've been able to pinpoint. Thanks to their gamma ray emmited during the collision of the two neutron stars that luckily crossed earth some 144 millions years later.

2 2 Rate This

Reply

23 February, 2025 at 4:44 pm

Terence Tao

I am sure there are other videos or other popular presentations by professional astronomers who would do such a deep dive; I would welcome any such links by readers of this blog to be posted as comments here.

3 1 Rate This

Reply

24 February, 2025 at 7:36 pm

Craig Cahillane

Hi Terence, thanks to you and Grant for this wonderful video! I was very excited to see LIGO and standard sirens mentioned in the video as an independent measure of the Hubble constant.

Here is the link to the paper where the famous GW170817 LIGO gravitational-wave from a neutron-star merger detected at nearly the same time as a GRB was used to estimate the Hubble constant:
Nature: https://www.nature.com/articles/nature24471
arXiv: https://arxiv.org/abs/1710.05835

Some cliff notes from the paper:

1. In the near universe, $v_H = H_0 d$,
   where v_H is the "Hubble flow" velocity of the source away from us,
   H_0 is the Hubble constant,
   and d is the distance from us to the source.
2. Gravitational wave detections are *very good* at specifying the distance d that the astrophysical source merged away from us. This is due to the very predictable GW waveform of a binary neutron star merger, which precisely constrains the energy that the source can emit.
   From this, a decreasing amplitude of GW detected tells us how far away the GW source was.
3. Thanks to the GW, GRB, and subsequent telescope follow-up, we were able to pinpoint the source galaxy as NGC 4993. This galaxy has been part of some telescopic surveys, whose results were used to precisely estimate v_H.

I like this paper because while it represents the culmination of an enormous level of work and preparedness, the physics itself is so straightforward. We hope to catch more multimessenger sources in the very near future to further constrain the Hubble constant.

Another option is the so-called *dark sirens* which rely on binary black hole mergers distance and sky locations estimates convolved with galaxy catalogs to estimate the Hubble constant: https://arxiv.org/abs/1901.01540. While each individual estimate is lightly constraining, once you pile on a hundred detections, a good estimate might be converged upon.

Very interesting stuff, and thanks again for the wonderful explainer.

3 0 Rate This
Reply

23 February, 2025 at 9:07 am
Anonymous

Hi Terrence, hope you read this comment. I have always wondered how much relativity plays a role in the way light get's affected by it. If light was affected by say density from a cosmological point of view and not just empty space in our solar system point of view, we would have to take into account the density of galaxies depending on the direction we are looking at (measuring the distance of stars). What I mean by that is, I think that light is affected by relativity even based on the mass of the galaxy. So light would behave differently based on the Milky Way's mass compared to if we were in Andromeda or some other galaxy. This would also be true depending on the direction of the sky we are looking at. If we are looking perpendicular to the arms of the Milky Way then there's not enough density compared to looking through the arms of the Milky Way and seeing how much light bends or changes. And then this when blown up to the edge of the visible universe, we would have to take into account the red shift of the light we receive from distant galaxies (depending on how many galaxies are around the path of light and the relative density of galaxy [as in the number of galaxies surrounding that ray of light]), thus adding more error to our calculations. I don't know how much analysis we have done on this but is it potentially something worth looking into? This could maybe explain the Dark Matter observations we see in the universe. Maybe gravity (thus our observations) are different at cosmological scales relating to space between galaxies and galaxy densities. Why do spiral galaxy arms behave the way they do? Could this explain it?

2 3 Rate This
Reply

23 February, 2025 at 4:27 pm
Terence Tao

Yes, variation in the gravitational field does lead to measurable effects, most notably gravitational lensing, that add some additional complexity to the observations, though to my understanding the measurement of these additional effects have ended up confirming the standard cosmological model (which contains both dark matter and dark energy) rather than suggesting a replacement for it.

5 1 Rate This
Reply

23 February, 2025 at 4:41 pm
Terence Tao

I believe such theories are generally known as "tired light" theories. As I mentioned in another comment, one can use such theories to "fix" an individual observational anomaly that is not explained by the standard model, but to my knowledge every such proposed fix has actually created worse problems in other observations that fit the standard model very well, but not the newly proposed model. Improving a model that already explains a large number of disparate phenomena to reasonable accuracy is actually quite difficult!

6 9 Rate This
Reply

24 February, 2025 at 4:45 pm
Sahil Gupta

Hi Terence, the standard model does critique 'tired light', but with straw man arguments. Each of the tired light criticisms can be debunked (lack of blurriness, SN light curves, and CMB).

For starters, there's the straw man argument that tired light by interaction with intergalactic electrons would cause blurring. But Point 1: it is already known that photons interact with intergalactic electrons (see dispersion measure, dispersion slope), and evidently galaxies are not blurred. So that knocks out that argument. And Point

2: it is not a hard-and-fast claim of tired light that the redshift is caused by interaction with intergalactic electrons. The model leaves open the cause. It might even be the case that the photon simply decays on its own, somewhat like nuclear decay (which too is well-modeled by exponential decay).

https://cosmology.info/redshift/rebut/errorswright.html

6 4 Rate This

Reply

25 February, 2025 at 7:50 am
Terence Tao

The tired light theories have so far failed to follow the Lambda-CDM model into the modern era of high-precision cosmology, in which extremely detailed data sets from WMAP, Planck, JWST, SDSS, etc. are available. With the best-fit parameters, Lambda-CDM fits most data (in particular, CMB power spectrum, galactic structure, and baryon acoustic oscillations) to within 1% accuracy, with the one main outlier – the "Hubble tension" of the predicted value of H_0 being about 10% off from more direct measurements – being of sufficient inaccuracy to be deemed a "crisis in cosmology" in some circles. There is no proposed tired light model of sufficient mathematicial precision that achieves anywhere close to even this "crisis-level" 10% accuracy against *all* available data, though one can imagine that for any particular cherry-picked piece of data one can tailor a model to explain just that one individual measurement accurately, at the expense of accuracy in all other measurements. (But many proposed models are so qualitative in nature that no meaningful numerical measure of accuracy can even be assigned at all.)

5 5 Rate This

Reply

25 February, 2025 at 8:47 am
Sahil Gupta

Hi — You've extended the discussion without commenting on how the criticisms of tired light (like the lack of blurriness criticism) I mentioned are wrong.

Even still, I can answer your new points.

On WMAP and Planck and more generally the CMB: The instruments are precise but have not proven that the source of the incoming microwave radiation is some multi-billion-light-year distance away from the Milky Way. Can you prove otherwise? If you look into how "z=1100" is determined, it is not actually independently confirmed. If you list the stack of supporting ideas, the "z=1100 redshift of the CMB" is a claim that assumes the incoming microwave radiation was from a big bang and got redshifted. The claim of "z=1100" hides a premise. Alternatively, in non-big-bang cosmology, there remains a possibility the CMB-emitter is nearer the Milky Way, as in within a 100 Mly radius. And on a historical note, a single-digit Kelvin background was predicted before the proposal of the Big Bang Theory and long before the measurements of Penzias and Wilson. Big bang cosmologists were not the first to predict the incoming microwave radiation, and neither the most accurate. This history, https://www.ifi.unicamp.br/~assis/Apeiron-V2-p79-84(1995).pdf, is super important.

On baryon acoustic oscillations: BAOs are grandiose label for what is de facto a interesting statistical pattern of galaxies and a pattern of the peaks and troughs of the microwave radiation angular power spectrum. It's a general phenomenon that does not strictly favor Big Bang LCDM.

On JWST: JWST images of the distant universe, which show fully developed disc and spiral galaxies, strictly contradict pre-2022 predictions of big bang cosmology. This is the bigger crisis in cosmology. Now, big bang cosmologists are post-facto creating galaxy formation models. Isn't it astonishing that they're fabricating ad hoc free parameters instead of pondering that maybe the Big Bang theory is deeply in error, and that we have to debug the entire stack of ideas, starting with the redshift interpretation?

I do understand how strange it is to see a blog post commenter trying to convey to you that the Big Bang Theory is completely false, but I hope you might understand why he's trying to reach Terence Tao to break through.

6 3 Rate This

25 February, 2025 at 10:57 am
Terence Tao

To quote George Box, all models are wrong; the more important question is which ones are useful.

As I mentioned before, for any specific piece of data, it is possible to locate an alternative explanation using hypotheses and parameters tailored to that data that can match or outperform the standard Lambda-CDM model. But that does not provide a coherent competing alternative to that model. For that, one would need a mathematically precise cosmological model, with explicit numerical values (complete with error bars) of key parameters that are consistently chosen across all tests of the theory, that can give falsifiable predictions for each of the major cosmological data sets (as interpreted using the hypotheses of the model, of course), which one can then test against each of these sets in order to report the error between the model predictions and the actual data (after performing reasonable fitting of local unknown parameters). Until that occurs, one only has a hypothetical competitor to Lambda-CDM, rather than an actual one.

4 5 Rate This

25 February, 2025 at 11:17 am
Sahil Gupta

I'm disappointed you again mischaracterized a photon continuous decay redshift model as an "explanation using parameters tailored to that data" when tailoring parameters to data is the unabashed behavior of Big Bang LCDM (tailoring parameters to data is literally what inflation is, what dark matter is (2 free parameters per galaxy), what dark energy is, and what is going on with the ad hoc galaxy formation models for the mature spiral galaxies JWST observed).

Do you know what the TL model is?

The TL redshift model is

E_t / E_0 = e^(-Ht) = e^(-d/(c/H))

where *E* is energy of the photon, *t* is time of travel, or *d* is distance of travel,

and *H* is one parameter (the only parameter the model asks for), which derived from uncalibrated supernovae data is ≈ 1/(13.8 Gy).

5 3 Rate This

23 February, 2025 at 9:10 am
Anonymous

Curious, so we know that some of the data came from India – are there any records of possibly parallel endaeavours in ancuent China or Japan?

0 1 Rate This
Reply

23 February, 2025 at 4:52 pm
Terence Tao

I believe there are Chinese records from the first millennium of planetary motions and eclipses of roughly comparable levels of thoroughness to the Greek, Indian, and Islamic traditions, although there was less emphasis on geometric measurements and modeling, or of the type of deductive reasoning favored by the Greeks. However, they did seem to have a sophisticated ability to make calendrical calculations and predict eclipses, for instance having accurate values of the axial tilt of the Earth.

I am not familiar with the Japanese astronomy of the period, but I would imagine it was strongly influenced by the Chinese tradition.

2 1 Rate This
Reply

23 February, 2025 at 9:23 am
Anonymous

Hello! I'm a 12 grader, I had a question.

The expansion of the universe by hubble, is it in a 3d or 4d form?

0 1 Rate This
Reply

23 February, 2025 at 5:06 pm
Terence Tao

The special and general theories of relativity make it natural to view the universe as a four-dimensional spacetime, though one can take "slices" of this spacetime to create more familiar three-dimensional "snapshots" of the universe at a given instant of time (here there is some subtlety due to the fact that time and space are measured differently by different observers, but ignore this for now). So one can view the universe either as a "static" four dimensional object, or a "dynamic" three dimensional object. Neither perspective is "wrong", but the former is a more convenient perspective in some contexts, and the latter in others. One analogy is with a flip book: this can either be thought of as a three-dimensional object (consisting of pages that contain static images), or as a moving two-dimensional object, which changes as one flips through the pages of the book.

From the dynamic three-dimensional point of view, the universe is expanding, somewhat similarly to a balloon being inflated over time (except that the balloon is essentially two-dimensional instead of three-dimensional, and also sits inside an ambient three-dimensional space, whereas there is no reason to believe the universe sits inside a larger-dimensional space at cosmological scales). If one instead adopts the static four-dimensional point of view, the expansion of the universe is reflected in a certain "curvature" of spacetime, although describing exactly what curvature means here requires some rather sophisticated Lorentzian geometry. (But roughly speaking, much as curvature of a surface would cause bends in what would otherwise be a "straight" line on that surface, the curvature of spacetime will cause bends in the geometry of light rays, and other worldlines of inertial (free-falling) objects.)

9 7 Rate This
Reply

23 February, 2025 at 9:26 am
Anonymous

According to Matt O'Dowd, the scientific consensus on the size of the universe is more like 500 times larger in each direction than the observable universe. I have also heard this figure from Don Lincoln. The reasoning is that the flatness of the universe makes it at least 500 times larger in each direction, giving it a total of 125,000,000 times larger than the observable universe.

2 1 Rate This
Reply

23 February, 2025 at 10:48 am
Anonymous

Truly awesome. Thank you _so much_ for your time, knowledge, and enthusiasm!

1 1 Rate This
Reply

23 February, 2025 at 11:02 am
Anonymous

I really enjoyed these videos, but I believe there are a few inaccuracies in the discussion of Hubble's Law that your comment on Part 2, timestamp 21:07 doesn't quite address, and that weaken the claim that this video shows how you can measure distances across the full observable universe. In 20:35 you say that Hubble measured a linear relationship between galaxies' recessional velocities and distances, and in 20:45, you say this proportionality exists because "the universe is expanding at a uniform rate." You then say we can apply this (assumed to be linear) law to determine the distance to a galaxy, having measured its redshift.

But as we see it, the universe is not expanding at a uniform rate; the expansion is accelerating, so the recessional velocity that we observe for distant galaxies is not simply proportional to their distance (no matter what measure of distance you want to use–luminosity distance, comoving distance, or angular diameter distance). The linear approximation of Hubble's Law is really only accurate for quite nearby sources (z<~0.1), which doesn't get you much farther out (a few hundred megaparsecs) than Cepheids. We can use the full, nonlinear version of Hubble's Law (which requires assumed/measured values for the densities of matter, dark energy, and curvature) to calculate how far away a truly distant source is, having measured its redshift. But making that full version of Hubble's Law requires Type Ia supernovae to get some estimates for those cosmological parameters. I know it's mentioned why the video left out discussion of Type 1a supernovae. However, even if the video doesn't discuss how these supernovae allow you to build up the nonlinear version of Hubble's Law that actually extends to the scale of the observable universe, it seems important to add a correction that you can't just apply the linear approximation of Hubble's law at these scales and that the expansion of the universe is actually accelerating.

With the linear version of Hubble's Law, you're still pretty stuck to the relatively nearby, recent universe.

[Some clarifications added to this portion of the FAQ -T.]

2 5 Rate This
Reply

23 February, 2025 at 11:03 am
Lior Silberman

Hi Terry. For the comment at the end it might be worth mentioning the Tully–Fisher and Faber–Jackson relations (on top of the Type Ia SN), as additional standard candles.

I haven't worked in cosmology other than in my undergrad thesis, so discount appropriately.

[Thanks, this is now added to the relevant FAQ entry. -T]

2 1 Rate This
Reply

23 February, 2025 at 11:31 am
Matthijs van Duin

Regarding the size of the observable vs entire universe, my understanding of the research you linked is that they're looking for evidence (or absence thereof) that the universe is finite without boundary (e.g. a 3-torus) with a total size that's *smaller* than the observable universe, which would mean it loops in a way that is potentially observable.

The smaller the total universe, the easier it should be to detect such looping, so when attempts to detect this turn out negative (as they have been) this non-detection can then be quantified by putting a lower bound on the size of the universe.

If the total universe is larger than the observable universe (which seems most likely) then it seems to me that the actual size would be inherently unknowable, including whether it is finite or infinite?

3 1 Rate This
Reply

23 February, 2025 at 1:12 pm
chocolatewhispersa7adabc2d6

/I have a question, Mr. Terrance Tao (if this is a repeated question, please ignore. I think my first comment was not uploaded properly). At the end of the video, it was mentioned that there's always a 10 % error using Hubble's constant. Assuming that everything else is sound, could it be that some measurements cause a deviation because there're pockets of "more concentrated" dark energy, causing them to deviate from the constant? That is, could it be that dark energy is not evenly distributed on the Universe?

0 1 Rate This

Reply

23 February, 2025 at 4:36 pm
Terence Tao

There are some "local void" proposals along these lines, but my understanding is that it has been difficult to reconcile such proposals with the extreme uniformity of the cosmic microwave background radiation, as well as some other observed data on the distribution of galaxies that are otherwise well explained by the standard cosmological model. In general, there are a variety of ways to add one or more tweaks to this model to resolve any given observational anomaly, but in many cases this tends to create several more anomalous deviations from observed data for each anomaly resolved, and so I don't believe there is yet a comprehensive rival model to the standard Lambda-CDM model that has widespread acceptance.

2 5 Rate This

Reply

23 February, 2025 at 1:15 pm
Anonymous

Thank you so much for this 2 part series it is amazing to have everything bundled up and so easily accessible.

With regards to Rhomer method of calculating the speed of light if I get it right the 20 minutes or so are cumulative differences in Io's orbits as they appear to a moving Earth. Cumulative differences from when Earth moves away from Jupiter vs from when Earth moves towards Jupiter.

There is not a 10 minute aparent difference when we are closer than when we are further away. If Earth would be stationary at the closest point or stationary at the further away point the time between Io's eclipses would be the same (even if light has to travel a longer path).

Am I missing something?

1 1 Rate This

Reply

23 February, 2025 at 4:24 pm
Anonymous

I have a question about the 20 min ahead schedule when Jupiter is in opposition of the Sun. I don´t understand how they perceived it, shouldn't he also be 20 min ahead when disappearing and thus canceling out?

I would like to say the video was phenomenal and one of the beast videos I ever saw.

[Some additional text added to the FAQ entry to help explain this. From one transit to the next, the effects of finite light speed do indeed largely cancel, leaving only a very small shift of seconds in the period; but cumulatively over months, the light speed delays begin to accumulate. -T]

2 1 Rate This

Reply

23 February, 2025 at 4:57 pm
Anonymous

Parallax remains the only direct measurement of distance, all others rely on relationships that are statistical, with non-measurement related uncertainties, and require various corrections. When I was a graduate student, the Hipparcos satellite hadn't flown yet, and so your statement about "a thousand" stars with measured parallax was about the right order of magnitude. However, Hipparcos, and then Gaia have achieved measurements I thought would be impossible in my lifetime – direct measurements of distance and proper motion of 6 orders of magnitude more stars!

1 1 Rate This

Reply

23 February, 2025 at 9:20 pm

From my understanding, spacetime can warp due to gravitational forces. Could

dmiller this have affected the parallax measurements for the super far away distances?

0 0 Rate This
Reply

23 February, 2025 at 10:06 pm
Terence Tao

This does have some small impacts that become important at higher levels of precision, and the most modern parallax measurements (such as those coming from Gaia) take into account gravitational lensing effects from the Sun, other stars, and even the planets.

3 1 Rate This
Reply

23 February, 2025 at 11:19 pm
Anonymous

Dear Terry. Maybe it is a big ask at this moment but I was wondering if you have somewhere a bibliography to share regarding your book project, or in the understandable absence of one at this moment a short list of the books you have found more useful/interesting about the cosmic ladder history.

0 0 Rate This
Reply

24 February, 2025 at 7:53 am
Terence Tao

This only concerns one aspect of the distance ladder, but I can recommend "The Glass Universe" by Dana Sobel.

2 0 Rate This
Reply

23 February, 2025 at 11:49 pm
Anonymous

Very enjoyable and informative videos! Too often, we are just asked to trust the science – i.e. an argument from authority – but it is more important to explain *why* we believe what we believe, and to propose experiments that at least in principle could be performed to confirm our beliefs.

0 0 Rate This
Reply

24 February, 2025 at 3:02 am
Anonymous

Really cool!

Also, "with corrections" is a typical thing for people like us, isn't it.

0 0 Rate This
Reply

24 February, 2025 at 5:21 am
Anonymous

Hi, would it be conceivable or, at least in theory make sense to measure some sort of apparent size of the universe based on the behaviour of our observable universe? so:

I've always imagined our observable universe as its own bubble – just like a planet orbits its sun, so does our observable universe some sort of larger structure. thus if we could observe how the whole obeservable universe behave, maybe we can infer some sort of direction of our bubble – maybe based on some global redshift direction, or perhaps a differential in the overall direction everything moves galaxies moving in in one way, and moving out in the opposite diretion? so basically using the observable universe a staging ground for the next level paralax?

0 0 Rate This

Reply

24 February, 2025 at 5:46 am
Petar Knezevich

Thank you for such a nice review of the cosmic distance ladder. It is inspiring to be reminded how many millennia of observing, analyzing and theorizing have been dedicated to furthering knowledge. We should have all been taught these histories of mathematics and science when we were in school.

0 0 Rate This
Reply

24 February, 2025 at 5:54 am
Anonymous

Quick question; you write that "the Nile ran from North to South" but it actually ran from South to North, right? This doesn't affect the rest of that section, of course.

[Corrected, thanks – T.]

0 0 Rate This
Reply

24 February, 2025 at 6:06 am
deliciouslye7bcae2778

Hi, would it make sense to imagine the sphere of our observable universe, like planet orbiting the sun, orbiting through space and infer some larger structure or and apparent size of the universe? Basically, we should see new things come into the observable universe on one side, and something else disappear on the other side, or something similar to that effect that perhaps could allow us to infer something about the behaviour of the observable universe as a whole?

0 0 Rate This
Reply

24 February, 2025 at 7:57 am
Terence Tao

This is roughly how the cosmic microwave background radiation is used to infer some lower bounds on the size of the actual universe, as mentioned in the 21:15 entry of the FAQ. There is an important complication to bear in mind, though: because of the finite speed of light, the observable universe should not be viewed as a sphere that moves instantaneously with us, but rather as a "backwards light cone" that extends further into the past as one moves further into the edge of the observable universe. In particular, at the edges of the observable universe we are now observing the very early stages of the universe, which is why the cosmic microwave background radiation, the features of which also mostly originate from that era, are the most important data source we have for that edge.

2 0 Rate This
Reply

24 February, 2025 at 7:09 am
Anonymous

These two videos were a treat! Grant's videos are always excellent and Terence's exposition was a beautiful fit.

How do astronomers compensate for attenuation in brightness due to intervening dust/gas clouds (when using standard candles)?

Is the redshift due to expansion of the universe "obvious"? I thought the expansion is sort of happening "between" galaxies. Also, that the speed of light "locally" (as it were), through space, is indeed still c. — I'm obviously not being clear here, almost deliberately confusing myself, but "the expansion of space" (ie. "the metric") isn't clearly analogous to the Doppler shift of a siren. Is it?

1 2 Rate This
Reply

24 February, 2025 at 8:11 am

To compensate for attentuation (also known as https://en.wikipedia.org/wiki/

Terence Tao Extinction_(astronomy)) one can work in multiple frequencies of the spectrum (visible, x-ray, radio, etc.), as some frequencies are less susceptible to extinction than others. From the discrepancies between these frequencies one can infer the amount of extinction, leading to "dust maps" that can then be used to facilitate such corrections for subsequent measurements in the same area of the universe. (More generally, the trend in modern astronomy is towards "multi-messenger astronomy" in which one combines together very different types of measurements of the same object to obtain a more accurate understanding of that object and its surroundings.)

As for your second question: it is true that an observed redshift could, in isolation, be explainable purely by the Doppler effect of a galaxy receding in a flat background universe, rather than by expansion in the intervening spacetime metric. However, this would not explain why distant galaxies would have a greater redshift than nearby galaxies, as in the absence of an expanding universe, there would be no mechanism to ensure that recessional velocities should be proportional to distance from the observer.

The "expanding universe" model was not immediately accepted after the discovery of Hubble's law (somewhat analogously to how the heliocentric model was not immediately accepted after Aristarchus's measurements of the Sun), as it was initially possible to explain this law from some other rival models. The later discovery of the cosmic microwave background, which was highly consistent with a "Big Bang" model but difficult to explain in other models, was a major factor in achieving consensus adoption of this model among astronomers; there is also other supporting evidence consistent with the standard cosmological model, such as the observed distribution of the galaxies and the abundance of hydrogen and helium.

4 6 Rate This
Reply

---

24 February, 2025 at 9:26 am
Anonymous

I liked the part of the "story", where Kepler`s motivations were mentioned.

The fact that Platonic solids can be inscribed in the orbits was in itself so beautiful and meaningful to him that he approached it also from a theological perspective (of course, a somewhat typical thing for his period).

This didn`t stop him though from conducting his research in the most honest way, paving the way to a discovery of another beautiful and unobvious thing, namely that orbits are in fact conic sections, which I find quite funny.

It is also a nice parallel of current scientific research, where there can be a lot of apparent elegance and beauty in a particular model, which can be a great motivator, but should never hinder the actuall pursuit of truth.

Thanks for all the work on this!

0 1 Rate This
Reply

---

24 February, 2025 at 12:13 pm
Anonymous

This was such a fun read with many points to tangent off into. I'd like to highlight a video on recent(ish) advancements on the "crisis in cosmology" by Dr. Becky Smethurst. Starting from 13:34, they give an introduction on the question, and then delve into how different teams of researchers came up with different estimates of the hubble constant using JWST data. I felt if anyone wanted to delve deeper into this topic it would be a good place to start.

2 0 Rate This
Reply

---

24 February, 2025 at 3:57 pm
Anonymous

Hello Terry, thanks for these fascinating videos. On part II, a the 12:40 mark one see a very distinct moving star against a background of more distant stars. There is no way this motion is produced by parallax, at least not by the Earth moving around its orbit. Indeed at the distance of Proxima Centauri, the total parallax angle seen from Earth is less than 2″ of arc, which is only about the FWHM visible diameter of a bright star taken from Earth even in good seeing conditions. The motion shown in the video is much more likely produced by the proper motion of the star as it orbits the centre of the galaxy.

The only such motion due to parallax so clearly visible on a photograph was taken by the probe New Horizon as it was cruising well beyond the orbit of Pluto. See this:

https://cdn.mos.cms.futurecdn.net/z3pC8PmqTZWxwBcHyHRkWZ-1200-80.gif

Indeed the Wikipedia page on star proper motion shows the example of Barnard's star, which is similar in magnitude to the motion you show in the video.

This does not deter in any way from the quality of your exchanges. I was just bothered by this. All the best !

1 0 Rate This

Reply

24 February, 2025 at 4:18 pm

Anonymous

Wow! I am only 13 and most of this went completely over my head, but I loved it you are answering so many questions and taking the time to explain these incremental beautiful discoveries to everyone!

1 0 Rate This

Reply

24 February, 2025 at 5:49 pm

Mr Chris

Possibly uneducated comment here, but;

Does the coalescing of matter in the universe into these "strands" impact the measurement of their red shift? If a galaxy were on the other side of a strand being drawn in our direction, would that movement have an effect on the red shift? Or is it too slow/big/etc.?

0 0 Rate This

Reply

24 February, 2025 at 7:47 pm

Terence Tao

There are some local motions of clusters of galaxies, known as bulk flows, that do make some noticeable impact on red shifts for nearby galaxies, but I don't believe they make a significant impact for very distant galaxies, particularly if one performs statistical averaging over many such galaxies to reduce the net magnitude of these effects.

3 0 Rate This

Reply

25 February, 2025 at 7:06 pm

Richard

At a "local" level, surveys do indeed reveal overdensities of matter (eg The Great Attractor) towards which the motions (ie redshifts) of more-than-statistically-expected galaxies redshifts "point".

Even larger scale galaxy clusters are revealed by the Sunyaev–Zeldovich effect on the cosmic microwave background radiation.

At the most cosmological levels, the Sachs–Wolfe effect causes uneven gravitational redshift of the cosmic microwave background in a dark-energy-dominated universe (one in which expansion is accelerating, which is our present "ΛCDM" consenus model), revealing the largest structures of the observable universe. CMB photons on their way to us from the Big Bang take so long to traverse these enormous structures (filaments or voids between filaments) that the *acceleration* of the expansion of the universe *during* the light traversal decays the gravitational potential of the structure and so blueshifts (or redshifts through voids) CMB photons' energy.

0 3 Rate This

Reply

24 February, 2025 at 11:45 pm
Anonymous

This may not be useful to many people, but for those who speak Spanish, you can go to the QuantumFracture YouTube channels to see more videos on this topic, including more information on type 1a supernovae and other "candles" used in the cosmic ladder.

https://www.youtube.com/@QuantumFracture

https://www.youtube.com/watch?v=LR_Gz1LsUpM

1 0 Rate This

Reply

25 February, 2025 at 3:36 am
pinktotallya7d9f24ba6

Fabulous talks. This extraordinary history has often been told but never with such concision. We get to get the gist and axsense of belonging. I think this might also be calling into question traditional schooling models as there is so much knowledge out there for the curious, even in deprived areas. We can work a lot out by ourselves with the right guidance. The collaborative format is a step towards this as well. Looking forward to the upcoming book for more nitty-gritty details indeed.

0 0 Rate This

Reply

25 February, 2025 at 8:53 am
Anonymous

Kepler's discoveries were paved by an immense amount of preparatory work with no guarantee of an outcome in the end. The planets orbit the common center of mass at/near the sun in an ellipse. In school I learned that an ellipse can be described by putting two pins in a board and tying each end of a string to the pins and then a pencil can be used in the stretched out string to scribe out the ellipse. The sun is at one pin, or focus, with nothing at the other pin, or focus. It has always bothered me that there is nothing at the other focus. Why should the orbit be an ellipse if there is nothing at the other focus? This has always bothered me. Intuitively it made no sense to me to have two foci with nothing at one of them. My AI does not seem to help resolve this. So it won't keep me up at nights, I have decided that the sun is at one focus and the system is balanced at the other focus by time. I'm not used to sensing time in the same way as I sense mass so maybe that is why intuitively it seemed wrong to have nothing at the other focus. Why should it make intuitive sense that there is nothing at one focus? Jimmy Ellis at utgeek@earthlink.net

0 1 Rate This

Reply

26 February, 2025 at 1:32 am
Anonymous

The orbit is not symmetrical – when Earth is closest to the focal point populated by the Sun (or more accurately by the center of mass of the system), it has the highest kinetic energy and lowest potential energy, with the opposite on the far side. Thus, there is really no reason for anything to be in the other focus.

0 1 Rate This

Reply

25 February, 2025 at 2:34 pm
Anonymous

Really appreciate all the work and thought and insight in this post and the videos; thank you to the team that put it together!

Sending fond greetings from Sydney! 👋

0 0 Rate This

Reply

25 February, 2025 at 5:11 pm

Anonymous

Very curious where "the observable universe is 20% of the universe" (paraphrasing) came from!?

0 0 Rate This

Reply

25 February, 2025 at 7:06 pm

Tara Li

How do the various levels of uncertainty between the different rungs on the ladder interact? It seems like these uncertainties could add up quickly, since each rung depends on the one before it. I've seen estimates of the distance to various stars (pre-Gaia data, I presume) that differed by a factor of two or more – and they're relatively close stars, such as Betelgeuse or Polaris. Wouldn't that kind of inaccuracy affect the next step – Cephids (ignoring the T1/T2 Cephid discrepancy), leading to even more uncertainty. Then there's things like 1a Supernovae, which keep getting glossed over as being the same brightness – but surely there would be at least some variation, even if just 1-5%.

How does this climbing accumulation of error ranges get resolved?

0 0 Rate This

Reply

26 February, 2025 at 7:53 am

Terence Tao

There is a lot of work on cross-calibrating between different distance measurement methods. For instance, if a single galaxy contains both Cepheids and tips of red giants then one can compare the two measurements to see if either of them contain undesirable variability. For errors that are not systemic in nature, statistical averaging over many different measurements can also reduce relative error. For errors that themselves can be modeled with reasonable precision, e.g., extinction effects modeled through dust map measurements, one can compensate for those contributions by inferring the parameters of the error model.

As gravitational wave distance measurements become more plentiful and accurate, this should provide a way to calibrate many rungs of the ladder at once, since this type of measurement avoids a lot of the problems you mentioned.

0 0 Rate This

Reply

25 February, 2025 at 8:25 pm

Anonymous

Terry I believe the visualizations in the last section of part 1 are incorrect; they show that the synodic position of mars is fixed for several relative earth positions; when the sidereal position of mars should have been fixed. It would have been impossible for Kepler to have known that mars was in the same place *a priori* that can only happen *a posteriori*

0 0 Rate This

Reply

26 February, 2025 at 7:58 am

Terence Tao

Kepler was using Copernicus's value for the sidereal period of Mars (687 days), not the synodic period (780 days). (In the video I had somehow managed to average these numbers and quote the period as 729 days, but this was an error.)

In short, the original Copernican model does indeed predict that Mars returns to the same location in the Solar system every 687 days. It turns out that even though the orbits are not perfectly circular, they do not exhibit significant precession (at least at the level of accuracy available in Kepler's day; there is of course the famous issue of Mercury's precession due to relativistic effects that was discovered later); because of this fortunate lack of precession, the Copernican sidereal periods could still be used to high accuracy in Kepler's analysis. (However, in an alternate universe where the law of gravity was sufficiently different that orbits did experience significant precession, then Kepler's task would have been significantly harder, and perhaps even impossible with the technology and mathematics of the era.)

1 0 Rate This
Reply

---

26 February, 2025 at 12:23 am
Anonymous

Cool video / conversation! You mention that Rømer's estimates have been updated from about 20 minutes to about 16 with modern technology. How does modern technology help here? Is it just more precise clocks? It doesn't seem like this calculation should be downstream of much else that could've been fuzzy at the time (unlike e.g. the speed of light, which is downstream of this time measurement + distance from earth to sun), and I'm surprised that the timing measurements alone would've been so far off.

Thanks!

0 0 Rate This
Reply

---

26 February, 2025 at 8:08 am
Terence Tao

There were several sources of error in Romer's original calculations. In addition to possible timekeeping errors, there were also measurement errors in timing the transits of Io perfectly. It is also difficult to measure transits accurately at times of conjunction (because the Sun is obstructing the transit) and to a lesser extent in opposition (because now Jupiter is obstructing the transit). So Romer actually relied on data coming from intermediate times between conjunction and opposition in which there was a viewing angle that avoided these difficulties. But then one has to do some trigonometric calculations to compensate for this angle. My understanding is that Romer made a simple assumption of circular orbits for both Earth and Jupiter, which of course is not quite right due to Kepler's laws.

Finally, with the passage of time, more and more transits can be observed, and non-systemic measurement errors can be averaged out. Romer's method was replicated by Jean Baptiste Joseph Delambre in 1809, and by that time there was enough data (as well as a precise value of the AU) that a quite accurate measurement of the speed of light was obtained (close to 300,000 km/sec), as well as the time needed for light to traverse an AU (just over 8 minutes). I believe modern recreations of the method match very closely with other measurements of light speed and the AU.

1 0 Rate This
Reply

---

26 February, 2025 at 1:47 am
Anonymous

Hello Mr. Tao,

I was wondering if you could share:

(a) your own personal interpretation of the result of the Michelson-Morley experiment

(b) why the result of the Michelson-Morley experiment was called a "null result"

(Def: In science, a "null result" is a result without the expected content.)

Thanks very much.

KB

0 0 Rate This

Reply

26 February, 2025 at 12:27 pm
Anonymous

Is there a formula or some sort of equation to estimate or to deduce the distance between the milky way and the nearby galaxies?

0 0 Rate This

Reply

26 February, 2025 at 4:18 pm
Terence Tao

Pretty much every rung of the modern distance ladder is basically an equation relating the distance to a given galaxy in terms of other things that we can measure. For instance, Hubble's law is an equation relating the distance to the redshift. If the galaxy contains a standard candle, then we have an equation relating the distance to the apparent magnitude and absolute magnitude of the standard candle; and similarly for standard sirens and standard rulers.

6 0 Rate This

Reply

2 March, 2025 at 12:01 pm
Anonymous

Thank you both, Terry and Grant, for the videos and this write-up. Beautiful and diligent work, much needed today!

Looking forward to your book about the Ladder, and I have a humble wish/suggestion:

If not happening already, can you perhaps collaborate with Grant to jazz up the book with some stunning and accurate graphics?

1. I think Grant's expertise in precise animations will translate well into making consistent still images for print, he already has a lot of the base code done, and it may well be his next new challenge as an educator to fit his work to the print medium.
2. Too many pop-sci books nowadays end up being walls of text, by trying to avoid "technical" tools of explanation, like diagrams and equations. IMO they underestimate how excited the average reader can be to see things they don't necessarily immediately understand but are willing to learn, and miss a great opportunity that way.

0 0 Rate This

Reply

2 March, 2025 at 8:52 pm
Anonymous

Are there other problems that you've enjoyed using the Cosmic Distance Ladder as a model for while thinking about it? To me it has the feeling of a metaphor for scientific discovery in general, slowly incrementing one's knowledge over time. But the metaphor is a little hollow, I'm struggling to populate it with specific examples.

0 0 Rate This

Reply

3 March, 2025 at 8:02 am
George Lowther

Hi Terence,

I really enjoyed watching the videos and learned some new stuff.

Btw, I'm George. Not important, but just mentioning as you called me Greg in the links at the end!

Thanks

George (I.e., Almost Sure)

[Corrected, thanks – T.]

0 0 Rate This

Reply

7 March, 2025 at 7:09 am
March '25 Links | The Tao of Gaming

[…] the Cosmic Ladder — how do we deduce astronomical distances? Part 1. Part 2. The (other) Tao's FAQ on the videos. Terence Tao lecture on the state of the art on Machine-assisted […]

0 0 Rate This

Reply

8 March, 2025 at 6:52 pm
MP

Professor Tao, thank you very much for the effort put here.

I tried discussing the topic of periodicities in the cosmic microwave radiation with ChatGPT, but I am not sure I 100% understood. Apparently, the idea is that if the universe would be finite, a photon would be able to go back to the same place in the universe just by going straight, just like one does in a sphere.

But how can physicists be sure that it'd be the case? From my humble vantage point, it seems like a big extrapolation just from the fact we don't see periodicities in the CMR.

0 0 Rate This

Reply

9 March, 2025 at 9:35 am
Terence Tao

The lack of periodicities in the CBR is evidence against the universe being too small – roughly speaking, it can rule out certain scenarios in which the diameter of the universe is smaller than twice the diameter of the observable universe (assuming a standard cosmological model like Lambda-CDM). But it is consistent with other scenarios in which the universe is finite, but significantly larger than the observable universe.

More recently, experimental evidence from Planck suggests that the flatness of the universe is surprisingly close to zero, which suggests an extremely large inflationary effect in the early universe which in turn hints that the universe is indeed much larger than the observable universe, or possibly even infinite; but the error bars are still too large to draw definitive conclusions.

1 0 Rate This

Reply

26 March, 2025 at 10:52 pm
Anonymous

"Redshift" is the speed. Explain please how, knowing the speed, you can calculate the distance? If a car is traveling at 100 km/h at the other end of the city, can you calculate the distance to it?

"The second point" is, explain how the stars and the clouds of hydrogen that surround them move together at the same speed – how the redshift does work?

“The third one.” Has anyone measured the blue shift of Andromeda, which is approaching us?

0 0 Rate This
Reply

27 March, 2025 at 6:26 am
Terence Tao

Yes, redshift only directly computes recession velocity, not distance. However, as observed by Hubble after computing both the redshift (and thus implied recession velocity) and distance of multiple nearby galaxies and nebulae, he found that the redshift or recession velocity was approximately linearly proportional to the distance; more distant galaxies tended to have more distant recession velocities. See for instance the graphs at this link.

This empirical finding was later explained by the famous “Big Bang” model of the universe, later refined to the modern “Lambda-CDM” model which also incorporates a “dark energy” term that generates some additional acceleration in the expansion of the universe.

As you note in your other questions, galaxies also exhibit peculiar velocities that are in addition to the general recession caused by Hubble’s law, so the law is not a perfect fit for the nearest galaxies. However, as one moves out to further and further distances, the effects of the expanding universe dominate the effects of peculiar velocities (which stay more or less the same magnitude across the universe), and so the law becomes more accurate (in terms of relative error) as one goes further out (again, see the graph in the previous link provided, or even the original graph of Hubble displayed in the video). Also, if one statistically averages over many galaxies instead of focusing on individual galaxies, the effects of peculiar velocities (which can either increase or decrease the redshift) tend to cancel out, and so the reliability of Hubble’s law improves if one works with statistical averages.

(See also the FAQ entries at 20:37, 20:46, and 21:07 of part 2.)

2 0 Rate This
Reply

28 March, 2025 at 8:07 am
Anonymous

>**Isn’t the illuminated portion of the Moon, as well as the visible portion of the Moon, slightly smaller than half of the entire Moon**

>In reality, the Sun turns out to be about 86,000 Moon radii away from the Moon, so asserting that half of the Moon is illuminated by the Sun is actually a very good first approximation.

As the sun is somewhat larger than the moon, one might expect the illuminated portion of the moon to be *larger* than half, rather than smaller. If the sun is a plane, for instance, there is exactly one point on the moon that is not illuminated by it.

[Fair point; I expanded the FAQ entry to mention this. -T]

0 0 Rate This
Reply

28 March, 2025 at 8:10 pm
Anonymous

You could argue that the moon is illuminated by the sun only when the full sun is visible, I suppose, but for me the distinction between being able to see some portion of the sun, and no portion of the sun, is night and day.

0 0 Rate This
Reply