Early history[edit]

The hypothesis of dark matter has an elaborate history.[19] In the appendices of the book *Baltimore lectures on molecular dynamics and the wave theory of light* where the main text was based on a series of lectures given in 1884,[20] Lord Kelvin discussed the potential number of stars around the Sun from the observed velocity dispersion of the stars near the Sun, assuming that the Sun was 20 to 100 million years old. He posed what would happen if there were a thousand million stars within 1 kilo-parsec of the Sun (at which distance their parallax would be 1 milli-arcsec). Lord Kelvin concluded "Many of our supposed thousand million stars, perhaps a great majority of them, may be dark bodies".[21][22] In 1906, Henri Poincaré in "The Milky Way and Theory of Gases" used the French term *matière obscure* ("dark matter") in discussing Kelvin's work.[23][22] He found that the amount of dark matter would need to be less than that of visible matter.[24]

The second to suggest the existence of dark matter using stellar velocities was Dutch astronomer Jacobus Kapteyn in 1922.[25][26] A publication from 1930 points to Swedish Knut Lundmark being the first to realise that the universe must contain much more mass than can be observed.[27] Dutchman and radio astronomy pioneer Jan Oort also hypothesized the existence of dark matter in 1932.[26][28][29] Oort was studying stellar motions in the local galactic neighborhood and found the mass in the galactic plane must be greater than what was observed, but this measurement was later determined to be erroneous.[30]

In 1933, Swiss astrophysicist Fritz Zwicky, who studied galaxy clusters while working at the California Institute of Technology, made a similar inference.[31][32] Zwicky applied the virial theorem to the Coma Cluster and obtained evidence of unseen mass he called *dunkle Materie* ('dark matter'). Zwicky estimated its mass based on the motions of galaxies near its edge and compared that to an estimate based on its brightness and number of galaxies. He estimated the cluster had about 400 times more mass than was visually observable. The gravity effect of the visible galaxies was far too small for such fast orbits, thus mass must be hidden from view. Based on these conclusions, Zwicky inferred some unseen matter provided the mass and associated gravitation attraction to hold the cluster together.[33] Zwicky's estimates were off by more than an order of magnitude, mainly due to an obsolete value of the Hubble constant;[34] the same calculation today shows a smaller fraction, using greater values for luminous mass. Nonetheless, Zwicky did correctly conclude from his calculation that the bulk of the matter was dark.[22]

Further indications of mass-to-light ratio anomalies came from measurements of galaxy rotation curves. In 1939, Horace W. Babcock reported the rotation curve for the Andromeda nebula (known now as the Andromeda Galaxy), which suggested the mass-to-luminosity ratio increases radially.[35] He attributed it to either light absorption within the galaxy or modified dynamics in the outer portions of the spiral and not to the missing matter he had uncovered. Following Babcock's 1939 report of unexpectedly rapid rotation in the outskirts of the Andromeda galaxy and a mass-to-light ratio of 50; in 1940 Jan Oort discovered and wrote about the large non-visible halo of NGC 3115.[36]

1960s[edit]

Early radio astronomy observations, performed by Seth Shostak, later SETI Institute Senior Astronomer, showed a half-dozen galaxies spun too fast in their outer regions, pointing to the existence of dark matter as a means of creating the gravitational pull needed to keep the stars in their orbits.[37]

1970s[edit]

Vera Rubin, Kent Ford, and Ken Freeman's work in the 1960s and 1970s[38] provided further strong evidence, also using galaxy rotation curves.[39][40][41] Rubin and Ford worked with a new spectrograph to measure the velocity curve of edge-on spiral galaxies with greater accuracy.[41] This result was confirmed in 1978.[42] An influential paper presented Rubin and Ford's results in 1980.[43] They showed most galaxies must contain about six times as much dark as visible mass;[44] thus, by around 1980 the apparent need for dark matter was widely recognized as a major unsolved problem in astronomy.[39]

At the same time Rubin and Ford were exploring optical rotation curves, radio astronomers were making use of new radio telescopes to map the 21 cm line of atomic hydrogen in nearby galaxies. The radial distribution of interstellar atomic hydrogen (H^{I}) often extends to much greater galactic distances than can be observed as collective starlight, expanding the sampled distances for rotation curves – and thus of the total mass distribution – to a new dynamical regime. Early mapping of Andromeda with the 300 foot telescope at Green Bank[45] and the 250 foot dish at Jodrell Bank[46] already showed the H^{I} rotation curve did not trace the expected Keplerian decline. As more sensitive receivers became available, Roberts & Whitehurst (1975)[47] were able to trace the rotational velocity of Andromeda to 30 kpc, much beyond the optical measurements. Illustrating the advantage of tracing the gas disk at large radii; that paper's *Figure 16*[47] combines the optical data[41] (the cluster of points at radii of less than 15 kpc with a single point further out) with the H^{I} data between 20 and 30 kpc, exhibiting the flatness of the outer galaxy rotation curve; the solid curve peaking at the center is the optical surface density, while the other curve shows the cumulative mass, still rising linearly at the outermost measurement. In parallel, the use of interferometric arrays for extragalactic H^{I} spectroscopy was being developed. Rogstad & Shostak (1972)[48] published H^{I} rotation curves of five spirals mapped with the Owens Valley interferometer; the rotation curves of all five were very flat, suggesting very large values of mass-to-light ratio in the outer parts of their extended H^{I} disks.[48]

1980s[edit]

A stream of observations in the 1980s supported the presence of dark matter, including gravitational lensing of background objects by galaxy clusters,[49] the temperature distribution of hot gas in galaxies and clusters, and the pattern of anisotropies in the cosmic microwave background. According to consensus among cosmologists, dark matter is composed primarily of a not-yet-characterized type of subatomic particle.[50][51] The search for this particle, by a variety of means, is one of the major efforts in particle physics.[52]

Twenty-first century[edit]

While primordial black holes were long considered possibly important if not nearly exclusive components of dark matter,[53][54][55][56] the latter perspective was strengthened by both LIGO/Virgo interferometer gravitational wave and James Webb Space Telescope (JWST) observations.[13][15] Early constraints on PBHs as dark matter usually assumed most black holes would have similar or identical ("monochromatic") mass, which was disproven by LIGO/Virgo results, and further suggestions that the actual black hole mass distribution is broadly platykurtic were evident from JWST observations of early large galaxies.[57][58][59]

Technical definition[edit]

See also: Friedmann equations

In standard cosmological calculations, *"matter"* means any constituent of the universe whose energy density scales with the inverse cube of the scale factor, i.e., $\rho \propto a^{-3}$. This is in contrast to *"radiation"*, which scales as the inverse fourth power of the scale factor $\rho \propto a^{-4}$, and a cosmological constant, which does not change with respect to a ($\rho \propto a^{0}$). The different scaling factors for matter and radiation are a consequence of radiation redshift: For example, after gradually doubling the diameter of the observable Universe via cosmic expansion of General Relativity, the scale, a, has doubled. The energy of the cosmic microwave background radiation has been halved (because the wavelength of each photon has doubled);[60] the energy of ultra-relativistic particles, such as early-era standard-model neutrinos, is similarly halved.[d] The cosmological constant, as an intrinsic property of space, has a constant energy density regardless of the volume under consideration.[61][e]

In principle, "dark matter" means all components of the universe which are not visible but still obey $\rho \propto a^{-3}$. In practice, the term "dark matter" is often used to mean only the non-baryonic component of dark matter, i.e., excluding "missing baryons". Context will usually indicate which meaning is intended.

Observational evidence[edit]

Galaxy rotation curves[edit]

Main article: Galaxy rotation curve

Animation of rotating disc galaxies. Dark matter – shown in red – is more concentrated near the center and it rotates more rapidly.

The arms of spiral galaxies rotate around the galactic center. The luminous mass density of a spiral galaxy decreases as one goes from the center to the outskirts. If luminous mass were all the matter, then we can model the galaxy as a point mass in the centre and test masses orbiting around it, similar to the Solar System.[f] From Kepler's Third Law, it is expected that the rotation velocities will decrease with distance from the center, similar to the Solar System. This is not observed.[62] Instead, the galaxy rotation curve remains flat as distance from the center increases.

If Kepler's laws are correct, then the obvious way to resolve this discrepancy is to conclude the mass distribution in spiral galaxies is not similar to that of the Solar System. In particular, there is a lot of non-luminous matter (dark matter) in the outskirts of the galaxy.

Velocity dispersions[edit]

Main article: Velocity dispersion

Stars in bound systems must obey the virial theorem. The theorem, together with the measured velocity distribution, can be used to measure the mass distribution in a bound system, such as elliptical galaxies or globular clusters. With some exceptions, velocity dispersion estimates of elliptical galaxies[63] do not match the predicted velocity dispersion from the observed mass distribution, even assuming complicated distributions of stellar orbits.[64]

As with galaxy rotation curves, the obvious way to resolve the discrepancy is to postulate the existence of non-luminous matter.

Galaxy clusters[edit]

Galaxy clusters are particularly important for dark matter studies since their masses can be estimated in three independent ways:

- From the scatter in radial velocities of the galaxies within clusters
- From X-rays emitted by hot gas in the clusters. From the X-ray energy spectrum and flux, the gas temperature and density can be estimated, hence giving the pressure; assuming pressure and gravity balance determines the cluster's mass profile.
- Gravitational lensing (usually of more distant galaxies) can measure cluster masses without relying on observations of dynamics (e.g., velocity).

Generally, these three methods are in reasonable agreement that dark matter outweighs visible matter by approximately 5 to 1.[65]

Gravitational lensing[edit]

One of the consequences of general relativity is massive objects (such as a cluster of galaxies) lying between a more distant source (such as a quasar) and an observer should act as a lens to bend light from this source. The more massive an object, the more lensing is observed.

Strong lensing is the observed distortion of background galaxies into arcs when their light passes through such a gravitational lens. It has been observed around many distant clusters including Abell 1689.[66] By measuring the distortion geometry, the mass of the intervening cluster can be obtained. In the dozens of cases where this has been done, the mass-to-light ratios obtained correspond to the dynamical dark matter measurements of clusters.[67] Lensing can lead to multiple copies of an image. By analyzing the distribution of multiple image copies, scientists have been able to deduce and map the distribution of dark matter around the MACS J0416.1-2403 galaxy cluster.[68][69]

Weak gravitational lensing investigates minute distortions of galaxies, using statistical analyses from vast galaxy surveys. By examining the apparent shear deformation of the adjacent background galaxies, the mean distribution of dark matter can be characterized. The mass-to-light ratios correspond to dark matter densities predicted by other large-scale structure measurements.[70] Dark matter does not bend light itself; mass (in this case the mass of the dark matter) bends spacetime. Light follows the curvature of spacetime, resulting in the lensing effect.[71][72]

In May 2021, a new detailed dark matter map was revealed by the Dark Energy Survey Collaboration.[73] In addition, the map revealed previously undiscovered filamentary structures connecting galaxies, by using a machine learning method.[74]

An April 2023 study in *Nature Astronomy* examined the inferred distribution of the dark matter responsible for the lensing of the elliptical galaxy HS 0810+2554, and found tentative evidence of interference patterns within the dark matter. The observation of interference patterns is incompatible with WIMPs, but would be compatible with simulations involving 10^{-22} eV axions. While acknowledging the need to corroborate the findings by examining other astrophysical lenses, the authors argued that "The ability of (axion-based dark matter) to resolve lensing anomalies even in demanding cases such as HS 0810+2554, together with its success in reproducing other astrophysical observations, tilt the balance toward new physics invoking axions."[12][75]

Cosmic microwave background[edit]

Main article: Cosmic microwave background

Although both dark matter and ordinary matter are matter, they do not behave in the same way. In particular, in the early universe, ordinary matter was ionized and interacted strongly with radiation via Thomson scattering. Dark matter does not interact directly with radiation, but it does affect the cosmic microwave background (CMB) by its gravitational potential (mainly on large scales) and by its effects on the density and velocity of ordinary matter. Ordinary and dark matter perturbations, therefore, evolve differently with time and leave different imprints on the CMB.

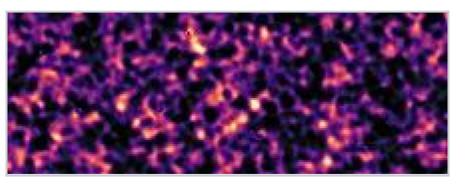
The cosmic microwave background is very close to a perfect blackbody but contains very small temperature anisotropies of a few parts in 100,000. A sky map of anisotropies can be decomposed into an angular power spectrum, which is observed to contain a series of acoustic peaks at near-equal spacing but different heights. The series of peaks can be predicted for any assumed set of cosmological parameters by modern computer codes such as CMBFAST and CAMB, and matching theory to data, therefore, constrains cosmological parameters.[76] The first peak mostly shows the density of baryonic matter, while the third peak relates mostly to the density of dark matter, measuring the density of matter and the density of atoms.[76]

The CMB anisotropy was first discovered by COBE in 1992, though this had too coarse resolution to detect the acoustic peaks. After the discovery of the first acoustic peak by the balloon-borne BOOMERanG experiment in 2000, the power spectrum was precisely observed by WMAP in 2003–2012, and even more precisely by the *Planck* spacecraft in 2013–2015. The results support the Lambda-CDM model.[77][78]

The observed CMB angular power spectrum provides powerful evidence in support of dark matter, as its precise structure is well fitted by the lambda-CDM model,[78] but difficult to reproduce with any competing model such as modified Newtonian dynamics (MOND).[78][79]

Structure formation[edit]

Main article: Structure formation



Dark matter map for a patch of sky based on gravitational lensing analysis of a Kilo-Degree survey.[80]

Structure formation refers to the period after the Big Bang when density perturbations collapsed to form stars, galaxies, and clusters. Prior to structure formation, the Friedmann solutions to general relativity describe a homogeneous universe. Later, small anisotropies gradually grew and condensed the homogeneous universe into stars, galaxies and larger structures. Ordinary matter is affected by radiation, which is the dominant element of the universe at very early times. As a result, its density perturbations are washed out and unable to condense into structure.[81] If there were only ordinary matter in the universe, there would not have been enough time for density perturbations to grow into the galaxies and clusters currently seen.

Dark matter provides a solution to this problem because it is unaffected by radiation. Therefore, its density perturbations can grow first. The resulting gravitational potential acts as an attractive potential well for ordinary matter collapsing later, speeding up the structure formation process.[81][82]

Bullet Cluster[edit]

Main article: Bullet Cluster

If dark matter does not exist, then the next most likely explanation must be that general relativity – the prevailing theory of gravity – is incorrect and should be modified. The Bullet Cluster, the result of a recent collision of two galaxy clusters, provides a challenge for modified gravity theories because its apparent center of mass is far displaced from the baryonic center of mass.[83] Standard dark matter models can easily explain this observation, but modified gravity has a much harder time,[84][85] especially since the observational evidence is model-independent.[86]

Type Ia supernova distance measurements[edit]

Main articles: Type Ia supernova and Shape of the universe

Type Ia supernovae can be used as standard candles to measure extragalactic distances, which can in turn be used to measure how fast the universe has expanded in the past.[87] Data indicates the universe is expanding at an accelerating rate, the cause of which is usually ascribed to dark energy.[88] Since observations indicate the universe is almost flat,[89][90][91] it is expected the total energy density of everything in the universe should sum to 1 ($\Omega_{tot} \approx 1$). The measured dark energy density is $\Omega_\Lambda \approx 0.690$; the observed ordinary (baryonic) matter energy density is $\Omega_b \approx 0.0482$ and the energy density of radiation is negligible. This leaves a missing $\Omega_{dm} \approx 0.258$ which nonetheless behaves like matter (see technical definition section above) – dark matter.[92]

Sky surveys and baryon acoustic oscillations[edit]

Main article: Baryon acoustic oscillations

Baryon acoustic oscillations (BAO) are fluctuations in the density of the visible baryonic matter (normal matter) of the universe on large scales. These are predicted to arise in the Lambda-CDM model due to acoustic oscillations in the photon–baryon fluid of the early universe, and can be observed in the cosmic microwave background angular power spectrum. BAOs set up a preferred length scale for baryons. As the dark matter and baryons clumped together after recombination, the effect is much weaker in the galaxy distribution in the nearby universe, but is detectable as a subtle (≈1 percent) preference for pairs of galaxies to be separated by 147 Mpc, compared to those separated by 130–160 Mpc. This feature was predicted theoretically in the 1990s and then discovered in 2005, in two large galaxy redshift surveys, the Sloan Digital Sky Survey and the 2dF Galaxy Redshift Survey.[93] Combining the CMB observations with BAO measurements from galaxy redshift surveys provides a precise estimate of the Hubble constant and the average matter density in the Universe.[94] The results support the Lambda-CDM model.

Redshift-space distortions[edit]

Large galaxy redshift surveys may be used to make a three-dimensional map of the galaxy distribution. These maps are slightly distorted because distances are estimated from observed redshifts; the redshift contains a contribution from the galaxy's so-called peculiar velocity in addition to the dominant Hubble expansion term. On average, superclusters are expanding more slowly than the cosmic mean due to their gravity, while voids are expanding faster than average. In a redshift map, galaxies in front of a supercluster have excess radial velocities towards it and have redshifts slightly higher than their distance would imply, while galaxies behind the supercluster have redshifts slightly low for their distance. This effect causes superclusters to appear squashed in the radial direction, and likewise voids are stretched. Their angular positions are unaffected. This effect is not detectable for any one structure since the true shape is not known, but can be measured by averaging over many structures. It was predicted quantitatively by Nick Kaiser in 1987, and first decisively measured in 2001 by the 2dF Galaxy Redshift Survey.[95] Results are in agreement with the lambda-CDM model.

Lyman-alpha forest[edit]

Main article: Lyman-alpha forest

In astronomical spectroscopy, the Lyman-alpha forest is the sum of the absorption lines arising from the Lyman-alpha transition of neutral hydrogen in the spectra of distant galaxies and quasars. Lyman-alpha forest observations can also constrain cosmological models.[96] These constraints agree with those obtained from WMAP data.