

Dark energy: how the paradigm shifted

The idea that our universe is dominated by mysterious dark energy was revealed by two paradigm-shifting studies of supernovae published in 1998. **Lucy Calder** and **Ofer Lahav** explain how the concept had in fact been brewing for at least a decade before – and speculate on where the next leap in our understanding might lie

Lucy Calder has an MSc in astrophysics from University College London and is currently a freelance writer in London, UK. **Ofer Lahav** is Perren Professor of Astronomy and head of astrophysics at University College London, e-mail lahav@star.ucl.ac.uk

Arguably the greatest mystery facing humanity today is the prospect that 75% of the universe is made up of a substance known as “dark energy”, about which we have almost no knowledge at all. Since a further 21% of the universe is made from invisible “dark matter” that can only be detected through its gravitational effects, the ordinary matter and energy making up the Earth, planets and stars is apparently only a tiny part of what exists. These discoveries require a shift in our perception as great as that made after Copernicus’s revelation that the Earth moves around the Sun.

Just 25 years ago most scientists believed that the universe could be described by Albert Einstein and Willem de Sitter’s simple and elegant model from 1932 in which gravity is gradually slowing down the expansion of space. But from the mid-1980s a remarkable series of observations was made that did not seem to fit the standard theory, leading some people to suggest that an old and discredited term from Einstein’s general theory of relativity – the “cosmological constant” or “lambda” (Λ) – should be brought back to explain the data.

This constant had originally been introduced by Einstein in 1917 to counteract the attractive pull of

gravity, because he believed the universe to be static and eternal. He considered it a property of space itself, but it can also be interpreted as a form of energy that uniformly fills all of space; if Λ is greater than zero, the uniform energy has negative pressure and creates a bizarre, repulsive form of gravity. However, Einstein grew disillusioned with the term and finally abandoned it in 1931 after Edwin Hubble and Milton Humason discovered that the universe is expanding. (Intriguingly, Isaac Newton had considered a linear force behaving like Λ , writing in his *Principia* of 1687 that it “explained the two principal cases of attraction”.)

Λ resurfaced from time to time, seemingly being brought back into cosmology whenever a problem needed explaining – only to be discarded when more data became available. For many scientists, Λ was simply superfluous and unnatural. Nevertheless, in 1968 Yakov Zel’dovich of Moscow State University convinced the physics community that there was a connection between Λ and the “energy density” of empty space, which arises from the virtual particles that blink in and out of existence in a vacuum. The problem was that the various unrelated contributions to the vacuum energy meant that Λ , if it existed, would be up to 120 orders of magnitude greater than observations suggested. It was thought there must be some mechanism that cancelled Λ exactly to zero.

In 1998, after years of dedicated observations and months of uncertainty, two rival groups of supernova hunters – the High- Z Supernovae Search Team led by Brian Schmidt and the Supernova Cosmology Project (SCP) led by Saul Perlmutter – revealed the astonishing discovery that the expansion of the universe is accelerating. A cosmological constant with a value different to that originally proposed by Einstein for a static universe – rebranded the following year as “dark energy” – was put forward to explain what was driving the expansion, and almost overnight the scientific community accepted a new model of the universe. Undoubtedly, the supernova observations were crucial in changing people’s perspective, but the key to the rapid acceptance of dark energy lies in the decades before.

At a Glance: The paradigm shift to dark energy

- Dark energy is a mysterious substance believed to constitute 75% of the current universe. Proposed in 1998 to explain why the expansion of the universe is accelerating, dark energy has negative pressure and causes repulsive gravity
- Data suggest that dark energy is consistent (within errors) with the special case of the cosmological constant (Λ) that Einstein introduced in 1917 (albeit for a different reason) and then abandoned. Λ can be interpreted as the vacuum energy predicted by quantum mechanics, but its value is vastly smaller than anticipated
- During the 20th century, Λ was reintroduced a number of times to explain various observations, but many physicists thought it a clumsy and *ad hoc* addition to general relativity
- The rapid acceptance of dark energy a decade ago was largely due to the work of researchers in the 1980s and early 1990s who concluded that, in spite of the prejudice against it, Λ was necessary to explain their data
- We still have no fundamental explanations for dark energy and dark matter. The next paradigm shift could be equally astonishing and we must be ready with open minds

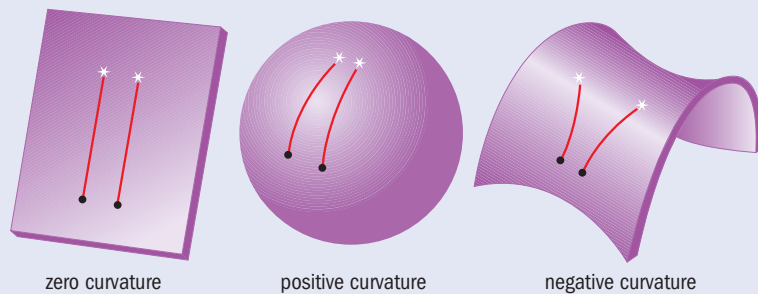


T Abbott/NOAO/AURA/NSF

Next steps

The Blanco 4 m telescope in Chile is being upgraded to conduct the five-year Dark Energy Survey, which in 2011 will begin mapping some 300 million galaxies. It is one of a number of new projects designed to learn more about dark energy.

1 Geometry of the universe



In flat, Euclidean space (left), light always travels in straight lines, but in the presence of a gravitational field, space is curved. This causes light rays to bend when passing near to a massive body such as the Sun. If the matter/energy density in the universe is greater than a critical value, then space will be positively curved, like a 3D version of the surface of a sphere, and light rays will converge on each other (middle). But if there is less matter and energy in the universe than the critical density, space will be negatively curved, like a saddle, and light rays will diverge (right).

Inflation and cold dark matter

Our story begins in 1980 when Alan Guth, who was then a postdoc at the Stanford Linear Accelerator Center in California, suggested a bold solution to some of the problems with the standard Big Bang theory of cosmology. He discovered a mechanism that would cause the universe to expand more, in a time interval of about 10^{-35} s just after the Big Bang, than it has done in the estimated 13.7 billion years since. The implications of this “inflation” were significant.

Einstein’s general theory of relativity, which has so far withstood every test made of it, tells us that the curvature of space is determined by the amount of matter and energy in each volume of that space – and that only for a specific matter/energy density is the geometry Euclidean or “flat” (figure 1). In inflationary cosmology, space is stretched so much that even if the geometry of the observable universe started out far from flat, it would be driven towards flatness – just as a small patch on the surface of a balloon looks increasingly flat as the balloon is blown up. By the mid-1980s a modified version of Guth’s model was overwhelmingly accepted by the physics community.

The problem was that while inflation suggested that the universe should be flat and so be at the critical density, the actual density – calculated by totting up the number of stars in a large region and estimating their

masses from their luminosity – was only about 1% of the required value for flatness. In other words, the observed mass density of conventional, “baryonic” material (i.e. protons and neutrons) appeared far too low. Moreover, the amount of baryonic matter in the universe is constrained by the theory of nucleosynthesis, which describes how light elements (hydrogen, helium, deuterium and lithium) formed in the very early universe. The theory can only match observations of the abundances of light elements if the density of baryonic matter is 3–5% of the critical density, with the actual value depending on the rate of expansion.

To make up the shortfall, cosmologists concluded that there has to be a lot of extra, invisible non-baryonic material in the universe. Evidence for this dark matter had been accumulating since 1932, when Jan Oort realized that the stars in the Milky Way are moving too fast to be held within the galaxy if the gravitational pull comes only from the visible matter. (At about the same time, Fritz Zwicky also found evidence for exotic hidden matter within clusters of galaxies.) Inevitably, the idea of dark matter was highly controversial and disputes over its nature rumbled on for the next 50 years. In particular, there were disagreements about how fast the dark-matter particles were moving and how this would affect the formation of large-scale structure, such as galaxies and galaxy clusters.

Then in March 1984, a paper by George Blumenthal, Sandra Faber, Joel Primack and Martin Rees convinced many scientists that the formation of structure in the universe was most likely if dark-matter particles have negligible velocity, i.e. that they are “cold” (*Nature* **311** 517). They found that a universe with about 10 times as much cold dark matter (CDM) as baryonic matter correctly predicted many of the observed properties of galaxies and galaxy clusters. The only problem with this “CDM model” was that the evidence pointed to the total matter density being low – barely 20% of the critical density. However, because of the constraints of inflation, most scientists hoped that the “missing mass density” would be found when measurements of dark matter improved.

Problems with the standard theory

At this point the standard cosmological model was a flat universe with a critical density made up of a small amount of baryonic matter and a majority of CDM. Apart from the fact that most of the matter was thought to be peculiar, this was still the Einstein–De Sitter model, and the theoretical prejudice for it was strong. Unfortunately for the inflation plus CDM model, it came with one very odd prediction: it said that the universe is no more than 10 billion years old, whereas, at the time, some stars were thought to be much older. For this reason, and because observations of the distribution of matter favoured a low mean mass density, the US cosmologists Michael Turner, Gary Steigman and Lawrence Krauss published a paper in June 1984 that investigated the possibility of a relic cosmological constant (*Phys. Rev. Lett.* **52** 2090).

The presence of Λ would cause a slight gravitational repulsion that acts against attractive gravity, meaning that the expansion of the universe would slow down less quickly. This implied that the universe was older than

Unfortunately for the inflation plus cold-dark-matter model, it came with one very odd prediction: that the universe is no more than 10 billion years old, whereas, at the time, some stars were thought to be much older

people thought at the time and so could accommodate its most ancient stars. Although Turner, Steigman and Krauss realized that Λ could solve some problems, they were – considering the constant’s chequered past – still wary about including it in any sensible theory. Indeed, the trio paid much more attention to the possibility that the additional mass density required for a flat universe was provided by relativistic particles that have been created recently (in cosmological terms) from the decay of massive particles surviving from the early universe.

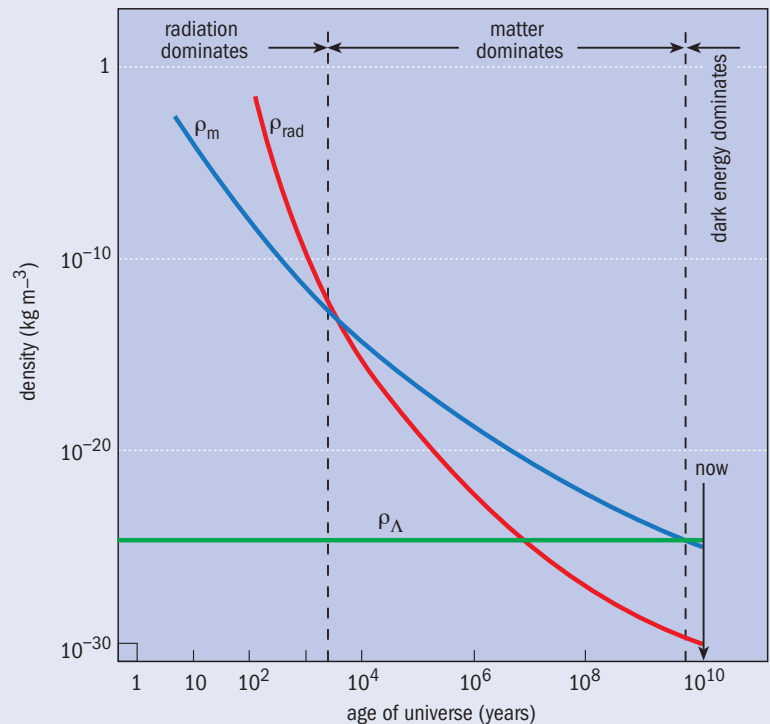
One of the other people to tentatively advocate the return of the cosmological constant was James Peebles from Princeton University, who at the time was studying how tiny fluctuations in the density of matter would grow, due to gravitational attraction, then ultimately collapse to form galaxies. Writing in a September 1984 paper in *The Astrophysical Journal* (284 439), he likewise deduced that the data pointed to an average mass density in the universe of about 20% of the critical value – but he went further and said it might be reasonable to invoke a non-zero cosmological constant in order to meet the new constraints from inflation. Although Peebles was fairly cautious about the idea, this paper helped bring Λ out of obscurity and began to pave the way to the acceptance of dark energy.

It would be tempting to think that the path to dark energy was now clear. However, astronomers realized that the mean mass density of the universe could still be as high as the critical density if there was a lot of extra dark matter hidden in the vast spaces, or voids, between clusters of galaxies. Indeed, when Marc Davis, George Efstathiou, Carlos Frenk and Simon White (following work by Nick Kaiser) ran computer simulations of the evolution of a universe dominated by CDM, they found dark matter and luminous matter were distributed differently, with more CDM in the voids. If galaxies formed only where the overall mean density was high, the simple Einstein–De Sitter flat cosmology could still agree with observations and we would not need to invoke the idea of dark energy at all.

But for those opposed to the idea of dark energy, the problem was that there was no sign of lots of missing mass in the voids. Indeed, when Lev Kofman and Alexei Starobinskii calculated the size of the tiny temperature variations in the cosmic microwave background (CMB) radiation, using different models of the universe, they found that adding Λ to the CDM model predicted fluctuations that would provide a better explanation of the observed distribution of galaxy clusters. Even if Λ were not included in the theory, observations in the late 1980s suggested that cosmological structure on very large scales could be more readily explained by a low-density universe and this, obviously, was incompatible with inflation.

Nevertheless, many people continued to believe that the idea of introducing another parameter, such as Λ , went against the principle of Occam’s razor, given that the data were still so poorly determined. It was not so much that physicists were deeply attached to the standard model, more that, like Einstein, they did not want to complicate the theory unnecessarily. Indeed, at the time, almost anything seemed preferable to the addition of Λ . As George Blumenthal, Avishai Dekel and Primack commented in 1988, introducing Λ would re-

2 Density evolution



In the Einstein–De Sitter model the universe consists only of components – matter and radiation – that give rise to ordinary, attractive gravity. As the universe expands, gravity slows the expansion and the mass density, ρ_m (blue), and radiation density, ρ_{rad} (red), of the universe decrease, at a rate given by Einstein’s general theory of relativity. In this model there is nothing special about the era in which humans have evolved. If, however, there is a non-zero cosmological constant, it will always keep the same value. So, it seems an unbelievable coincidence that we are living during the brief cosmological era when the dark energy density, ρ_Λ (green), and the mass density are about the same size. The repulsive gravity of Λ has only recently (in cosmological timescales) begun to dominate and is causing the expansion of the universe to accelerate. Adapted from R A Freedman and W J Kaufmann *Universe* (2005 W H Freeman)

quire “a seemingly implausible amount of fine-tuning of the parameters of the theory” (*Astrophys. J.* 326 539).

They instead proposed that a low-density, negatively curved universe with zero cosmological constant could explain the observed properties of galaxies, even up to large scales, if CDM and baryons contributed comparably to the mass density. This model, they admitted, conflicted with nucleosynthesis bounds, with inflation’s prediction that the universe is flat and with the small observational limit on the size of fluctuations in the CMB, but they believed that there were potential solutions. It seemed so much more aesthetically pleasing for Λ to simply be zero.

Surprising results

The quiet breakthrough came in 1990. Steve Maddox, Will Sutherland, George Efstathiou and Jon Loveday published the results of a study of the spatial distribution of galaxies, based on 185 photographic plates obtained by the UK Schmidt Telescope Unit in Australia (*Mon. Not. R. Astron. Soc.* 242 43). High-quality, glass copies of the plates were scanned using an automatic plate measuring (APM) machine that had recently been developed at Cambridge University by Edward Kibblewhite and his group. This remarkable

Future paradigm shifts?

Cosmologists still have no real idea what dark energy is and it may not even be the answer to what makes up the bulk of our universe. Here are a few potential paradigm shifts that we may have to contend with.

Violation of the Copernican principle

At the moment we assume that the Milky Way does not occupy any special location within the universe. But if we happen to be living in the middle of a large, underdense void, then it could explain why the type 1a supernovae (our strongest evidence for cosmic acceleration) look dim, even if no form of dark energy exists. However, requiring our galaxy to occupy a privileged position goes against the most basic underlying assumption in cosmology.

Is dark energy something other than vacuum energy?

Although vacuum energy is mathematically equivalent to Λ , the value predicted by fundamental theory is orders of magnitude larger than observations can possibly permit and there is no accepted solution to this problem. Many interesting ideas have been proposed, including time-varying dark energy, but even they do not address the “coincidence” problem of why the present epoch is so special.

Modifications to our understanding of gravity

It may be that we have to look beyond general relativity to a more complete theory of gravity. Exciting new developments in “brane” theory suggest the influence of extra spatial dimensions, but it is likely that the mystery of dark energy and cosmic acceleration will not be solved until gravity can successfully be incorporated into quantum field theory.

The multiverse

Λ can have a dramatic effect on the formation of structure in the universe. If Λ is too large and positive, it would have prevented gravity from forming large galaxies and life as we know it would never have emerged. Steven Weinberg and others used this anthropic reasoning to explain the problems with the cosmological constant and predicted a value for Λ that is remarkably close to what was finally observed. However, this use of probability theory predicted an infinite number of universes in which Λ takes on all possible values. Many scientists mistrust anthropic ideas because they do not make falsifiable predictions and seem to imply some sort of life principle or intention co-existing with the laws of physics. Nevertheless, string theory predicts a vast number of vacua with different possible values of physical parameters, and to some extent this legitimates anthropic reasoning as a new basis for physical theories.

survey – the largest in more than 20 years – covered more than 4300 square degrees of the southern sky and included about two million galaxies, looking deep into space and far back in time.

Astonishingly, the results from the APM galaxy survey did not match the standard CDM plus inflation model at all. On angular scales greater than about 3° , the survey provided strong evidence for the existence of galaxy clustering that was simply not predicted by the standard model. In 1990 Efstathiou, Sutherland and Maddox wrote a forthright letter to *Nature* (348 705), in which they argued that CDM and baryons accounted for only 20% of the critical density. The remaining 80%, they inferred, was provided by a positive cosmological constant, and this soon became known as the Λ CDM model.

The case for a low-density, CDM model came from the APM galaxy survey and a redshift survey of over 2000 galaxies detected by the infrared astronomical satellite (IRAS). The case for a positive cosmological constant now had several arguments in its favour: infla-

tion, which required a flat universe; the small size of the temperature fluctuations in the CMB; and the age problem. “A positive cosmological constant”, wrote Efstathiou, Sutherland and Maddox, “could solve many of the problems of the standard CDM model and should be taken seriously.” This was the strongest appeal yet made in favour of bringing Einstein’s Λ back into cosmology. The APM result for a low mass density of the universe was later confirmed by the 2dF Galaxy Redshift Survey and the Sloan Digital Sky Survey.

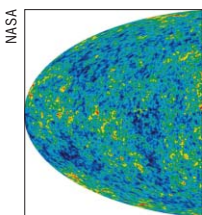
Return of the cosmological constant

Soon, others also began looking seriously at the case for Λ CDM. For example, in 1991 one of the present authors (OL), with Per Lilje, Primack and Rees, studied the implications of the cosmological constant on the growth of structure, and concluded it agreed with data available at the time (*Mon. Not. R. Astron. Soc.* 251 128). But researchers were still reluctant to embrace Λ fully. In 1992 Sean Carroll, William Press and Edwin Turner underlined the problems of considering Λ as the energy density of the vacuum – the coincidence problem (figure 2) and the fact that quantum mechanics predicts a far higher value than observations permit (*Ann. Rev. Astron. Astrophys.* 30 499). They pointed out that a flat universe plus Λ model effectively required the inclusion of non-baryonic CDM, which meant there were then two highly speculative components of the universe.

In 1993 an article appeared in *Nature* (366 429) on the baryon content of galaxy clusters, written by Simon White, Julio Navarro, August Evrard and Carlos Frenk. They studied the Coma cluster of galaxies, which is about 100 megaparsecs away from the Milky Way and contains more than 1000 galaxies. It is assumed, from satellite evidence, to be typical of all clusters rich in galaxies and its mass has three main components: luminous stars; hot, X-ray emitting gas; and dark matter. White and colleagues realized that the ratio of baryonic to total mass in such a cluster would be fairly representative of the ratio in the universe as a whole. Plugging in the baryon mass fraction from the nucleosynthesis model would then give a measure of the universe’s mean mass density.

After taking an inventory, using the latest data and computer simulations, the group concluded that the baryonic matter was a larger fraction of the total mass of the galaxy cluster than was predicted by a combination of the nucleosynthesis constraint and the standard CDM inflationary model. Baryons could have been produced in the cluster during its formation (by cooling, for example) but the number created would not have been enough to explain the discrepancy. The most plausible explanations were either that the usual interpretation of element abundances (nucleosynthesis theory) was incorrect, or that the mean matter density fell well short of the critical density. Once again, the standard CDM model, with the mass density equal to the critical density, was inadequate. The way to satisfy the constraint from inflation that the universe is flat would be to add a cosmological constant.

Luckily for the theorists, astronomers continued to refine and improve their observations through the development of new equipment and techniques. In par-



Cosmic clues

Fluctuations in the cosmic microwave background reveal that the universe is flat but that the mass density is far short of the critical density.

ticular, tiny variations in the CMB – as measured by NASA's Cosmic Background Explorer (COBE) satellite and later by the Wilkinson Microwave Anisotropy Probe (WMAP) satellite – dramatically showed that the mass density fell far short of the critical density and yet favoured a universe with an overall flat geometry. Something other than CDM must be providing the mass/energy needed to reach the critical value.

Although there were still a number of proposed variations on the CDM model, Λ CDM was the only model to fit all the data at once. It appeared that matter accounted for 30–40% of the critical density and Λ , as vacuum energy, accounted for 60–70%. Jeremiah Ostriker and Paul Steinhardt succinctly summed up the observational constraints in 1995 in an influential letter to *Nature* (377 600). The case rested strongly on measurements of the Hubble constant and the age of the universe, and results from the Hipparcos satellite in 1997 finally brought age estimates of the oldest stars down to 10–13 billion years.

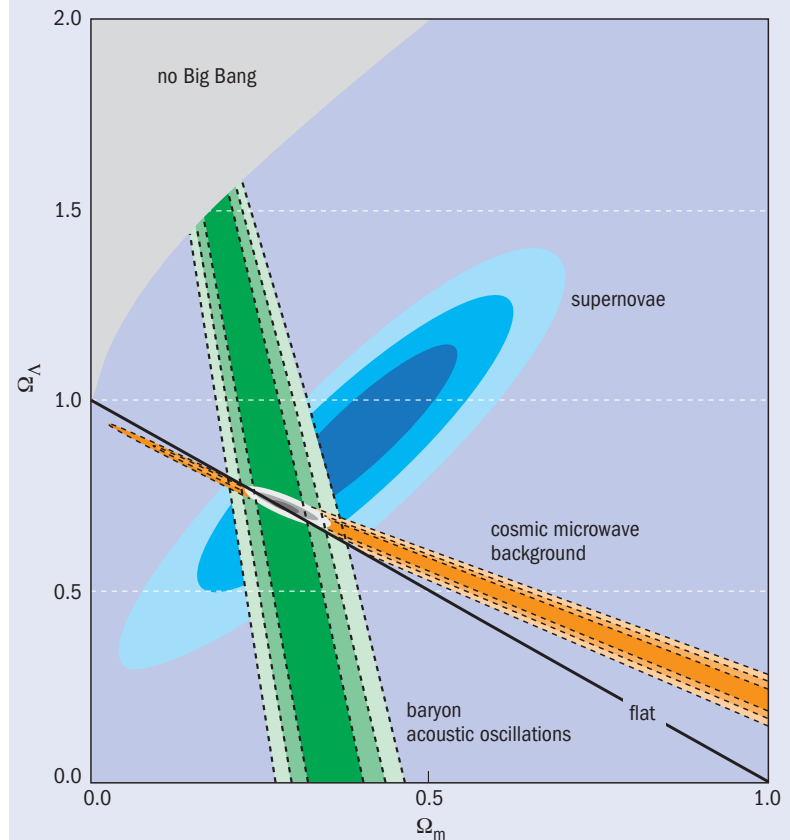
Most physicists were still reluctant to consider the idea that Λ should be brought back into cosmology, but the stage was now set for a massive shift in opinion (see “Dark energy” by Robert P Crease *Physics World* December 2007 pp19–22). Some people suggested other possibilities for the missing component, and the name “dark energy” was introduced by Michael Turner in 1999 to encompass all the ideas. When the supernova data of the High-Z and SCP teams indicated that the expansion of the universe is accelerating, the rapid embrace of dark energy was in large part due to the people who had argued for its return in the 1980s and 1990s.

Towards a new paradigm shift?

If recent history can teach us anything, it is to not ignore the evidence when it is staring us in the face. While the addition of two poorly understood terms – dark energy and dark matter – to Einstein's theory may spoil its intrinsic elegance and beauty, simplicity is not in itself a law of nature. Nevertheless, although the case for dark energy has been strengthened over the past dozen years, many people feel unhappy with the current cosmological model. It may be consistent with all the current data but there is no satisfactory explanation in terms of fundamental physics. As a result, a number of alternatives to dark energy have been proposed and it looks likely that there will be another upheaval in our comprehension of the universe in the decade ahead (see box opposite).

The whole focus of cosmology has altered dramatically and many astronomical observations now being planned or under way are mainly aimed at discovering more about the underlying cause of cosmic acceleration. For example, the ground-based Dark Energy Survey (DES), the European Space Agency's proposed Euclid space mission, and NASA's planned space-based Joint Dark Energy Mission (JDEM) will use four complementary techniques – galaxy clustering, baryon acoustic oscillations, weak gravitational lensing and type 1a supernovae – to measure the geometry of the universe and the growth of density perturbations (figure 3). The DES will use a 4 m telescope in Chile with a new camera that will peer deep into the southern sky

3 Evidence for cosmic acceleration



Data from type 1a supernovae (blue), baryon acoustic oscillations (green) and the cosmic microwave background (orange) provide independent information on the nature of dark energy. Only a very small region where the amount of matter, $\Omega_m \approx 0.25$, and the amount of dark energy, $\Omega_\Lambda \approx 0.75$, matches all three data sets. They converge at the value needed for a flat universe: $\Omega_m + \Omega_\Lambda = 1$. Adapted from M Kowalski et al. 2008 *Astrophys. J.* **686** 749

and will map the distribution of 300 million galaxies over 5000 square degrees (an eighth of the sky) out to a redshift of 2. The five-year survey involves over 100 scientists from the US, the UK, Brazil and Spain, and is due to begin in 2011.

The mystery of dark energy is closely connected to many other puzzles in physics and astronomy, and almost any outcome of these surveys will be interesting. If the data show there is no longer a need for dark energy, it will be a major breakthrough. If, on the other hand, the data point to a new interpretation of dark energy, or to a modification to gravity, it will be revolutionary. Above all, it is essential that astrophysics continues to focus on a diversity of issues so that individuals have the chance to do creative research and suggest new ideas. The next paradigm shift in our understanding may not come from the direction we expect. ■

More about: The paradigm shift to dark energy

L Calder and O Lahav 2008 Dark energy: back to Newton?

Astron. Geophys. **49** 1.13–1.18

B Carr and G Ellis 2008 Universe or multiverse? *Astron. Geophys.* **49** 2.29–2.33

E V Linder and S Perlmutter 2007 Dark energy: the decade ahead *Physics World* December pp24–30

P J E Peebles and B Ratra 2003 The cosmological constant and dark energy arXiv:astro-ph/0207347v2