It has been called the biggest problem in physics. The universe is accelerating thanks to an infinitesimal amount of negative energy. That sounds weird, but the key revelation of Einstein is that energy has mass. So the tiniest energy field has a mass. Now mass causes matter to collapse via gravity. It is an attractive force. But physics allows us to have negative energy. This seems weird, but that is exactly what potential energy is, in a gravity field. Tension in a rubber band is an example of potential energy. A catapult has potential energy.

The source of gravity is matter. But it is really the sum of matter and pressure. Normally pressure is positive, it pushes a system apart. But it can also be negative, as in the example of a rubber band under tension.

If positive pressure is dominant, it adds to the gravity. But any contribution by negative pressure does the opposite.

If there is more negative than positive energy in a system it will expand, not collapse, and even accelerate if the positive contribution to energy, coming say from the density of matter, which accounts for gravity, decreases.

There is energy in nothing, so the quantum theory asserts. Heisenberg’s uncertainty principle is at the core of quantum theory. Our inability to do a precise measurement of the position of an electron is because the act of observation itself, via transmission of a photon, induces an intrinsic uncertainty in position and time. This uncertainty is equivalent to tiny fluctuations in space and time. These fluctuations generate a tiny energy. Hence even the most perfect vacuum, as long as it is in space and time, contains an infinitesimal amount of energy. But in the early twentieth century, Einstein did not know about this. He developed a cosmology that contained only matter and energy, which acted as the source of gravity. For Einstein space, matter and gravity were interchangeable. Matter could be replaced by the curvature of space, and the motion of matter or rays of light were directed by the curvature of space.

Let’s go back to 1916, when Einstein developed his new theory of gravity, called general relativity because it generalized his earlier (1905) theory of special relativity that told us that physics measurements should be independent of the observer’s velocity, with a consequence that no material object could ever go faster than light. Einstein’s stated goal was to establish “whether the relativity concept can be followed through to the finish, or whether it leads to contradictions.”

The prediction that matter curved space was confirmed in 1919. A war-weary world awoke to triumphant newspaper headlines reporting the results of two expeditions that verified Einstein’s prediction that the positions of stars near the solar limb would be displaced when their positions were revealed during a total eclipse of the sun. Another major prediction of general relativity, that of black holes, was only confirmed a century later by the discovery of gravitational waves. Einstein regarded the universe as a test bed of his new theory. In 1917, he applied general relativity to the universe as a whole. He assumed the universe was the same everywhere and in all directions, a concept he called the cosmological principle. Einstein showed that his theory of gravity naturally led to an explanation of the universe of galaxies. To see how this works, let us symbolically denote Einstein’s equations of general relativity, the theory of gravity, as follows.

\[ \text{GRAVITY/CURVATURE OF SPACE} = \text{MASS/ENERGY}. \]

Gravity is defined by matter curving space, so Einstein told us in 1916, and mass-energy is its source. But there was a problem. Einstein soon realized that since the universe contains matter and gravity is an attractive force, this equation implies a collapsing universe. Mass and energy are also attractive, as far as their contribution to gravity goes. Indeed, one revelation of Einstein’s theory was even that ordinary pressure due to matter is primarily attractive in a gravity field. I could write this equation more rigorously in tensor form but it is not really necessary. Indeed it has been said that the tensor notation was designed to keep theologians out of modern cosmology.

Einstein was convinced the universe was stationary, as were essentially all of his contemporaries. He addressed the problem of the stability of the universe in his first paper on cosmology in 1917, and introduced a cosmological constant term, LAMBDA to counter gravity, because without this, the universe would be unstable and collapse.

Now we have in Einstein’s language:

\[ \text{GRAVITY/CURVATURE OF SPACE} - \text{LAMBDA} = \text{MASS/ENERGY}. \]

Think of LAMBDA as antigravity. The situation however soon became more complex when in the same year the
Dutch physicist Willem de Sitter noticed that the introduction of LAMBDA also allowed the universe to be empty. Mass and energy would be irrelevant. There was one bonus from de Sitter’s solution, as it predicted that radiation emitted by test particles in the empty universe would be redshifted. This result intrigued astronomers because of the emerging evidence being accumulated in particular by US astronomer Vesto Slipher at Flagstaff, Arizona that the spectra of the extragalactic nebulae, now called galaxies, seemed to be systematically shifted towards red wavelengths.

The next step in the saga was taken by the Russian mathematical physicist Alexander Friedmann in 1922 at the age of 34, who discovered expanding cosmological solutions of Einstein’s equations of general relativity. Einstein at first refused to accept the new solutions, suggesting a mathematical error, then dismissed them in 1923 as being unrealistic. In fact it was Einstein who had made the mathematical error in overlooking the expanding solution.

Unfortunately Friedmann died prematurely in 1925, and his work was ignored in the west. It was left to Georges Lemaître, a Belgian physicist and recently ordained priest, to republish in 1927, unaware of Friedmann’s work, the solutions to the expanding universe. There was one notable difference. Lemaître knew about Slipher’s results on the recession of the extragalactic nebulae as well as of Hubble’s measurements of their large distances. Lemaître presented the expanding universe models, and demonstrated that an expanding universe fit the astronomical data without any recourse to a cosmological constant. He talked on this topic at the 1927 Solvay conference on physics.

Einstein, in the audience, was unimpressed. He famously responded

“Vos calculs sont corrects, mais vôtre physique est abominable”.

When Hubble announced the discovery of the recession of the galaxies in 1929, Einstein realized that the expansion of space would stop its collapse. There was no more need for LAMBDA, which Einstein regretted. He is quoted by George Gamow in his posthumously-published autobiography as saying that Einstein “remarked that the introduction of the cosmological term was the biggest blunder of his life.”

By 1931, Einstein had become reconciled with the expanding universe model. A couple of years later, travelling with Lemaître on a lecture tour to California, Einstein commented on Lemaitre’s presentation “This is the most beautiful and satisfactory explanation of creation to which I have ever listened”. But Lemaître would never agree about the notion of “creation”: for him, the equations pointed to a physical beginning of the universe, and said nothing about any prior event. Which is more or less the modern scientific view of the beginning of the universe, at least until someone comes out with a satisfactory marriage of quantum theory and gravity.

Let us fast forward some 70 years. The acceleration of the universe was discovered in 1998. Two independent teams of mostly US astronomers were involved. Both found that distant supernovae were about 20% fainter than they should have been. Supernovae, or exploding stars, of a certain type were believed to be incredibly precise beacons, exploring via the decay of half a solar mass or so of radioactive nickel created in the precursor star’s collapse. Because of their immense luminosity, such supernovae are visible far away in distant galaxies.

The conclusion was not easy to justify. One had to be certain that the selection of distance supernovae gave objects identical to nearby ones. Then one could be certain of the standard bomb hypothesis, verified for nearby examples.

Next, one had to eliminate other explanations. For example, interstellar or intergalactic dust could be responsible for the diminution in light. The more distant supernovae were produced by a host population of stars that was younger and more metal-deficient. This could lead to appreciable evolutionary and environmental differences between nearby and remote supernovae. All of these objections were systematically eliminated. One conclusion remained. The systematic dimming of the otherwise identical distant supernovae could only be explained if the universe was accelerating. There was indeed a non-zero value of the cosmological constant. For this discovery, the Nobel Prize in physics was awarded in 2011 to Saul Perlmutter, Brian Schmidt and Adam Riess.

This is how LAMBDA was revived, but now it is interpreted as dark energy. Energy can accelerate if it is negative. The cosmological constant is interpreted as the energy of the vacuum, which turns out to be negative. Vacuum energy is identified, in Einstein’s equations of gravity, as the cosmological constant. It acts against gravity. The new cosmology looks like:

**GRAVITY/CURVATURE OF SPACE = LAMBDA + MASS/ENERGY.**

So now Einstein’s erstwhile demon was resuscitated and means that two-thirds of the mass-energy in the universe reappears as LAMBDA and acts with the expansion to counter gravity.

The discovery of acceleration was a revolution in cosmology. It meant that two-thirds of the mass energy in the universe could not be accounted for in classical physics. For this new component, it soon became clear, was present in the vacuum, where there was no matter of any consequence. It is an energy field that does not possess any attractive gravity of its own, rather it is repulsive.

Here is one reason why one may accept the tiny value of the cosmological constant. It is indeed a new field that
generates acceleration. But a similar phenomenon occurred long ago. It is called inflation, and results in an
accelerating phase of the universe. Indeed so much acceleration occurs that inflation results in a huge universe,
vastly larger than we can see with the largest telescopes. Inflation works because a constant energy field is
invented that temporarily dominates the energy density of the universe. If the energy is constant, the universe
expands exponentially rapidly. We accept inflation because it accounts for a number of deep questions. Let us
enumerate them. There are six reasons for our trust in inflation.

- The size of the universe. Why is it so large?
- The flatness of space. Why is space Euclidean?
- The seeds of galaxies. Where did the fluctuations come from?
- The infinitesimal temperature fluctuations in the microwave background. Why is their strength described by
  just a few numbers?
- Why do these fluctuations continue to such large scales? These scales had no causal connection at early
times.
- Why are these fluctuations completely random? We see no patterns in their distribution.

All of these questions are answered by the theory of inflation. We cannot prove the theory is correct. In fact we
cannot prove any theory is true. But it is strongly supported by the astronomical evidence.

Perhaps inflation is not the ultimate theory. Critics maintain that it is built on shaky foundations. But it is the best
we have, and merits our trust. It certainly is our closest approximation to the truth. It describes the universe at
$10^{-36}$ second after the beginning. And a similar, but much, much smaller, energy field to that postulated for
inflation reasserted itself, smaller a few billion years later. This is why the cosmological constant should not come
as a total surprise.

There is a key difference, however. The scale of inflation is a natural scale in terms of particle physics. It
corresponds, at least in order of magnitude, to the energy scale where the fundamental forces of nature, apart
from gravity, first became unified. However the energy scale corresponding to the cosmological constant is
smaller than any natural scale from particle physics by some 120 factors of 10. There is no physics explanation
for the scale of the cosmological constant. We have two choices. We can accept it as just one more number or
constant of nature that describes the universe. Perhaps one day an ultimate theory will emerge, sometimes
dubbed the Theory of Everything, that will specify the fundamental constants of nature, and in particular the
value of the cosmological constant. Or we can accept one of the precepts of string theory, that the constants of
nature may vary according to the precise 3-dimensional universe that results after compactification from ten
spatial dimensions. We then appeal to inflation theory in its generic form, which experts believe predicts an
infinite number of separate universes. According to this argument, inflation is eternal, and it can occur at any
time and place.

This leads inevitably to the concept of the multiverse, the topic of a previous lecture in this series. There are an
infinite number of universes in the multiverse. None of these can communicate with each other, so we cannot
exactly prove this theory, or even detect another universe. But the lack of empirical evidence has not deterred
many physicists from believing in the existence of the multiverse.

Almost all of the universes in the multiverse will have values of the fundamental constants of nature, and in
particular the cosmological constant, that are inimical to the existence of life. They would be too hot, too cold, or
too short-lived. There will be just one universe in the multiverse, in this Goldilocks perspective, where stars and
planets are sufficiently long-lived and the constants of nature are not too different from those we measure, and
where life would emerge. And here we are, with the small value measured for the cosmological constant
explained anthropically, by a sort of cosmic selection effect.

The universe is just beginning to accelerate and will continue expand ever more rapidly. The acceleration began
about seven billion years ago, that is when we detect the first deviations from Hubble’s law of the linear relation
between distance and redshift or recession velocity. In 140 billion years from now, our universe will have
accelerated so much that our Milky Way and its close companions will be alone in the visible universe. All more
distant galaxies will have accelerated away outside our horizon. Only if the acceleration were to stop would we
renew contact with our neighbours. This happened with the first episode of inflation when the seed fluctuations
destined to form galaxies reentered the horizon. But we have no reason to believe we will be similarly saved in
the distant future. Of course the earth will be destroyed in 5 billion years when the sun dies as a red giant, so
there will be no terrestrial witnesses in the remote future.

Though it is encouraging to think that dwarf stars, some 20 percent of the mass of the sun, will live for a trillion
years or more. Such stars are far more common in the Milky Way than are solar-type stars. And many are
believed to possess habitable planets, with a triplet of exoplanets being recently discovered around the nearby
dwarf star TRAPPIST 1, some 29 light years away. So cosmology will be very different for any civilization
surviving in the distant future.