MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Physics Department

Physics 8.286: The Early Universe November 19, 2009
Prof. Alan Guth

Lecture Notes 8
THE COSMOLOGICAL CONSTANT

INTRODUCTION:

Much excitement has been generated since January 1998 over observations that show that the expansion of the universe today is accelerating, rather than decelerating. Two groups of astronomers,* with a total of 52 astronomers in the two groups, have reported evidence for such an acceleration, based on observations of distant ($z \lesssim 1$) Type Ia supernova explosions, which are used as standard candles. (Note that "Ia" is pronounced "one-A," not "eeya.") The first announcement was made at the AAS meeting in January of 1998, leading to news articles in *Science* on January 30 and February 27, 1998, and in *The New York Times* on March 3, March 8, April 21, and May 5, 1998. On May 15 one of the two groups (the The High Z Supernova Search Team) posted a paper on the web titled "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant".† The other group (The Supernova Cosmology Project) submitted its findings to the web on December 8, 1998.‡ *Science* magazine officially proclaimed this to be the "Breakthrough of the Year" for 1998.

The evidence for a cosmological constant has stood up firmly for the eleven years since 1998, and in fact it has gotten stronger. A particularly strong piece of evidence was uncovered in early 2001 by A.G. Riess, P.E. Nugent, et al.¶, who discovered in data from the Hubble Space Telescope a supernova at the colossal redshift of 1.7. Many cosmologists including me were skeptical in 1998, but now

---

* One group is the Supernova Cosmology Project, based at Lawrence Berkeley Laboratory and headed by Saul Perlmutter. Their web page is
http://www-supernova.lbl.gov/
The other group is the The High Z Supernova Search Team, led by Brian Schmidt, with web page
http://cfa-www.harvard.edu/cfa/oir/Research/supernova/HighZ.html.

† http://arXiv.org/abs/astro-ph/9805201, later published as Riess *et al.*, *Astronomical Journal* **116**, No. 3, 1009 (1998).

‡ "Measurements of Ω and Λ from 42 High-Redshift Supernovae," http://arXiv.org/abs/astro-ph/9812133, later published as Perlmutter *et al.*, *Astrophysical Journal* **517**:565–586 (1999).

¶ "The Farthest Known Supernova: Support for an Accelerating Universe and a Glimpse of the Epoch of Deceleration," astro-ph/0104455, Riess *et al.*, *Astrophysical Journal* **560**, 49–71 (2001).

essentially all of us are all convinced that the universe has a nonzero cosmological constant, or perhaps something called *quintessence,* which has very nearly the same effect. (Quintessence refers to a slowly evolving scalar field that permeates the universe and fills it with a nearly uniform energy density — we'll get back to that idea when we talk about inflation near the end of the course.) Since no one is sure what exactly is driving this acceleration, the term "dark energy" has been invented to describe the stuff that is driving the acceleration, whatever it might be. A cosmological constant is the simplest explanation, and that is what will be discussed in this set of lecture notes.

BACKGROUND:

The cosmological constant was first proposed by Albert Einstein in 1917, when he was trying for the first time to apply his newly invented theory of general relativity to the universe as a whole. At the time he believed that the universe was static, since it appeared static and there was no evidence to the contrary. However, when he worked out the consequences of his theory, he discovered the equation that we first wrote as Eq. (7.38):

$$\frac{d^2a}{dt^2} = -\frac{4\pi}{3}G\left(\rho + \frac{3p}{c^2}\right)a \ , \tag{8.1}$$

where a is the scale factor, t is time, G is Newton's gravitational constant, ρ is the mass density, p is the pressure, and c is the speed of light. Taking $\rho > 0$ and $p \approx 0$, Einstein was forced to the conclusion that $d^2a/dt^2 < 0$, so a static ($a = constant$) solution did not exist. The problem, essentially, was that gravity is an attractive force, so an initially static universe would collapse.

Einstein's solution was to modify his equations of general relativity by adding an extra term, which he called the cosmological term, which could create a repulsive force that could be adjusted in strength so that it could prevent the universe from collapsing. The coefficient of this term was called the cosmological constant and assigned the symbol Λ (capital Greek lambda).

Einstein's static model seemed viable for about a decade, but during the 1920s astronomers discovered that the universe was not static after all. In 1929 Edwin Hubble published his famous paper announcing what we now know as Hubble's law. Einstein was quick to accept Hubble's findings, and discarded his cosmological term as unwarranted.

COSMOLOGICAL EQUATIONS WITH A COSMOLOGICAL CONSTANT:

Although Einstein did not look at the cosmological constant this way, from a modern perspective the cosmological constant is interpreted as an energy density attributed to the vacuum. Since everything that we see can be described as particles moving through the vacuum, the vacuum energy density becomes a uniform contribution to the total energy, at all points in space and at any time. The relation between Einstein's original symbol Λ and the vacuum energy density $u_{\rm vac}$, or the vacuum mass density $\rho_{\rm vac}$, is given by

$$u_{\rm vac} = \rho_{\rm vac} c^2 = \frac{\Lambda c^4}{8\pi G} \; . \tag{8.2}$$

Einstein's constant Λ has the units of $(\text{length})^{-2}$, while $u_{\rm vac}$ and $\rho_{\rm vac}$ of course have the usual units for energy density and mass density. When the contribution of such a vacuum energy density is included in the Einstein field equation, which describes how matter causes spacetime to curve, it produces a term identical to the cosmological term added by Einstein. The pressure that corresponds to this vacuum energy can be obtained by applying the equation of energy conservation, using the fact that the energy density of the vacuum is fixed. On Problem 1 of Problem Set 7 (2009), we learned that conservation of energy in a Robertson-Walker universe takes the form

$$\frac{d}{dt}\left(a^3 \rho c^2\right) = -p \frac{d}{dt}(a^3) \; , \tag{8.3}$$

or equivalently

$$\dot\rho = -3\frac{\dot a}{a}\left(\rho + \frac{p}{c^2}\right) \; , \tag{8.4}$$

where the overdot denotes a time derivative. Setting $\dot\rho_{\rm vac} = 0$ gives

$$p_{\rm vac} = -\rho_{\rm vac} c^2 = -\frac{\Lambda c^4}{8\pi G} \; . \tag{8.5}$$

The relation between the pressure and energy density is the same as the relation that we will later discuss for the false vacuum that is responsible for driving the accelerated expansion of the inflationary universe model. From Eq. (8.1), one can see that a negative pressure can drive an acceleration. We must add the contributions of the vacuum energy density and pressure to the right-hand side, so for clarity we will use the symbols ρ_n and p_n to denote the mass density and pressure of normal matter, where *normal* refers to all forms of energy other than the cosmological constant. One then has

$$\begin{aligned} \frac{d^2a}{dt^2} &= -\frac{4\pi}{3}G\left(\rho_n + \frac{3p_n}{c^2} + \rho_{\text{vac}} + \frac{3p_{\text{vac}}}{c^2}\right)a \\ &= -\frac{4\pi}{3}G\left(\rho_n + \frac{3p_n}{c^2} - 2\rho_{\text{vac}}\right)a \ , \end{aligned} \tag{8.6}$$

where we used Eq. (8.5) to eliminate p_{vac}.

We learned in Lecture Notes 7 that the first order Friedmann equation is not modified by pressure, so it is still written as it was first written, as Eq. (4.24):

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}G\rho - \frac{kc^2}{a^2} \ . \tag{8.7}$$

Since the right-hand side depends only on ρ, we find the contribution of the vacuum energy density by replacing ρ by $\rho_n + \rho_{\text{vac}}$:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}G(\rho_n + \rho_{\text{vac}}) - \frac{kc^2}{a^2} \ . \tag{8.8}$$

Using Eqs. (8.6) and (8.8), we can be more precise about what it means to live in an accelerating universe. From Eq. (8.6), we see that $\ddot{a}$ can be positive if the ρ_{vac} term is positive and dominates the right-hand side, so under these circumstances one says that universe accelerates, meaning that the function $a(t)$ accelerates. Since the physical distance ℓ_p to a galaxy at coordinate distance ℓ_c is given by

$$\ell_p(t) = a(t)\ell_c \ ,$$

we see that in an accelerating universe, the relative velocity between galaxies increases with time.

On the other hand, from Eq. (8.8) we see that this acceleration does not necessarily mean that H increases with time. We can more easily see the behavior of Eq. (8.8) if we replace ρ_n by $\rho_m + \rho_{\text{rad}}$, where ρ_m is the mass density of nonrelativistic (pressureless) matter and ρ_{rad} is the mass density of radiation:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}G\,(\underbrace{\rho_m}_{\propto\frac{1}{a^3(t)}} + \underbrace{\rho_{\text{rad}}}_{\propto\frac{1}{a^4(t)}} + \rho_{\text{vac}}) - \frac{kc^2}{a^2} \ . \tag{8.9}$$

For an open ($k < 0$) or flat ($k = 0$) universe, the right-hand side of Eq. (8.9) contains only positive terms, each of which decreases as the universe expands. Thus

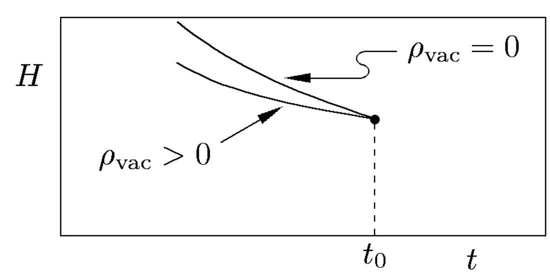
H decreases monotonically in such a universe, even if the universe is accelerating. The matter, radiation, and curvature terms all approach zero as $a \to \infty$, so asymptotically

$$H = \frac{\dot{a}}{a} \xrightarrow[a\to\infty]{} \sqrt{\frac{8\pi}{3} G \rho_{\text{vac}}} \qquad \text{from above.} \tag{8.10}$$

Note that $H = \dot{a}/a$ can decrease even when $\dot{a}$ is increasing, as long as a is increasing faster. For a closed universe ($k > 0$) it is possible for H to increase as the universe expands, but this happens only if the last term of Eq. (8.9) is large enough in magnitude so that it dominates the rate of change of H. Our universe could be closed, but the last term of Eq. (8.9) is known to be small, so H for our universe is decreasing. If it is true that the acceleration is caused by vacuum energy density, then Eq. (8.10) describes the asymptotic future of our universe, whether it is open, closed, or precisely flat. However, we should certainly keep in mind that predictions about the infinite future are very dicey. It is possible, for example, that the state that we call the vacuum might not really be stable, but might instead decay into a lower energy state after 10^{1000} years, falsifying our prediction.

THE COSMOLOGICAL CONSTANT AND THE AGE OF THE UNIVERSE:

One effect of a positive cosmological constant is an increase in the age of the universe that is inferred from a given value of the Hubble constant. This effect can be understood qualitatively by remembering that the cosmological constant causes the universe to accelerate. Suppose, then, that we calculated the age of the universe as we learned in Lecture Notes 5, assuming that there was no cosmological constant. Then suppose that we add a vacuum energy term, keeping fixed the current value of the Hubble expansion rate H_0 and the current mass density of nonrelativistic matter and radiation. The new energy contribution adds a positive term to $\ddot{a}$, which means that H has not been falling as fast as it had in the previous $\rho_{\text{vac}} = 0$ calculation. Then, as can be seen from the following sketch,



the Hubble expansion rate in the past would be lower in the new calculation than it was in the first calculation. The slower decrease in H would mean that it takes longer for H to reach its present value, since in both models H starts at infinity at the instant of the big bang. Similarly, the lower value of H in the past would mean that it takes longer for the universe to reach its present mass density. Thus, the new calculation implies a universe which is older than we had calculated in the absence of a cosmological constant.

Quantitatively, we can calculate the age of the universe from Eq. (8.9). To be completely explicit about the time-dependence of each term, we write

$$\begin{aligned} \rho_m(t) &= \left[\frac{a(t_0)}{a(t)}\right]^3 \rho_{m,0} \\ \rho_{\rm rad}(t) &= \left[\frac{a(t_0)}{a(t)}\right]^4 \rho_{\rm rad,0} \\ \rho_{\rm vac}(t) &= \rho_{\rm vac,0} \ . \end{aligned} \tag{8.11}$$

Here we are using the convention that a subscript 0 denotes the present value of any quantity, so for example $\rho_{m,0}$ denotes the present value of the mass density of nonrelativistic matter. Each of the above equations reflects the known dependence on $a(t)$ for each contribution to the mass density, with the constant of proportionality written so that $\rho_X(t_0) = \rho_{X,0}$, for each type of matter X. Mass densities are usually tabulated as fractions of the critical density,

$$\rho_c = \frac{3H^2}{8\pi G}, \tag{8.12}$$

using the convention that for each type of mass density X,

$$\Omega_X \equiv \rho_X/\rho_c \ . \tag{8.13}$$

So, we rewrite Eqs. (8.11) by replacing each $\rho_{X,0}$ by $\Omega_{X,0}\rho_{c,0}$:

$$\begin{aligned} \rho_m(t) &= \frac{3H_0^2}{8\pi G}\left[\frac{a(t_0)}{a(t)}\right]^3 \Omega_{m,0} \\ \rho_{\rm rad}(t) &= \frac{3H_0^2}{8\pi G}\left[\frac{a(t_0)}{a(t)}\right]^4 \Omega_{\rm rad,0} \\ \rho_{\rm vac}(t) &= \frac{3H_0^2}{8\pi G}\Omega_{\rm vac,0} \ . \end{aligned} \tag{8.14}$$

Defining

$$x \equiv \frac{a(t)}{a(t_0)} \ , \tag{8.15}$$

so that x varies from 0 to 1 as the universe evolves from the big bang to the present, Eq. (8.9) can be rewritten as

$$\left(\frac{\dot a}{a}\right)^2 = H_0^2\left(\frac{\Omega_{m,0}}{x^3} + \frac{\Omega_{\rm rad,0}}{x^4} + \Omega_{\rm vac}\right) - \frac{kc^2}{a^2} \ . \tag{8.16}$$

It is convenient to rewrite the curvature term in the same form as the other terms, by defining

$$\Omega_{k,0} \equiv -\frac{kc^2}{a^2(t_0)H_0^2} \ , \tag{8.17}$$

so

$$\begin{aligned} H^2 = \left(\frac{\dot a}{a}\right)^2 = \left(\frac{\dot x}{x}\right)^2 &= H_0^2\left(\frac{\Omega_{m,0}}{x^3} + \frac{\Omega_{\rm rad,0}}{x^4} + \Omega_{\rm vac} + \frac{\Omega_{k,0}}{x^2}\right) \\ &= \frac{H_0^2}{x^4}\left(\Omega_{m,0}x + \Omega_{\rm rad,0} + \Omega_{\rm vac,0}x^4 + \Omega_{k,0}x^2\right) \ . \end{aligned} \tag{8.18}$$

By specializing this formula to $t = t_0$, for which $x = 1$, one finds $1 = \Omega_{m,0} + \Omega_{\rm rad,0} + \Omega_{\rm vac,0} + \Omega_{k,0}$, so

$$\Omega_{k,0} = 1 - \Omega_{m,0} - \Omega_{\rm rad,0} - \Omega_{\rm vac,0} \ . \tag{8.19}$$

$\Omega_k > 0$ for an open universe, $\Omega_k < 0$ for a closed universe, and $\Omega_k = 0$ for a flat universe. The present age of the universe can then be found by taking the square root of Eq. (8.18),

$$\frac{\dot x}{x} = \frac{H_0}{x^2}\sqrt{\Omega_{m,0}x + \Omega_{\rm rad,0} + \Omega_{\rm vac,0}x^4 + \Omega_{k,0}x^2} \ , \tag{8.20}$$

or

$$x\frac{dx}{dt} = H_0\sqrt{\Omega_{m,0}x + \Omega_{\rm rad,0} + \Omega_{\rm vac,0}x^4 + \Omega_{k,0}x^2} \ . \tag{8.21}$$

This equation can be rearranged as

$$dt = \frac{1}{H_0}\frac{x\,dx}{\sqrt{\Omega_{m,0}x + \Omega_{\rm rad,0} + \Omega_{\rm vac,0}x^4 + \Omega_{k,0}x^2}} \ , \tag{8.22}$$

which can be integrated over the range of x from the big bang to the present to give

$$t_0 = \frac{1}{H_0}\int_0^1 \frac{x\,dx}{\sqrt{\Omega_{m,0}x + \Omega_{\rm rad,0} + \Omega_{\rm vac,0}x^4 + \Omega_{k,0}x^2}} \ . \tag{8.23a}$$

The above form is probably the easiest to integrate, but for some purposes it is useful to rewrite it by changing variables of integration to z, where

$$1 + z = \frac{a(t_0)}{a(t)} = \frac{1}{x} \ .$$

The integral then becomes

$$t_0 = \frac{1}{H_0} \int_0^\infty \frac{dz}{(1+z)\sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{\text{rad},0}(1+z)^4 + \Omega_{\text{vac},0} + \Omega_{k,0}(1+z)^2}} \ . \tag{8.23b}$$

In this form one could also find the "look-back time" to any particular redshift z by stopping the integration at that point:

$$t_{\text{look-back}}(z) = \frac{1}{H_0} \int_0^z \frac{dz'}{(1+z')\sqrt{\Omega_{m,0}(1+z')^3 + \Omega_{\text{rad},0}(1+z')^4 + \Omega_{\text{vac},0} + \Omega_{k,0}(1+z')^2}} \ . \tag{8.24}$$

The look-back time is defined as the time interval between the era that we observe at redshift z and the present.

The general case of the integrals in Eqs. (8.23) and (8.24) can be computed only by numerical integration, but various special cases can be carried out analytically. The case of a matter-dominated universe ($\Omega_{\text{rad}} = \Omega_{\text{vac}} = 0$) was done in Lecture Notes 5. The case of a flat universe composed of nonrelativistic matter and vacuum energy (i.e., $\Omega_{\text{rad}} = \Omega_k = 0$, $\Omega_m + \Omega_{\text{vac}} = 1$) can also be integrated analytically, yielding

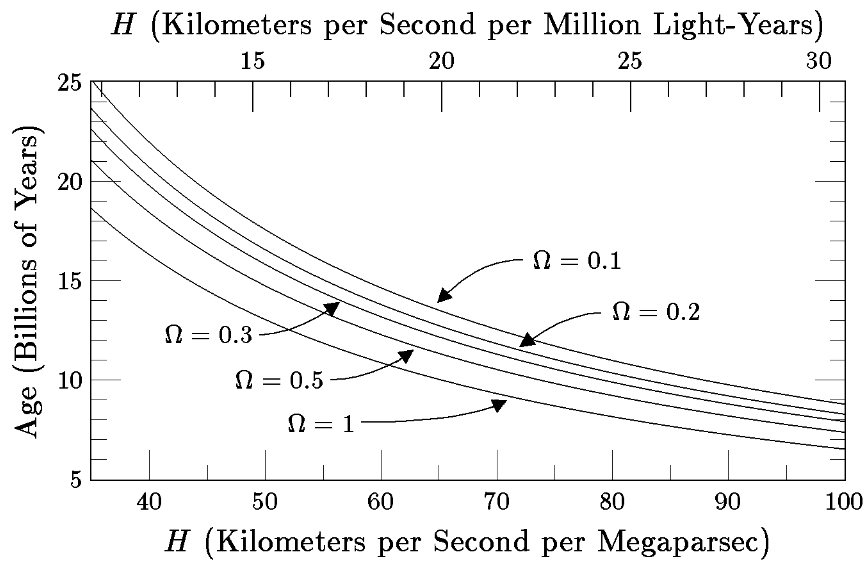
$$t_0 = \begin{cases} \dfrac{2}{3H_0} \dfrac{\tan^{-1}\sqrt{\Omega_{m,0} - 1}}{\sqrt{\Omega_{m,0} - 1}} & \text{if } \ \Omega_{m,0} > 1, \ \Omega_{\text{vac}} < 0 \\ \dfrac{2}{3H_0} & \text{if } \ \Omega_{m,0} = 1, \ \Omega_{\text{vac}} = 0 \\ \dfrac{2}{3H_0} \dfrac{\tanh^{-1}\sqrt{1 - \Omega_{m,0}}}{\sqrt{1 - \Omega_{m,0}}} & \text{if } \ \Omega_{m,0} < 1, \ \Omega_{\text{vac}} > 0 \ . \end{cases} \tag{8.25}$$

Note that inverse hyperbolic tangents can also be expressed in terms of logarithms, so the answer for the $\Omega_{m,0} < 1$ case can also be written as

$$t_0 = \frac{2}{3H_0} \frac{\ln\left(\sqrt{1 - \Omega_{m,0}} + 1\right) - \ln\sqrt{\Omega_{m,0}}}{\sqrt{1 - \Omega_{m,0}}} \ . \tag{8.26}$$

Although Eq. (8.25) expresses t_0 in terms of three different expressions, the function is actually continuous, so the value for $\Omega_{m,0} = 1$ can be obtained as the limit of either the expression for $\Omega_{m,0} > 1$ or the expression for $\Omega_{m,0} < 1$.

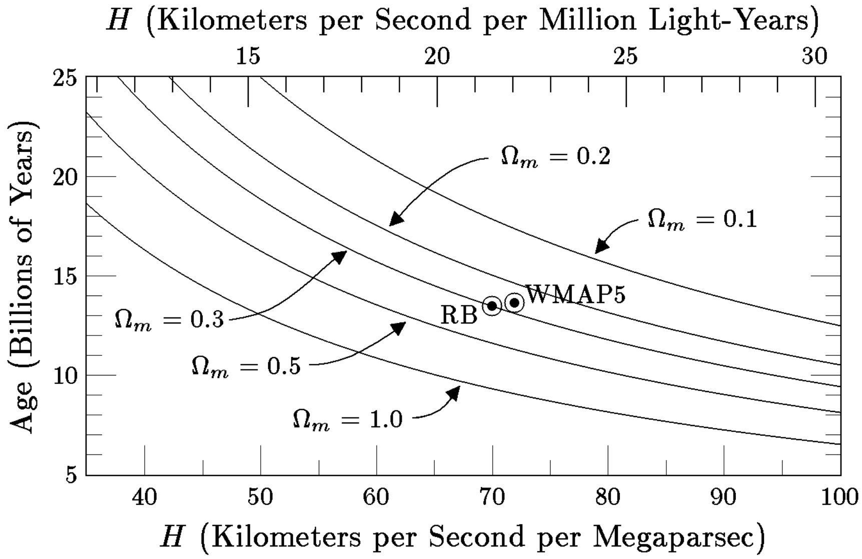
A graph of the age of the universe as a function of the Hubble constant, for a matter-dominated universe **without** a cosmological constant, is given by Eq. (5.46) and is shown here:



The age of an open ($\Omega < 1$), closed ($\Omega > 1$), or flat ($\Omega = 1$) universe containing only nonrelativistic matter.

For the parameter choice of $H_0 = 72\ \text{km}{\cdot}\text{s}^{-1}{\cdot}\text{Mpc}^{-1}$ (the Hubble Space Key Project value) and $\Omega = 1$, this gives $t_0 = 9.1 \times 10^9$ yr, which is much younger than the 11.2 billion year mininum age determined by Krauss and Chaboyer, based on age of the oldest stars, as discussed on pp. 2-3 of Lecture Notes 5.

However, this discrepancy of age estimates goes away when one attributes approximately 70% of Ω to a cosmological constant. A graph of the age of the universe, for a flat universe composed of nonrelativistic matter and a cosmological constant, is given by Eq. (8.25) and is shown below. Note that Ω_m refers to the mass density of nonrelativistic matter only. For all the model universes shown on this graph, the total Ω (including nonrelativistic matter and vacuum mass density) is one, which is in accord with the predictions of the simplest inflationary models (which will be discussed at the end of the term).



The age of a flat universe containing nonrelativistic matter and vacuum energy.

The graph also shows two data points: the point labeled RB refers to the Ryden Benchmark Model (from Barbara Ryden, *Introduction to Cosmology*), and the point labeled WMAP3 is the best fit model to the WMAP 5-year data set.* The parameters associated with these two models are as follows:

Parameters	**Ryden Benchmark**	**WMAP 5-Year Best Fit**
$\boldsymbol{H_0}$	70	71.9 ± 2.7 km·s^{-1}·Mpc^{-1}
Baryonic matter $\boldsymbol{\Omega_b}$	0.04	0.0441 ± 0.0030
Dark matter $\boldsymbol{\Omega_{\rm dm}}$	0.26	0.214 ± 0.027
Total matter $\boldsymbol{\Omega_m}$	0.30	0.258 ± 0.027
Vacuum energy $\boldsymbol{\Omega_{\rm vac}}$	0.70	0.742 ± 0.030

The WMAP 5-year data best fit is generally regarded as the most reliable estimate of cosmological parmeters that we currently have. Using the parameters from this

* G. Hinshaw et al., "Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, and Basic Results," Table 7, 'WMAP Only' column, *Ap. J. Supp.* **180,** 225-245 (2009), http://lambda.gsfc.nasa.gov/product/map/dr3/map_bibliography.cfm.

table, Eq. (8.25) gives a current age t_0 of 13.5 billion years for the Ryden Benchmark model, and 13.7 billion years for the WMAP 5-year best fit model. The Hinshaw et al. WMAP paper gives a best fit value for the age of the universe of 13.69 ± 0.13 billion years, where the quoted uncertainty of 0.9% is considerably smaller than would be obtained by compounding the uncertainties of the parameters shown in the table: 3.8% for H_0 and 10.5% for Ω_m. Thus, the WMAP group is asserting that their data set allows a determination of the age of the universe with much greater precision than it can determine the other parameters.

(To get some feeling for the stability of these numbers, we can compare with the previous WMAP data set, the 3-year data.* In that paper the numbers were reported as $H_0 = 73.5 \pm 3.2\ \text{km}\cdot\text{s}^{-1}\cdot\text{Mpc}^{-1}$, $\Omega_b = 0.041$, $\Omega_{\rm dm} = 0.196$, $\Omega_m = 0.237 \pm 0.034$, $\Omega_{\rm vac} = 0.763 \pm 0.034$, and $t_0 = 13.73^{+0.16}_{-0.15}$ billion years.)

We will discuss the physics underlying the WMAP measurements near the end of the term, but for now it is worth mentioning that WMAP refers to the Wilkinson Microwave Anisotropy Probe, a NASA satellite experiment dedicated to measuring the anisotropies (i.e., nonuniformities) of the cosmic microwave background radiation. While the CMB is uniform in all directions to an accuracy of one part in 100,000, the nonuniformities can nonetheless be measured to a high degree of accuracy, providing important information about the early universe. WMAP was launched on June 30, 2001, and is still in orbit and taking data. Its first year data set was released in February 2003, and its three-year data set was released in March 2006, after a rather long period of analysis by the WMAP group. The five-year data set was released in March 2008. The orbit of WMAP is unique, called the L2 Lagrange Point. The satellite is located at a position approximately 1.5 million km from Earth, in a direction opposite to the Sun. It follows the orbit of the Earth around the Sun once per year, always maintaining its position along a radial line drawn from the Sun through the Earth. L2 is an ideal location for astronomy, because the satellite can look outward away from the Sun, so at any time it can view half of the sky with no interference from the Sun, Earth, or Moon. Over the course of one year, the entire sky can be viewed under these ideal conditions.

It may seem strange that a satellite measuring the anisotropies of the CMB can give us values for parameters such as H_0 and Ω_m, but such parameters can be extracted if one has a theoretical prediction for the anisotropies that depends on these parameters. However, such a theoretical model exists, and it fits the data extraordinarily well. We will come back to this topic at the very end of the course.

---

* D.N. Spergel et al., "Wilkinson Microwave Anisotropy Probe (WMAP) Three Year Observations: Implications for Cosmology," Table 5, 'WMAP Only' column, *Ap. J. Supp.* **170,** 377 (2007), http://arxiv.org/abs/astro-ph/0603449v2.

Returning to the calculation of the age of the universe, we should discuss briefly the role of radiation, which was omitted from the calculations described above. Since we discovered in Lecture Notes 7 that the radiation-dominated period was finished at $t \approx 50,000$ years, we would not expect radiation to have much effect on the age calculation for a 13.7 billion-year-old universe. Continuing our discussion of parameters, the best estimate of the CMB temperature comes from a final analysis of the COBE data done in 1999*, which gives a value of $T_0 = 2.725 \pm 0.002$ K. Using Eq. (7.48) for u, combined with $g = 2$ for photons, $g = 21/4$ for neutrinos (Eqs. (7.54) and (7.56)), and $T_\nu = (4/11)^{1/3}\, T_\gamma$ (Problem Set 8, Problem 1, 2009), one has

$$\begin{aligned} \rho_{\rm rad,0} &= \left[2 + \frac{21}{4}\left(\frac{4}{11}\right)^{4/3}\right] \frac{\pi^2}{30}\frac{(kT_\gamma)^4}{\hbar^3 c^5} \\ &= 7.804 \times 10^{-34}\ \text{g/cm}^3\ , \end{aligned} \tag{8.27}$$

which can be combined with the WMAP 5-year value for H_0 listed above to give

$$\Omega_{\rm rad,0} = 4.154 \times 10^{-5}\, h_0^{-2} = 8.0 \times 10^{-5}\ . \tag{8.28}$$

These results agree with the numbers given in Eqs. (7.22) and (7.23).

One can include $\Omega_{\rm rad,0}$ in the age calculation of Eq. (8.23) by doing the integral numerically. If one includes it with the WMAP parameters, adjusting $\Omega_{\rm vac,0}$ to keep the universe exactly flat, one finds that the age is decreased by 6 million years, which is beyond the level of accuracy of the calculation.

THE HUBBLE DIAGRAM — RADIATION FLUX VERSUS REDSHIFT:

The claims that the cosmological constant is nonzero are based on the Hubble diagram, the graph which shows the measurements of the radiation flux of sources as a function of their redshift z. To understand how the cosmological constant affects this diagram, we need to derive the formula for the received radiation flux of a specified source, in a model universe which includes a cosmological constant. In principle we need to consider closed, flat, and open universes, but I will show the calculation in detail only for the case of a closed universe. The open-universe case is very similar, so I will merely describe the differences and state the answer for this case. The flat universe is the borderline case between open and closed, so

---

* J.C. Mather, D.J. Fixsen, R.A. Shafer, C. Mosier, and D.T. Wilkinson, "Calibrator Design for the COBE Far-Infrared Absolute Spectrophotometer (FIRAS)," *Ap. J.* **512,** 511 (1999), http://arxiv.org/abs/astro-ph/9810373.

it can be treated as a limiting case of either open or closed universes, or it can be done as a separate calculation.

The Robertson-Walker metric for a closed universe was given as Eq. (6.34):

$$ds^2 = -c^2\,dt^2 + a^2(t)\left\{\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta\,d\phi^2\right)\right\} \;. \qquad (6.34)$$

The cosmological constant will affect the evolution of $a(t)$, but the form of the metric was determined by the symmetries of homogeneity and isotropy, and will not be changed.

We will be interested in tracing the trajectories of photons traveling along radial lines, so for this purpose it will be useful to introduce the radial coordinate ψ, defined by

$$\sin\psi \equiv \sqrt{k}\,r \;.$$

One finds

$$d\psi = \frac{\sqrt{k}\,dr}{\cos\psi} = \frac{\sqrt{k}\,dr}{\sqrt{1-kr^2}} \;.$$

The metric then simplifies to

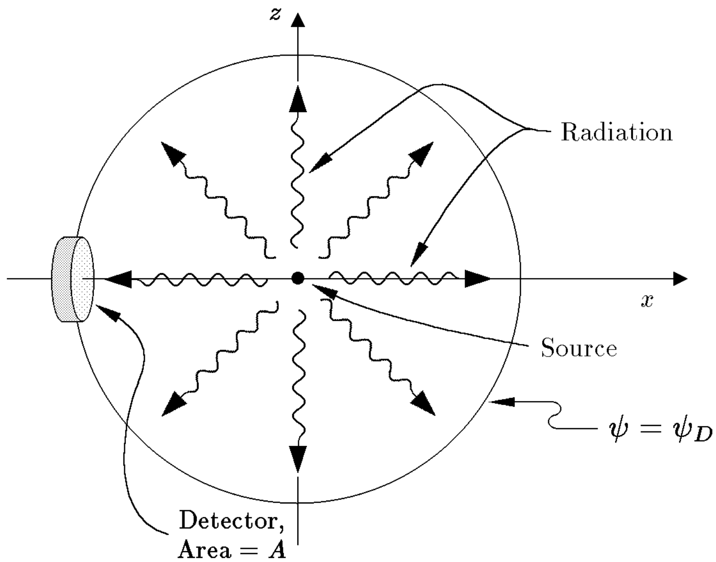
$$ds^2 = -c^2\,dt^2 + \tilde{a}^2(t)\left\{d\psi^2 + \sin^2\psi\left(d\theta^2 + \sin^2\theta\,d\phi^2\right)\right\} \;, \qquad (8.29)$$

where the new scale factor $\tilde{a}(t)$ is related to the scale factor $a(t)$ by

$$\tilde{a}(t) \equiv \frac{a(t)}{\sqrt{k}} \;. \qquad (8.30)$$

This form of the metric is useful for investigating radial motion, because the radial part of the metric is very simple. (You might recall that the closed universe metric was constructed in Lecture Notes 6 by first considering a sphere in 4 Euclidean dimensions. The coordinate ψ defined here is precisely the same as the angle ψ that was used in that construction—it is the angle between the w-axis and a line joining the origin of the 4-dimensional coordinate system to the point in question.)

Now consider a source emitting photons with total power P. Choose the comoving coordinate system so that the source is at the origin, $\psi = 0$. The radial coordinate of the detector, on Earth, will be ψ_D. The diagram also shows a sphere at the same radial coordinate, ψ_D:



Since the speed of light is independent of angle, all the photons that left the source at some particular time t_S are arriving at the $\psi = \psi_D$ sphere at the present time t_0. To calculate the power received by the detector, we need to know what fraction of those photons hit the detector. The fraction is simply the area of the detector divided by the area of the sphere, or

$$\text{fraction} = \frac{\text{area of detector}}{\text{area of sphere}} = \frac{A}{4\pi \tilde{a}^2(t_0)\sin^2\psi_D} \; .$$

(The area of the sphere at radial coordinate ψ_D is given by $4\pi\tilde{a}^2(t_0)\sin^2\psi_D$, because the part of the metric (8.29) that depends on $d\theta$ and $d\phi$ is equal to $\tilde{a}^2(t)\,\sin^2\psi$ times the metric for a sphere of unit radius.) The power hitting the detector is further reduced by one factor of

$$1 + z_S = \frac{a(t_0)}{a(t_S)} \; , \tag{8.31}$$

because the frequency, and hence the energy, of each photon is reduced by this factor. In addition, the power is reduced by another factor of $(1 + z_S)$ because the rate of arrival of photons is reduced by this factor. Thus, if P is the power that the source was emitting at time t_S, then the power received by the detector today is

$$P_{\text{received}} = \frac{P}{(1+z_S)^2}\frac{A}{4\pi\tilde{a}^2(t_0)\sin^2\psi_D} \; . \tag{8.32}$$

The flux is given by

$$J = \frac{P_{\text{received}}}{A} = \frac{P}{4\pi(1+z_S)^2 \tilde{a}^2(t_0)\sin^2\psi_D} \ . \tag{8.33}$$

Eq. (8.33) is the answer to our question, but it is not yet expressed in terms of useful variables — we cannot look up the values of $\tilde{a}(t_0)$ or ψ_D in standard tables, so we need to express them in terms of variables that we can look up. Specifically, we will be able to express the right-hand side of Eq. (8.33) in terms of P, z_S, H_0, and the various contributions to the current value of Ω.

Using the definition of $\tilde{a}(t)$ given by Eq. (8.30), one sees that its present value $\tilde{a}(t_0)$ can be related to $\Omega_{k,0}$, which was defined by Eq. (8.17). With a little rearranging, the relation becomes

$$\tilde{a}(t_0) = \frac{cH_0^{-1}}{\sqrt{-\Omega_{k,0}}} \ . \tag{8.34}$$

(Note that for a closed universe, $\Omega_k < 0$, so the denominator could have been written as $\sqrt{|\Omega_{k,0}|}$.)

Finally, we need to evaluate ψ_D, which we expect to be determined by the redshift z_S and cosmological parameters:

$$\psi_D = \psi(z_S) \ , \tag{8.35}$$

where $\psi(z_S)$ is defined as the ψ coordinate traversed by radial light pulses that are now reaching us with redshift z_S. These light pulses travel along a null trajectory, where the word "null" means that $ds^2 = 0$. Given the metric (8.29), a radial null trajectory is described by

$$0 = -c^2\,dt^2 + \tilde{a}^2(t)\,d\psi^2 \quad \Longrightarrow \quad \frac{d\psi}{dt} = \frac{c}{\tilde{a}(t)} \ . \tag{8.36}$$

The evolution equation for $\tilde{a}(t)$ is identical to the evolution equation for $a(t)$ that was given as Eq. (8.18):

$$H^2 = \left(\frac{\dot{\tilde{a}}}{\tilde{a}}\right)^2 = \frac{H_0^2}{x^4}\left(\Omega_{m,0}x + \Omega_{\text{rad},0} + \Omega_{\text{vac},0}x^4 + \Omega_{k,0}x^2\right) \ , \tag{8.37}$$

where

$$x = \frac{a(t)}{a(t_0)} = \frac{\tilde{a}(t)}{\tilde{a}(t_0)} \ . \tag{8.38}$$

Since the light pulse travels from time $t = t_S$ to $t = t_0$, the radial coordinate that it traverses can be found by integrating Eq. (8.36) to find

$$\psi(z_S) = \int_{t_S}^{t_0} \frac{c}{\tilde{a}(t)}\,dt \ . \tag{8.39}$$

Since we are hoping to express the answer in terms of the redshift of the source z_S, it is useful to change the variable of integration to z, where

$$1 + z = \frac{\tilde{a}(t_0)}{\tilde{a}(t)} \ . \tag{8.40}$$

Then

$$dz = -\frac{\tilde{a}(t_0)}{\tilde{a}(t)^2}\dot{\tilde{a}}(t)\,dt = -\tilde{a}(t_0)H(t)\,\frac{dt}{\tilde{a}(t)} \ . \tag{8.41}$$

The integration becomes

$$\psi(z_S) = \frac{1}{\tilde{a}(t_0)}\int_0^{z_S} \frac{c}{H(z)}\,dz \ . \tag{8.42}$$

In this expression we can replace $\tilde{a}(t_0)$ using Eq. (8.34), and we can replace $H(z)$ using Eq. (8.37), recognizing that $x = 1/(1+z)$. This gives our final expression for $\psi(z_S)$:

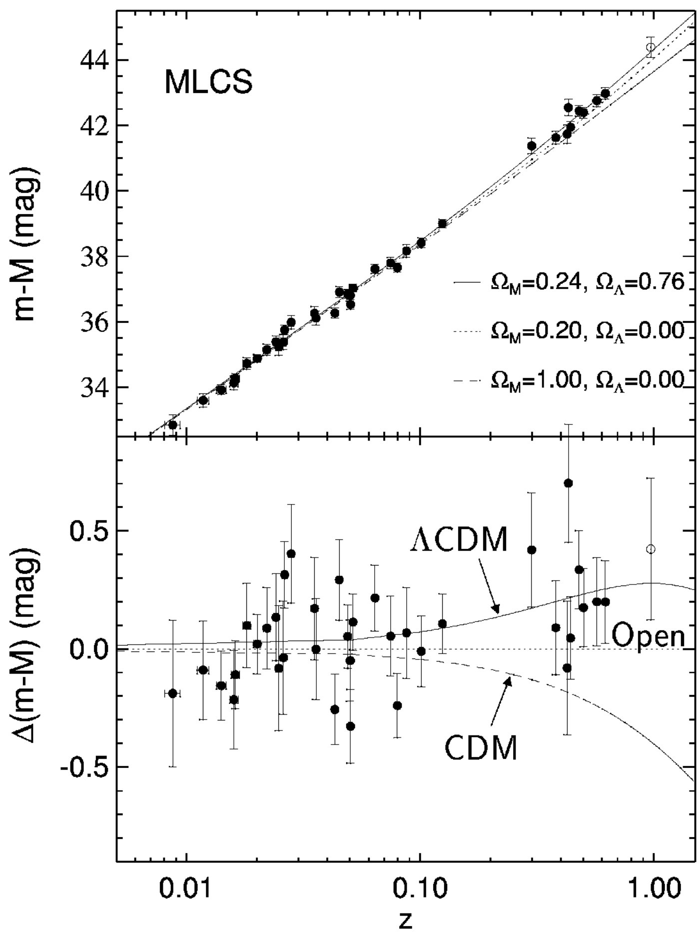
$$\boxed{\begin{aligned}\psi(z_S) &= \sqrt{|\Omega_{k,0}|} \\ &\qquad \times \int_0^{z_S} \frac{dz}{\sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{\text{rad},0}(1+z)^4 + \Omega_{\text{vac},0} + \Omega_{k,0}(1+z)^2}} \ .\end{aligned}} \tag{8.43}$$

We can now go back to our answer, expressed as Eq. (8.33), and eliminate the unwanted variables. $\tilde{a}(t_0)$ is replaced using Eq. (8.34), and $\sin^2\psi_D$ can be replaced by $\sin^2\psi(z_S)$, giving

$$\boxed{J = \frac{PH_0^2|\Omega_{k,0}|}{4\pi(1+z_S)^2c^2\sin^2\psi(z_S)} \ ,} \tag{8.44}$$

where $\psi(z_S)$ is given by Eq. (8.43).

For a sample of the recent data, I include a graph of the Hubble diagram from the paper by Riess *et al.* (1998) that was cited at the beginning of these lecture notes:



Shown at the top is a graph of magnitude vs. redshift for a sample of supernovae. The vertical axis represents distance as inferred from the brightness, with larger distances at the top.* Each increase of 5 magnitudes corresponds to the brightness decreasing by a factor of 100, so one magnitude corresponds to a factor of 2.512,

---

* More precisely, $m - M$ is the distance modulus, which is related to the *luminosity distance* d_L by

$$m - M = 5 \log_{10} \left(\frac{d_L}{1 \text{ Mpc}} \right) + 25 \ ,$$

where the luminosity distance is defined as the distance at which the object would have to be located to result in the observed brightness, if we were living in a static Euclidean universe. In such a universe the energy flux J at a distance d would be given by

$$J = \frac{P}{4\pi d^2} \ ,$$

and an increase by a tenth of a magnitude corresponds to about a 10% descrease in brightness. Shown on the same graph are three theoretical curves, calculated from Eq. (8.44), using different theoretical parameters. The lowest curve represents a matter-dominated flat universe (CDM = "cold dark matter"), with no cosmological constant. The middle curve represents an open matter-dominated universe, with $\Omega_m = \Omega_{\rm tot} = 0.2$, a value which was observationally plausible before the presence of dark energy became convincing. The uppermost curve, which seems to be the best fit to the data, represents a flat universe which includes nonrelativistic matter and a cosmological constant (ΛCDM = cosmological constant + cold dark matter), with the nonrelativistic matter comprising 0.24 of the critical density, and the vacuum mass density of the cosmological constant comprising 0.76 of the critical density. These ratios were chosen as a best fit to the data, within the class of flat models with these two components. Note that these numbers agree perfectly with the WMAP 5-year best fit model, even though the observations used to determine the parameters are completely different. The initials "MLCS" at the top stand for "Multi-Color Light Curve Shape," a method of analysis that the authors employed to compensate for small differences in the brightness of the supernovae based on the duration of the light output. The graph at the bottom shows the same data, but in a way that visually emphasizes the differences between the three curves. On this graph the middle curve is plotted as a straight line, and the other curves are shown as offsets relative to the middle curve. Note that the curves differ by two or three tenths of a magnitude, indicating that the brightness differences are only 20 to 30%. That is, the measured brightnesses of the distant supernovae are 20 to 30% dimmer than would be expected in the open universe $\Omega_m = \Omega_{\rm tot} = 0.2$ model.

The connection between this effect and acceleration is a little hard to see, but it can be seen most clearly if one thinks about the appearance of supernovae with a fixed magnitude, and hence a fixed distance as measured by the luminosity. Then the measured points lie to the left of the open universe $\Omega_m = \Omega_{\rm tot} = 0.2$ model, which means that the redshift is lower than expected. Lower redshift means smaller velocities, and hence the universe in the past was expanding more slowly than

---

so the luminosity distance is given by

$$d_L = \sqrt{\frac{P}{4\pi J}} \ .$$

Thus,

$$m - M = -\frac{5}{2}\log_{10}\left(\frac{4\pi J \times (1\ \mathrm{Mpc})^2}{P}\right) + 25 \ .$$

expected. If the universe in the past was expanding more slowly than expected on the basis of the current expansion rate, it means that some accelerating influence must have been at work.

The graph may not appear to be very conclusive, but nonetheless the data, if taken at face value, is statistically very significant. Especially when this data is combined with the data from the other group, the possibility that we are seeing a statistical fluke is very small. Nonetheless, there are possible systematic errors that are hard to evaluate. The observed effect is simply the fact that distant supernovae, at a given redshift, appear slightly dimmer (by about 20 to 30%) than expected. One alternative explanation might be that there is dust that obscures our view, causing the supernovae to appear dimmer than they really are. The problem with this explanation is that most forms of dust distort the spectrum of the light, absorbing more of the shorter wavelengths, resulting in a "reddening" of the received light. Since this reddening is not observed, the dust must be "gray," the word that is used to describe a filter that absorbs equally across the spectrum. It is physically possible for dust to be gray if the grains are large enough, but such dust is not known to exist. Another difficulty with the dust hypothesis is that if most of the dust is located in the host galaxy of the supernova, as one would expect, then the amount of absorption should depend on where the supernova is located within the galaxy. This would in turn produce scatter in a graph like the one shown above, while the amount of scatter seen is consistent with the known sources of uncertainty. Another totally different explanation for the observations is the possibility that it is caused by galactic evolution. Heavy elements are produced in stars, so 5 billion years ago there was a noticeably lower abundance of heavy elements in galaxies. If the lower abundance of heavy elements could lead to dimmer supernovae explosions, then this evidence for a cosmological constant would disappear. However, astronomers have looked hard to find any visible differences between the early supernovae at large redshift and the recent supernovae nearby, and so far they have found nothing significant. Further, in the nearby universe there are galaxies with a range of abundances of heavy elements, and this has not been observed to produce a difference in the brightness of supernova explosions.

On balance, I think it is fair to say that currently most cosmologists regard the supernova data as persuasive, but not, by itself, irrefutable. However, there is also increasingly strong evidence from observations of the cosmic microwave background radiation, which we have summarized earlier in these lecture notes in terms of the WMAP results. These measurements provide a measurement of the amount of vacuum energy that agrees beautifully with the supernova results. In addition, they provide very strong evidence that the universe is flat. There is also much evidence from extragalactic astronomy that there is not enough matter in the universe, even including the dark matter, to make up the critical density that is required by general relativity for a flat universe. If this is right, then vacuum energy becomes the most straightforward explanation of where the mass density is hidden. Also, as we have discussed, the inclusion of vacuum energy makes the calculation of the age of the universe from the Hubble expansion rate consistent with the estimated ages of the oldest stars. With all the evidence combined, there seems to be no alternative to the belief that about 70% of the mass density of the universe is in the form of dark energy — negative pressure material that is either vacuum energy (also called a cosmological constant), or perhaps "quintessence," which we will discuss later.

THE PARTICLE PHYSICS OF A COSMOLOGICAL CONSTANT:

While the observational evidence for a cosmological constant seems strong, the underlying physics of a cosmological constant remains very mysterious. From the point of view of modern particle physics it is not at all strange that the vacuum should have a nonzero mass density, but it is very hard to imagine any reason why it should have a value anywhere near the value that is being observed.

According to modern particle physics, the vacuum is actually a very complicated state. It is defined as the state with the lowest possible energy density, but it is not "empty" in any conventional sense. For example, the electric and magnetic fields are constantly fluctuating in the vacuum, because the uncertainty principles of quantum theory do not allow them to remain at zero value. These fluctuations give a positive contribution to the vacuum energy. The calculation of this contribution is formally infinite, since each mode of oscillation contributes, and there are an infinite number of modes at arbitrarily short wavelengths. It seems reasonable, however, to truncate this infinite sum at what is called the Planck length,

$$\lambda_{\text{Planck}} = \sqrt{\frac{\hbar G}{c^3}} = 1.6 \times 10^{-33} \text{ cm} . \tag{8.45}$$

This is the scale at which quantum gravity effects are believed to become important, and even the very notion of classical space presumably breaks down. With this cutoff the answer becomes finite, but it is more than 120 orders of magnitude larger than the energy density associated with the observed cosmological constant! You will have a chance to work this out in detail on Problem Set 8.

There are known negative contributions to the vacuum energy density as well, coming from fermions, such as the electron. Fermions give a huge negative contribution to the energy density of the vacuum, an effect that can be understood intuitively in terms of a metaphor known as the "Dirac sea". That is, the Dirac equation which describes relativistic electrons has both positive energy and negative energy solutions. These solutions are viewed as the possible energy levels of particles. When one of the positive energy levels becomes occupied by a particle, the overall energy of the state increases. But the overall energy is lowered whenever one of the negative energy levels becomes occupied by a particle. The vacuum, therefore, is the state in which all the negative enery levels are filled. The occupation of one of the positive energy levels then corresponds to an electron, which can be present in an otherwise vacuum state. The overall energy can also be increased by vacating one of the negative energy levels, leaving behind a "hole in the Dirac sea." Such a hole corresponds to a positron, the antiparticle of the electron.

This "filling of the Dirac sea" gives a negative energy density to the vacuum, since the filling of each negative energy level decreases the overall energy. Like the positive contribution of the electromagnetic field oscillations, the magnitude of this contribution is formally infinite. When it is cut off at the Planck length it becomes finite and comparable in magnitude to the positive contribution of the electromagnetic field.

There is a possibility that these huge positive and negative contributions could somehow cancel each other almost but not quite exactly, but no one knows why. In the absence of any real understanding, physicists had until recently assumed that the positive and negative contributions most likely cancel exactly because of some unknown symmetry principle. Even if that were the case, it would of course be an important challenge to understand why. If there really is a cosmological constant, then it looks like the positive and negative contributions to the vacuum energy density cancel to an accuracy of 120 decimal places, but miss in the 121st decimal place. Or maybe we are just looking at this all wrong.

At the present time, the cosmological constant problem is perhaps the most significant outstanding problem in our understanding of fundamental physics.