Carnegie Observatories Astrophysics Series, Vol. 2:
Measuring and Modeling the Universe, 2004
ed. W. L. Freedman (Cambridge: Cambridge Univ. Press)

Measuring and Modeling the Universe: A Theoretical Perspective

ROGER D. BLANDFORD
Kavli Institute for Particle Astrophysics and Cosmology, Stanford University

Abstract

A brief summary of a symposium discussing the measuring and modeling of the Universe is presented from a theoretical perspective. This whole field has been transformed by the cumulative success of microwave background observations and the large-scale galaxy surveys. After a long history of confusion and contradiction, the standard model of the Universe that has emerged is internally consistent with prior theoretical expectation, save for the now dominant feature of dark energy that accounts for over 70 percent of the energy density of the Universe. The dark matter that is required to account for local dynamics within galaxies and clusters is found to be present on the cosmological scale and has a quite plausible identity as a stable supersymmetric particle that has been invoked within particle physics. The prospects for significantly improving the precision of these measurements over the next decade and for identifying a dark matter particle are good. Considerable theoretical challenges also lie ahead. An immediate task is to reexamine the physics that fixes the two numbers that most strongly characterize our Universe, the entropy per baryon and the amplitude of the fluctuation spectrum. Another goal is to probe the very early Universe and to provide further corroboration of the principle that the Universe must have undergone a period of inflationary expansion. The hardest challenge of all is to account for the apparent acceleration of the Universe and its interpretation in terms of a dark energy density some 120 orders of magnitude smaller than the natural, Planckian value.

1.1 Introduction

Scepticism is the chastity of the intellect and it is shameful to surrender it too soon or to the first comer. G. Santayana

It is both an honor and a pleasure to participate in the celebration of the Centennial of the Carnegie Observatories. Collectively, the world of astronomy, and more recently that of physics, owes an immense debt to the Carnegie Institution and its distinguished astronomical faculty for discovering an impressive fraction of what we now take for granted about the Universe and its major constituents. There is also a personal debt that I would like to acknowledge. As a child growing up in the north Birmingham (England) suburb of Erdington, I obtained my very first books about astronomy and physics from the local public library which had been endowed with a grant from Andrew Carnegie.*

The title of this symposium highlights two important and, I believe, distinguishable fea-

* Malcolm Longair tells a similar story in his introductory talk.

tures of the recent explosive progress in cosmology. The first is that we are now making pretty secure and reproducible kinematic measurements. This is an activity that is sometimes called “cosmography” and it has long been associated with the Carnegie Observatories, especially through the work of Allan Sandage and Wendy Freedman. Cosmography is basically surveying and is not especially concerned with the dynamics of the Universe and the social histories of its inhabitants. The second feature is encapsulated in the somewhat pejorative term “modeling.” This is what we do when we are ignorant of the physics, either because it is too complicated to consider from first principles, as is true of galaxy evolution, or because it is too mysterious, which is true of the phenomenon described by the term “dark energy” or “quintessence.” That we are now measuring with such confidence is the giant step forward; that we are once again reduced to modeling much of the Universe is a prudent step backward. Put another way, when we are confident of the physics as, happens with the microwave background, we can construct elaborate deductive models and then test them observationally. When we do not have this confidence, it is better to restrict our science to simple inferences from direct observations.†

When I was asked to give this talk, I was suggested that I try to take a critical and provocative look at recent developments from the perspective of one who has worked more on the fringes of cosmology than at its core. This means that I will inevitably raise some concerns that are apparently not shared by other theorists. In order to assist myself in this unpopular task, I will appeal to a few apposite quotations from one of the more pungent academic commentators, George Santayana.

1.2 The Standard Model

The world is a perpetual caricature of itself; at every moment it is the mockery and the contradiction of what it is pretending to be. G. Santayana

There is some similarity between the development of the standard model of particle physics and what is becoming the standard model of cosmology, which, for the purposes of this talk, I will define as ΛCDM and a scale-invariant scalar fluctuation spectrum. Both represent a stage of organization and measurement allowing predictions to be made and tested, although cosmology is trailing particle physics by roughly 20 years. Both standard models are refutable and are characterized by a small set of parameters that can be measured. In particle physics, none of the experimental challenges have stood up although there is evidence of and hopes for new physics beyond the standard model, specifically neutrino mass and supersymmetry, respectively. In cosmology, there is impressive progress toward a concordance, and none of the serious observational challenges so far have been upheld. In addition, in both fields there are fundamental questions, for example the origin of mass in particle physics and the value of the scale of the density fluctuation spectrum in cosmology, that cannot be understood at present and may turn out to be metaphysical. Indeed most of the exciting theoretical (as opposed to phenomenological) research today in cosmology is addressing issues beyond the standard model, and the goals of the next generation of accelerators and space missions are to discover and characterize supersymmetry and see direct evidence for tensor modes from inflation.

As brought out in the excellent historical talks by Longair, Christiansen, and Trimble, it

† It is an ironic commentary on recent progress that we are more confident in the application of physics to the Universe at large before reionization than after.

is surprising how deep are the standard model's intellectual roots. The cosmological constant derives from the first relativistic cosmology, the Einstein static universe, and figured prominently in world models introduced over the following decade, most notably the de Sitter model. McCrea in 1951 understood the point that a substance with $P = -\rho$ would have similar dynamical consequences. This offered a more physically palatable cosmology than Hoyle's C-field. Bondi, in the first edition of his cosmology textbook, published in 1952, presents the analytical solution for the scale factor of the ΛCDM world model:

$$a(t) = \left(\frac{\Omega_0}{1-\Omega_0}\right)^{1/3} \sinh^{2/3}\left[\frac{3(1-\Omega_0)H_0 t}{2}\right], \tag{1.1}$$

where H_0 is the contemporary Hubble constant and Ω_0 is the current matter density parameter. He also understood the distinction between hot and cold matter.

However, all of this was comfortably theoretical. It took the discovery of the microwave background in 1965 to give a firm observational basis to physical cosmology. Within a few years, the theory of helium formation, recombination, the transfer of microwave background radiation, and gravitational instability were all given a secure foundation by Zel'dovich, Sunyaev, Peebles, and many others. This foundation that has been built upon over the last 35 years by successive generations of theorists is a remarkable, collective achievement, to provide the robust framework that can be used to interpret the equally impressive observations of the microwave background fluctuation and galaxy power spectra, most recently from *WMAP* and SDSS. The result, which was celebrated at this meeting and in papers submitted soon afterwards, is that, out of the suite of possible cosmological models, the one that describes the contemporary Universe is as generally anticipated (flat, dark matter-dominated, older than its constituents, endowed with scale-free perturbations). However, it is accelerating, which was not expected. As discussed here by Filippenko, Perlmutter, and Phillips, prior credit for this discovery should go to the Type Ia supernova teams who demonstrated that these explosions were far more standardized than it seemed reasonable to hope and that, at the 10 percent level, their luminosity distances increased slower with redshift than in an Einstein-de Sitter model. (I must remark that X-ray astronomers also deserve credit for anticipating this discovery with their steadfast claim that, granted the theory of nucleosynthesis, the ratio of the dark to the luminous mass in rich clusters of galaxies is much smaller than needed to close the Universe. The intercept with the flat Universe solution agrees with the supernova intercept to within the errors.)

I will come to the implications of this discovery later, but, for the moment I would like to consider a more limited approach to discussing the observations. This is based upon admitting that we do not understand the dynamics at all and so should just work in terms of kinematics. As we are claiming that the Universe was decelerating in the past and is now accelerating, it seems natural to introduce the third derivative of $a(t)$. Engineers often call this quantity the jerk, and following cosmological tradition it is natural to introduce a jerk parameter

$$j(t) = \frac{a''' a^2}{a'^3} \quad . \tag{1.2}$$

One reason why this is a nice thing to do, at least for observational cosmologists, is that the standard model—ΛCDM–has the curious property that $j = 1$.

We can use this as a basis for a purely kinematic description of the expansion of the Universe. The standard, dynamical first-order differential equation for the scale factor is

$$a' = H_0 \left[\Sigma_i \Omega_i(a)\right]^{1/2}, \tag{1.3}$$

where the functions $\Omega_i(a)$ measure the various contributions to the density. In the standard model, these are $\Omega_0 a^{-1}$ and $(1-\Omega_0)a^2$. If we adopt the convention that $a(t_0) = 1$ as a boundary condition, then we measure the Hubble constant to give a scale to size and age of the Universe and fit a single parameter Ω_0 to give the shape of the expansion curve. (Note that the deceleration parameter is $q_0 = 1.5\Omega_0 - 1$.) If we go beyond the standard model and allow the dark energy to have a nontrivial equation state, then we must fit additional parameters.

In the kinematic approach, we replace this equation with the third-order differential equation

$$a^2 a''' = j a'^3, \tag{1.4}$$

with a constant parameter j. Now, we need three boundary conditions and choose $a(0) = 0$, $a(t_0) = 1$ and $a'(t_0) = H_0$. It is straightforward to express t_0 in terms of q_0 and to regard q_0 as a parameter. We can then compute a three-parameter family of world models $a(t; H_0, q_0, j)$, similar to the three-parameter family $a(t; H_0, \Omega_0, w)$ used in the currently standard formalism.

Now, most astronomical measurements produce values for the comoving distance

$$d \equiv a d_L \equiv d_A/a = \int_t^{t_0} dt'/a \tag{1.5}$$

as a function of $a \equiv 1/(1+z)$. We can then use the measured values of $d(a)$, for example from a sample of Type Ia supernovae (Filippenko, Perlmutter), to fit for j. Just as it is possible to introduce a second parameter w' to characterize the dark energy, it is possible to add additional parameters such as j' to the kinematic description and to test the hypothesis that $j = 1, j' = 0$. The family of world models described by this parametrization seems no less well-motivated than that described by the $w - w'$ formalism.

1.3 Standard Parameters

Science is nothing but developed perception, interpreted intent, common sense, rounded out and minutely articulated. G. Santayana

Arguably the most important recent deduction that has been made from microwave background observations is that the geometry of the Universe is spatially flat to an accuracy of about 2 percent. Not only will this make life easier for future generations of astronomy graduate students but it provides an affirmation of a principle that had been expected, although not universally, as a consequence of the theory of inflation, which offers a mechanism for achieving this flatness. The way it comes about is more complicated than generally appreciated, as it depends upon knowing the size of the measuring rod at the last scattering surface (which comes from the theory of acoustic oscillations) as well as the angular diameter distance of this surface. Its measurement is likely to get much more accurate, especially after the *Planck* mission, and I suppose that there is a strong prediction that the spatial curvature of the Universe, insofar as it can be isolated from the physics of recombination, should remain unmeasurable.

As I am not an angel, let me tread fearlessly into a Carnegie speciality, the measurement of the Hubble constant. All the measurements reported here are consistent with the

WMAP value of $h = 0.72$ (with 4 to 7 percent errors, depending upon whether or not non-*WMAP* information is included) (Page). However, the full quoted errors from some of the non-*WMAP* methods standing alone is still disappointingly large (Filippenko, Perlmutter, Phillips, Hamuy, Jensen, Reese). The problem is that all of these approaches are subject to systematic error that is as difficult to bound as to reduce. It is possible that theorists will be able to contribute more to this measurement through making better Cepheid models and providing comprehensive models of rich clusters of galaxies that address all of the puzzles that they now present when viewed throughout the electromagnetic spectrum.

Gravitational lens determinations are subject to similar problems (Kochanek). All approaches rely on measuring time delays, and the accuracy of these measurements in the best-studied of these systems is no longer limiting. The major problem is with the modeling. This is one of the few techniques that still has a chance of being competitive with the microwave background determination. Here, I think there are two approaches. The first relies on *a priori* models of the lensing galaxies. These are generally pretty simple. There has been much progress in understanding the sensitivity of the result to the details, and understanding the radial profile of the mass distribution is crucial. (Measuring the galaxy velocity dispersion and its radial variation can help a lot.) The image flux ratios are proving to be interesting as they can be diagnostic of the amount of substructure in the lens potential. These small-scale clumps are generally expected on the basis of numerical simulations, but they are under-represented in visible light. The best way to quantify the modeling uncertainty is to study enough lensing systems in detail, and the large CLASS and SDSS surveys (Bernardi, Colless) are proving invaluable.

The second approach to the modeling of lenses is nonparametric. Basically, when one resolve extended sources, whose unresolved nuclei vary, one can plot out a pattern of multiply imaged isophotes around the Einstein ring. As surface brightness is conserved, these isophotes map onto each other, and so the only freedom is to slide along the isophotes. However, it turns out that if, as is true for a single lens, the deflection field is the gradient of a scalar potential, then this freedom is restricted and subject to certain conditions, and it is possible to predict the ratio of time delays up to a constant inversely proportional to h. Perfect data should furnish a perfect answer. Of course, in the real world the method is compromised by finite pixelation, telescope aberration, reddening, difficulty in distinguishing the source from the image, and so on, and it is difficult to determine how useful this approach will be in the foreseeable future, using a combination of VLBI and ACS imaging on sources like B1608+656. However, it does come with enough internal consistency checks that it will be possible to assess this using relatively few examples.

The major source of systematic error will ultimately be the so-called mass-sheet degeneracy. This relies on the observation that a large sheet of matter in front of the source acts like a converging lens, which will appear to shrink the Universe. This mass can be contributed by matter along the line of sight, for example in an intervening, rich cluster of galaxies, just as much as by matter in the lens plane. This is very similar to the old uncertainty that existed when we did not know which world model to use, which afflicts all determinations when the uncertainty associated with the inhomogeneity of the Universe becomes important. It can be removed either statistically or by tracing the mass distribution out to large radius,

The *WMAP*-inferred age of the Universe, 13.7 Gyr, is closely linked to the Hubble constant, and it is a minor coincidence that it is so close to being its reciprocal. This is another pleasing consistency check, as it is now concordant with the best direct chronometers, a far cry from the days when the Einstein-de Sitter model held sway when the age was 30 percent smaller. As this is now more reliable than the direct measurements, it will be very useful in understanding the history of our Galaxy and its oldest stars.

The mean baryonic density was first supplied by the theory of nucleosynthesis when combined with the measurement of intergalactic deuterium (Steigman), as well as consistent determinations of the mean density of other light elements. This approach has received remarkable vindication from the microwave background fluctuation analysis, which is now the method of choice and gives $\Omega_{\rm b} = 0.044$. It is also consistent with the direct measurement of the column density of intergalactic gas in quasar absorption spectra when the Universe was ~ 4 Gyr old, provided one trusts the estimate of the intergalactic ultraviolet radiation density at this epoch, which is necessary to fix the ratio of ionized to neutral hydrogen. There was no appreciable amount of nonbaryonic dark matter. We can now turn the argument on its head and accept the microwave background determination and very simple classical physics to infer that nuclear physics cross sections were the same when the Universe was roughly a minute old as they are today.

The *WMAP* determination of the total matter density is close to $\Omega_0 = 0.27$, with and without priors supplied by other approaches. This is quite consistent with the X-ray measurements in the most relaxed clusters of galaxies (Silk). However, it is not consistent with local determinations of the intergalactic hydrogen column density using quasar absorption lines. We can see no more than a quarter of now. This is not a mystery because numerical simulations show that the gas outside clusters collides and will shock to temperatures $\sim 10^6$ K, such that the sound speed is comparable with the virial velocities (~ 200 km s^{-1}) present in the nonlinearly clustered dark matter distribution. Gas at this temperature is rather hard to detect, being too high for ultraviolet detectors like *FUSE* and too soft for hard X-ray detectors. There are clear indications that it is present, especially from *XMM-Newton* observations, but its detailed study will, sadly, have to wait for the next generation of X-ray detectors like *Constellation-X* and *XEUS*.

There are still some puzzles associated with the behavior of this gas and, although it is not the subject of this conference, I would like to mention, in passing, one feature of the astrophysics that has not been widely appreciated. We know from X-ray and radio observations of supernova remnants that, when astrophysical plasma passes through a high-Mach number, collisionless shock front, cosmic rays are also accelerated. We may not understand the details of how this occurs, but it does seem that a partial pressure of perhaps 0.3 is transmitted in the form of mildly relativistic GeV particles. The slope of the momentum space cosmic ray distribution function power law, $-q$, is typically ~ -4.2, so that the effective specific heat ratio of this phase is $\gamma_{CR} = q/3 \approx 1.4$, significantly less than that of the thermal plasma. Now, intergalactic shock fronts should surely behave the same, and I see no escape from the conclusion that the intergalactic gas has an appreciable cosmic ray component. As the gas expands in the post-shock flow, the dynamical importance of these cosmic rays will diminish. However, when the gas recompresses as it falls into dark matter potential wells that are eventually associated with galaxies, groups, and clusters, it will compress to a density that is ultimately greater than that with which it started. Furthermore, the gas will start to cool, thereby increasing the importance of the cosmic rays that have no where to escape. This provides a different approach to understanding the characteristic of intergalactic gas that has been called an entropy floor.

1.4 Cold Dark Matter

For an idea to be fashionable is ominous, since it must afterward be always old-fashioned. G. Santayana

The idea of dark matter can be recognized in a Newtonian context in papers going back to the nineteenth century. However, it is Zwicky who is generally credited with making the quantitative case for its presence through measuring the mass-to-light ratio of the galaxies in the Coma cluster (Silk). Nowadays, we distinguish baryonic from nonbaryonic dark matter.

The astronomical options for dark matter are now mostly ruled out, leaving particle physics as the most likely place to turn (Silk). The most popular candidate is now surely an irreducible supersymmetric particle—perhaps, though not necessarily a neutralino. The mass of this particle is generally argued on physical and astrophysical grounds to lie in the range 30–1000 times the mass of the proton. In the solar neighborhood, these particles are expected to be roughly 3–10 cm apart on average. This is timely as the supersymmetric sector will be accessible to experiment using the Large Hadron Collider. Supersymmetric dark matter is also already accessible to terrestrial detectors that hope to measure the recoils associated with scattering. Furthermore, direct detection of dark matter particles is possible by measuring the recoil of atomic nuclei (Sadoulet). A somewhat longer shot is that annihilations of these particles will be seen by the *GLAST* gamma ray observatory due to be launched in 2006. This will be sensitive to gamma rays with GeV energies and will be looking for narrow lines from clusters of galaxies.

Although rather overshadowed by the evidence for dark energy in recent years, it is surely the quest to understand dark matter and the exciting prospect of being able to study the detailed properties of a new class of elementary particles that one must be more optimistic about. Convincing evidence for the direct detection of cosmological dark matter particles could come any day, while the accelerator position should become clear within five years. For astronomers, the challenge is to try to find new ways to probe its particulate nature that go beyond studying its overall distribution.

An alternative possibility for the dark matter is the axion, which was invoked theoretically in an extension to the standard model of particle physics to solve the strong CP problem. Axions are sought experimentally using resonant microwave cavities. Here the parameter space that must be searched is even larger, but upper limits are already being reported.

1.5 Large-scale Structure

Chaos is a name for any order that produces confusion in our minds. G. Santayana

It is remarkable that we can connect the ancient and the modern Universe in a rational and consistent fashion using an essentially Newtonian theory for the growth of perturbations as the Universe ages from four hundred thousand years to fourteen billion years. This is a remarkable consistency check and an affirmation of our tacit assumption that the evolution of the Universe between recombination and the present day followed the laws of relativistic dynamics. Small-scale perturbations actually grow according to the laws of Newtonian dynamics. However, perturbations entering the horizon at early times need relativity for their prescription. This overlap between cosmic microwave background and galaxy survey power spectra is possible over comoving length scales of $\sim 30-600$ Mpc. The total fluctuation spectrum is impressively consistent with the Harrison-Zel'dovich scale-invariant form at the

5 percent level over 4 decades in wavenumber if one accepts the assumptions that go into including Lyman α measurements of the fluctuation spectra.

One of the worries that I have about this procedure, though, is that the bias—the relationship between the clustering of the luminous galaxies and the underlying dark matter—could still be a free parameter and used to adjust the two spectra. The very fact that elliptical galaxies and spiral galaxies cluster differently and have different luminosity functions assures us that the considerations that led to the introduction of bias in the first place cannot easily be dismissed.

For this reason, I see the future of "precision" large-scale structure investigations as lying with weak lensing studies, which study the mass directly rather than the light. Remarkable advances have been made using this technique, and it looks ready for application to enormous data sets garnered from special purpose telescopes in space and on the ground, the former for high spatial frequencies, the later for low frequencies.

1.6 Galaxies

There is no cure for birth and death save to enjoy the interval. G. Santayana

For the purposes of this meeting, galaxies are mostly elementary particles (and hosts of supernovae). However, they do have their own peculiarities. Ellipticals are more gregarious than spirals, and these differences will have to be included in more detailed comparisons of the microwave background fluctuation spectra with the galaxy correlation function.

There are two puzzles that have come from the otherwise extremely well-interpreted microwave background observations. They may be linked. The first of these is that the optical depth to Thomson scattering from the cross correlation of the polarization and the density spectrum is larger than expected, implying that reionization occurred when the Universe was a hundred million years young (Page). The other is that the high-angular frequency perturbations are stronger than predicted. This may be due to integrated Sunyaev-Zel'dovich fluctuations along the line of sight, but, if so, the fluctuation level must be significantly larger than measured by other means ($\sigma_8 \approx 1$). A common explanation of both of these results is that we are seeing the consequences of the formation of the first, clustered superstars (Readhead). Future observations will surely clarify this matter.

Looking at the other end of the galaxy life cycle, it is natural to ask if there is a significant population of dead galaxies that only formed high-mass stars and burned out. The evidence from gravitational lensing is negative, but this might be worth considering more carefully.

1.7 Inflation

Fanaticism consists in redoubling your efforts when you have forgotten your aim. G. Santayana

The microwave background observations are widely taken as a corroboration of the principle of the theory of inflation. This is because the Universe is flat and, as noted above, because a scale-invariant spectrum is measured. However, these observations do not furnish a proof of inflation. I would argue that the strongest reasons, and they are very strong, for believing that the very early Universe passed through a phase of exponential expansion remain theoretical. Even the prediction of flatness was not universal. When confronted with the evidence that $\Omega_0 \approx 0.3$ and presented with the choice of a negatively curved, dark matter-dominated Universe and an exquisitely tuned dark energy-dominated Universe, many theorists took the first option. I do not personally see any compelling reason to regard tiny curvature as a much less likely outcome than a tiny dark energy density when we do not know when inflation occurred, have no accepted theory to describe the inflationary potential and the strength of the perturbations that it leaves us *ab initio*, and no good understanding of the Universe's exit strategy from exponential expansion.

Furthermore, the idea of a scale-invariant spectrum preceded the theory of inflation by several years and is far more general. By itself, it does not seem to be the smoking gun for the theory, even if the theory provides a good mechanism for producing it. This is also true for the tensor spectrum, which may one day be measured using the B-modes of the microwave background polarization (Zaldarriaga). A radio map of an extragalactic radio source produces B-modes in abundance. There is nothing very special about them, and I suspect that alternative cosmological sources could be found by devil's advocates. (The prospects for direct detection by gravitational radiation observatories seem particularly bleak [Flanagan, Lazzarini].) Not until we have a quantitative understanding of the physics, and this will require some understanding the form of the inflationary potential from first principles, as well as being able to compute the fate of quantum fluctuations, can we say that inflation has been observed. Perhaps the best hope for providing a good "proof" of inflation lies in general quantitative relationships between the fluctuation spectra that exist under restrictive assumptions.

1.8 Second Inflation

Those who cannot remember the past are condemned to repeat it. G. Santayana

Although theoretical physicists dislike this connection, it does seem as though we are entering upon a second epoch of inflation. (Perhaps this has happened more than twice, although there are clear constraints on our past lightcone against this having taken place after nucleosynthesis.) As we look to the future, we can survey a bleak and depressing eschatology of dilution and decay. Alternatively, we can take the optimistic viewpoint that we have got out of this mess once in the past and if we try hard enough, we can do it again!

More seriously, and for reasons that I still do not understand, the theoretical description of dark energy has followed a different path from the description of inflation. Now, the simplest and most natural possibility is that we are dealing with a cosmological constant, which we would interpret nowadays as a form of vacuum energy. The stress energy tensor for dark energy is simply proportional to the metric tensor and is the same in all frames and at all times. Dark energy is a simple fact of nature, like the existence of electrons. (This is not quite as mysterious as one might think when one considers a uniform magnetic field. It, too, has a net stress along the field that is the negative of the energy density. Of course, it is not isotropic.) This is the basis of the ΛCDM model, and there is no good evidence that it is wrong.* Now, as with general relativity itself, which is the simplest and least fussy theory of gravity, we need to provide a framework for checking if the simple view is correct and, if it is not, for providing a framework for relating it to the expansion of the Universe.

Most general descriptions of dark energy, for which most observing proposals are written, treat it as an isotropic fluid endowed with a pressure and a density endowed with an equation of state

* The recent detection of the integrated Sachs-Wolfe effect provides a very neat affirmation that it must be approximately correct.

$$P = w\rho \quad . \tag{1.6}$$

(This, surely, is the equation of state, *not* the parameter w, a usage favored by many authors.) What this equation is proposing is a modified Boyle's law

$$PV^{1+w} = \text{negative constant}, \tag{1.7}$$

where $w = -1$ corresponds to the cosmological constant, and we already know that $w < -0.8$. However, if we interpret w as a pure number, there is a problem. If we consider small perturbations $\propto \exp[i(\vec{k}\cdot\vec{x} - \omega t]$, then the dispersion relation is $\omega^2 = wk^2$ and there will be growth on all scales as long as $w < 0$.

Of course, this is not what was intended. This prescription is supposed to represent the dynamics of a scalar field ϕ with a potential $V(\phi)$. We use Noether's theorem to derive the stress energy tensor from the Lagrangian density, from which the density and pressure can be read off:

$$\rho = \phi'^2/2 + (\nabla\phi)^2/2 + V(\phi) \tag{1.8}$$

and

$$P = \phi'^2/2 - (\nabla\phi)^2/6 - V(\phi) \quad . \tag{1.9}$$

In addition, by setting the divergence of the stress energy tensor to zero, we can obtain the equation of motion for the field:

$$\phi'' + 3H\phi' - \nabla^2\phi + dV/d\phi = 0 \quad . \tag{1.10}$$

What is usually done is to ignore spatial gradients and the second derivative in the equation of motion. In this case, it is possible to relate w to V. This may be permissible on the scale of the horizon. However, the danger of instability remains, and, so far as I can see, the stability of the field and its impact on the growth of matter perturbations are not routinely considered. All of this must be computed directly from the field theory, and it does not seem to me that a fluid model of dark energy is of any more utility than a fluid model of the electromagnetic field has proven to be, despite the fact that it can be written down. Furthermore, the prescribed solutions with fixed w, say, may be quite different from those that could follow from a real field theory.

It turns out that, under reasonable assumptions, a knowledge of $a(t)$ can be used to infer $V(\phi)$ directly. It therefore seems much simpler to bypass the fluid treatment and either concentrate on kinematics, as I discussed above, or to introduce a simple set of parameterized potentials, as was done in nuclear physics. Observations can then be used to fit for these parameters, and theorists should then explore systematically the other consequences of these dynamical models.

There is no guarantee that second inflation is driven by a scalar field, and it is worth recalling that we have not detected any examples of fundamental scalar fields to compare with yet. To make this point, let us suppose that a blind astrophysicist was present during the radiation-dominated era and perceived the mean separation of electrons and protons increasing as the square root of time, and she tried to understand what was going on. She could find a solution to the scalar field equations that fit all the known facts—specifically it has $V(\phi)$ be a decaying exponential—and she would be completely wrong! Are we so sure that the only approach to explaining dark energy has to be cast in the language of a scalar field?

1.9 The String Connection

The cry was for vacant freedom and indeterminate progress. G. Santayana

The deductive approach to cosmology is surely from a fundamental theory of quantum gravity of which the overwhelmingly best-developed candidate is string theory (Schwartz). The challenge used to be understand why the cosmological constant vanishes. It has now become more interesting! Remarkably, there has been recent progress in understanding new classes of string vacua, some of which have quite unexpected classical behavior.

1.10 Anthropic Reasoning

Reason is not come to repeat the universe but to fulfil it. G. Santayana

There has recently been an upsurge in the use of anthropic arguments (Livio) and a serious reaction, particularly from theoretical particle physicists, who see it as dangerously close to religion and therefore a bad thing. I suspect that anthropic arguments do have a religious component but in a rather different sense from that attributed to them by the critics. Like many religions, anthropism provides a measure of solace in hard times and is set aside when things are going well. It is really hard to imagine that an anthropic explanation would be preferred to a rigorous physics calculation or that physicists would stop trying to make such calculations when told that the reason why things were as they were is because we are as we are. Nonetheless, anthropic arguments frequently illustrate beautifully the fundamentals of physics and may have made teleology respectable.

1.11 The Theoretical Challenge

Fashion is something barbarous for it produces innovation without reason and imitation without benefit. G. Santayana

At several times in the history of cosmology, there have been bold steps taken, for example by de Sitter, Friedman, Gamow, Guth, and Linde. Each of these advances was incomplete and had to be supplemented with corrections and amendments. However, they each had the characteristic that they were motivated by deep and fundamental physical questions that were real, and they faced up to these questions in ways that were constrained by what was then observed. Today, cosmology is a much more crowded and competitive field. It is also a more accomplished field, and much contemporary research is technically outstanding and reactive to detailed observational studies, sometimes retrofitting existing models, as in semi-analytical galaxy formation theory, occasionally making clean predictions and interpretations, as in the discussion of microwave background fluctuations.

However, the questions that cosmologists face are as deep as any in the history of the subject, and they are now quite crisply stated. At present it looks like the standard model is correct, and, if we set aside anthropic arguments and adopt physics as we find it locally, we are required to "explain" just two numbers, the entropy per baryon in units of k and the level of the fluctuation amplitude (in round numbers 10^{10} and 10^{-5}). The first of these requires a theory of baryogenesis, the second a theory that requires us to understand the connection to some form of dissipation at very early time. The standard ways of approaching these problems are to posit phase transitions at either the strong or the electroweak scale and to invoke a designer inflaton potential. I have no objections to either of these approaches and certainly have nothing special to offer in their place. However, I wonder if there are more possibilities (like, for example, the brane cosmologies) that promise quite different explanations, and which could impact physics on a scale of interest to laboratory physicists.

As we go beyond the standard model and anticipate even more accurate microwave background and large-scale structure observations, there is the prospect of measuring a slope parameter n different from 1 and detecting the tensor fluctuations. These would enable us to estimate the extent of the fluctuation spectrum and presumably the duration of exponential expansion under the inflationary paradigm. Features may appear in the primordial fluctuation spectrum, which would be incredibly prescriptive of new physics.

I am probably most optimistic about the prospects for studying supersymmetry using data acquired below, on, and above ground. Of course there are no guarantees and nothing is proven yet, but so many independent approaches point in this direction. If this happens, it will be a tremendous shot in the arm for particle physicists who, having developed one standard model, have to start all over again. It is nice to think of cosmologists repaying some of the debt that they owe particle physicists in this enterprise.

However, the biggest challenge of all is to understand why the Universe appears to be accelerating. For reasons to which I have alluded, I personally find the alternatives to pure vacuum energy unappealing. If it turns out that $w = -1$ (or $j = 1$), then we have a third single number to explain, the ratio of the dark energy density to the Planck density, which is 10^{-120} also in round numbers. This is surely an even greater challenge to fundamental thinking.

I started this review with a statement of my charge to be critical and provocative. I think I have largely failed. Perhaps it is the vicarious pleasure in seeing a theoretical (and observational) job well done (at least so far) and the clarity of the questions that have been uncovered in the process. Perhaps it is the realization that we have not reached the end of the road in terms of achievable observations and affordable facilities and that the next decade will surely provide several of the cosmological measurements that we will need. Perhaps I have been swept along by the hyperbolic public statements to which cosmologists are prey. However, the time does seem ripe for a new radical idea that might take our understanding of the Universe to a higher level and remove the need, which we currently have, to model as well as to measure.

Acknowledgements. I am indebted to Wendy Freedman and Luis Ho for their encouragement and patience and many of my theoretical colleagues for their instruction. I acknowledge support under NSF grant AST-0206286.