Analysis and Synthesis III: Cosmic Groundwork

Frank Wilczek

Most of cosmology can be captured in a few parameters, but those parameters are themselves mysterious, as I discussed in my previous column (see PHYSICS TODAY, July 2003, page 10). Now I'll survey some prospects for understanding them better.

Digging deeper

A big reason for excitement and optimism among physicists is the emerging possibility of forging links between fundamental physics and cosmology through models of inflation. Several assumptions in our cosmological models, specifically uniformity, spatial flatness, and the Harrison–Zeldovich spectrum, were originally suggested on grounds of simplicity, expediency, or aesthetics. They can be supplanted with a single dynamical hypothesis: that very early in its history, the universe underwent a period of superluminal expansion, or inflation. Such a period could have occurred while a matter field that was excited coherently out of its ground state permeated the universe.

Possibilities of that kind are easy to imagine in models of fundamental physics. For example, scalar fields are used to implement symmetry breaking even in the standard model and, theoretically, such fields can easily find themselves unable to shed energy quickly enough to stay close to their ground state as the universe expands. Inflation will occur if the approach to the ground state is slow enough. Fluctuations will be generated because the relaxation process is not quite synchronized across the universe.

Inflation is a wonderfully attractive, logically compelling idea, but very basic challenges remain. Can we be specific about the cause of inflation, and ground it in explicit, well-founded models of fundamental physics? To be concrete, can we calculate the correct amplitude of fluctuations convincingly? Existing implementations actually have a problem on this score; getting the amplitude sufficiently small takes some nice adjustment.

Frank Wilczek *is the Herman Feshbach Professor of Physics at the Massachusetts Institute of Technology in Cambridge, Massachusetts.*

More promising, perhaps, than the difficult business of extracting hard quantitative predictions from the broadly flexible idea of inflation is to follow up on the essentially new and surprising possibilities it suggests. The violent restructuring of spacetime attending inflation should generate detectable gravitational waves, and the nontrivial dynamics of relaxation should generate some detectable deviation from a strictly scale-invariant spectrum of fluctuations. Future precision measurements of polarization in the microwave background radiation and of the large-scale distribution of matter will be sensitive to these effects. Stay tuned!

There are many ideas for how an asymmetry between matter and antimatter might be generated in the early universe. Then after much mutual annihilation of particles and antiparticles, the asymmetry could be left as the present baryon density. Several of those ideas seem capable of accommodating the observed value. Unfortunately the answer generally depends on details of particle physics at energies that are unlikely to be accessible experimentally any time soon. So for a decision among them we may be reduced to waiting for a functioning Theory of (Nearly) Everything.

Dark matter

I'm much more optimistic about the dark matter problem. Here we have the unusual situation that two good ideas exist—which, according to William of Occam (the razor guy), is one too many.

The symmetry of the standard model can be enhanced, and some of its aesthetic shortcomings can be overcome, if we extend it to a larger theory. Two proposed extensions, logically independent of one another, are particularly specific and compelling. One incorporates a symmetry suggested by Roberto Peccei and Helen Quinn in 1977. Peccei–Quinn symmetry rounds out the logical structure of quantum chromodynamics by removing QCD's potential to support strong violation of time-reversal symmetry, which is not observed. This extension predicts the existence of a remarkable new kind of very light, feebly interacting particle: the axion.

The other proposal incorporates supersymmetry, an extension of special relativity, to include quantum spacetime transformations. Supersymmetry serves several important qualitative and quantitative purposes in modern thinking about unification; it relieves difficulties with understanding why W bosons are as light as they are and why the couplings of the standard model take the values they do. In many implementations of supersymmetry, the lightest supersymmetric particle, or LSP, interacts rather feebly with ordinary matter (though much more strongly than do axions) and is stable on cosmological time scales.

The properties of the particles, axion or LSP, are just right for dark matter. Moreover, you can calculate how abundantly each would be produced in the Big Bang. For both particles, the predicted abundance is also quite promising. Vigorous, heroic experimental searches are under way to observe dark matter in either of those forms. We will also get crucial information about supersymmetry once the Large Hadron Collider starts running in 2007. I will be disappointed—and surprised—if, a decade from now, we don't have a much more personalized portrait of the dark matter.

Dark energy

Now for a few words about the remaining parameter, the density of dark energy. Why is it so small? Why is it so big?

The standard model provides a great lesson: What we have been evolved to perceive as empty space is in fact a richly structured medium. It contains symmetry-breaking condensates associated with both electroweak superconductivity and spontaneous chiral symmetry breaking in QCD, an effervescence of virtual particles, and probably much more.

Because gravity is sensitive to all forms of energy, it really ought to see this stuff, even if our eyes don't. A straightforward estimation suggests that empty space should weigh several orders of magnitude of orders of magnitude (no misprint here!) more than it does. It "should" be much denser than a neutron star, for example. The expected energy of empty space acts like dark energy, with negative pressure, but far more is expected than is found.

To me, the discrepancy concerning the density of empty space is the most mysterious fact in all of physical science, the fact with the greatest potential to rock the foundations. We're obviously missing some major insight here. Given that, it's hard to know what to make of the ridiculously small amount of dark energy that presently dominates the universe.

Possible worlds

Discovery of definitive mathematical equations that govern the behavior of matter, and even mastery of techniques for solving them, if we ever achieved it, would by no means complete the program of understanding nature through analysis and synthesis. We would still need to address the problem of selection. How, among all possible solutions, do those that actually describe reality get selected out?

Niels Bohr defined a profound truth as a truth whose opposite is also a (profound) truth. In that spirit, I'd like to define a deep question as a question that doesn't make clear sense until after you've answered it. The selection problem is a deep question. In posing it, we take it for granted that we can make a distinction between what is "possible" and what is "real." On the face of it, that sounds pretty unscientific; the goal of science is to understand the real world, and only the real world is possible! But in practice a clear and useful distinction does emerge.

A famous episode from the early history of science will serve to illuminate the issues. In Johannes Kepler's first attempt to understand the Solar System, he postulated that the relative sizes of planetary orbits are determined as a system of concentric spheres, centered at the Sun, inscribed and circumscribed around a sequence of the five regular (Platonic) solids. To Kepler, the "fact" that there were six planets provided an impressive numerical confirmation of his idea. In light of later developments Kepler's model rates a bemused smile, but purely as a logical matter, it might have been right. And if such a model had provided the ultimate account of the Solar System, no useful distinction would have developed between possible solutions of its governing equations and the solution actually realized. Indeed, there wouldn't be equations to solve, as such—just the solution, perfect in its own terms, and incapable of further analysis.

Kepler's geometric model was soon made obsolete by the development of classical celestial mechanics, beginning with Kepler's own discoveries and culminating in Isaac Newton's world-system. That framework gives a clean separation between the governing laws and their specific realizations. The equations of classical mechanics can be solved for any number of planets, with any initial assignment of positions and velocities. So in the physics of 1700—as in the physics we use today—it was easy to imagine perfectly consistent planetary systems with different sizes and shapes than the Solar System. Galileo had already observed a very differently textured planetary system around Jupiter, and today astronomers are discovering new types surrounding distant stars.

Newton thought that God determined the initial conditions, by an act of will and creation. Nowadays most physicists consider the question of why the Solar System is precisely what it is as a bad question, or at least one that cannot be addressed from fundamentals, since it seems pretty clear that the answer depends largely on chance and accidents of history.

The Big Bang's peculiar foundation

The ultimate question of selection, of course, is how to select among possible candidate solutions to describe the universe as a whole.

Conventional Big Bang cosmology begins by assuming that, early on, matter throughout the universe was in thermal equilibrium at some high temperature, while at the same time space was (almost) perfectly uniform and had negligible spatial curvature. Those assumptions are consistent with everything we know and they lead to several successful predictions. They surely contain a large element of truth.

Yet from a fundamental perspective, these assumptions are most peculiar. According to general relativity, which of course underpins the whole discussion, spacetime curvature is a dynamical entity whose shape yields gravity. It would appear natural, therefore, to assume that spacetime too achieves some sort of equilibrium, through gravitational interaction. But gravity's acting on matter produces universal attraction, and if allowed to run to completion, that attraction will agglomerate the matter into lumps. The long-term equilibrium result is the very opposite of the smooth starting conditions commonly—and quite successfully—assumed.

Putting it another way, if the present universe were eventually to start contracting and evolve toward a Big Crunch, matter as it was squeezed together would tend to heap into bigger and bigger black holes, and the last moments would look nothing like the time-reverse of our best reconstruction of the beginning. That contrasting mix of maximum disorder for matter and perfect quiescence for gravity is, on the face of it, extremely difficult to reconcile with the idea that ultimately gravity is unified with the other interactions and should be treated equally with them in describing extreme conditions in the earliest universe. Where there's no distinction, there can be no difference!

In particle physics, we like to unify gravity with the other interactions, but in cosmology, we apparently have to assume they behaved very differently. Simple models of inflation begin to address this issue. During the inflationary phase, wrinkles in spacetime get smoothed out, while energy is frozen into a scalar "inflaton" field. The inflaton field eventually melts, and its energy pours into more-or-less conventional forms of matter, which interact and thermalize. As long as the melting does not produce too high a temperature, few gravitons are produced, which is to say that spacetime, having been stretched smooth, remains so. Some elements of this picture are quite speculative, of course. But it gives us a tangible, accessible approach to the problem of certifying the soundness of cosmology's foundation.

Even if this program of analysis and synthesis succeeds, questions will remain—in particular, we still lack a fundamental explanation of why the inflationary phase occurred. The general problem of how we select among solutions appears as one of the ultimate limits to the program. Ultimate limitations will be my concluding subject in this series. ■

Letters and opinions are encouraged and should be sent to Letters, PHYSICS TODAY, American Center for Physics, One Physics Ellipse, College Park, MD 20740-3842 or by e-mail to ptletter@aip.org (using your surname as "Subject"). Please include your affiliation, mailing address, and daytime phone number. We reserve the right to edit submissions.