The Future of Fundamental Physics

Nima Arkani-Hamed

Abstract: Fundamental physics began the twentieth century with the twin revolutions of relativity and quantum mechanics, and much of the second half of the century was devoted to the construction of a theoretical structure unifying these radical ideas. But this foundation has also led us to a number of paradoxes in our understanding of nature. Attempts to make sense of quantum mechanics and gravity at the smallest distance scales lead inexorably to the conclusion that space-time is an approximate notion that must emerge from more primitive building blocks. Furthermore, violent short-distance quantum fluctuations in the vacuum seem to make the existence of a macroscopic world wildly implausible, and yet we live comfortably in a huge universe. What, if anything, tames these fluctuations? Why is there a macroscopic universe? These are two of the central theoretical challenges of fundamental physics in the twenty-first century. In this essay, I describe the circle of ideas surrounding these questions, as well as some of the theoretical and experimental fronts on which they are being attacked.

NIMA ARKANI-HAMED, a Fellow of the American Academy since 2009, is a Professor in the School of Natural Sciences at the Institute for Advanced Study. His interests range from quantum field theory and string theory to cosmology and collider physics. He has published his work in the *Journal of High Energy Physics*, the *Journal of Cosmology and Astroparticle Physics*, and *Nuclear Physics*, among other places.

Ever since Newton realized that the same force of gravity pulling down on an apple is also responsible for keeping the moon orbiting the Earth, fundamental physics has been driven by the program of *unification*: the realization that seemingly disparate phenomena are in fact different aspects of the same underlying cause. By the mid-1800s, electricity and magnetism were seen as different aspects of electromagnetism, and a seemingly unrelated phenomenon – light – was understood to be the undulation of electric and magnetic fields.

Relativity and quantum mechanics pushed the trend toward unification into territory far removed from ordinary human experience. Einstein taught us that space and time are different aspects of a single entity: space-time. Energy and momentum are united analogously, leading to the famous equivalence between mass and energy, $E = mc^2$, as an immediate consequence. Einstein further realized that space-time is not a static stage on which physics unfolds, but a dynamic entity that can curve and bend. Gravity is understood as a manifestation of space-time curvature. This new picture of space-time made it possible to conceive of ideas that were impossible to articulate in the Newtonian picture of the world. Consider the most important fact about cosmology: we live in an expanding universe. The distance between two galaxies grows with time. But the galaxies are not rushing apart from each other into some preexisting space, as though blown out of an explosion from some common center. Rather, more and more space is being generated between the galaxies all the time, so from the vantage point of any one galaxy, the others appear to be rushing away. This picture, impossible to imagine in Newton's universe, is an inevitable consequence of Einstein's theory.

Quantum mechanics represented a more radical departure from classical physics, involving a completely new conceptual framework, both physically and mathematically. We learned that nature is not deterministic, and only probabilities can be predicted. One consequence is the famous uncertainty principle, by which we cannot simultaneously know the position and velocity of a particle to perfect accuracy. Quantum mechanics also allowed previously irreconcilable phenomena to be understood in a unified way: particles and waves came to be seen as limiting aspects of the underlying description where there are no waves at all, only quantum-mechanical particles.

The laws of relativity and quantum mechanics are the pillars of our current understanding of nature. However, describing physics in a way that is compatible with both of these principles turns out to be extremely challenging; indeed, it is possible only with an extremely constrained theoretical structure, known as *quantum field theory*. A quantum field theory is characterized by a menu of particles that interact with each other in various ways. The nature of the interactions is almost completely dictated by the rules of quantum mechanics, together with the requirement that the interactions take place at points in space-time, in compliance with the laws of special relativity. The latter requirement is known as the principle of *locality*.

One of the startling general predictions of quantum field theory is the existence of anti-particles such as the positron, which has the same properties as the electron but the opposite electric charge. This prediction has another striking consequence: namely, that even the vacuum has structure and dynamics.

Suppose we attempt to check that some small region of space-time is empty. Because of the uncertainty principle, we need higher energies to probe short distances. Eventually there is enough energy to make an electron and a positron, without violating either the conservation of energy or the conservation of charge. Instead of seeing nothing, probing the vacuum at small distances yields particle/anti-particle pairs. It is useful to think of the vacuum as filled with quantum fluctuations, with "virtual" particles and anti-particles popping in and out of existence on faster and faster timescales at shorter and shorter distances.

These quantum fluctuations give rise to measurable physical effects. For instance, the cloud of virtual electrons and positrons surrounding an electron is slightly perturbed by the electron's electric field. Any physical measurement of the electron's charge, then, will vary just slightly with distance, growing slowly closer in to the electron as more of the virtual cloud is pierced. These virtual effects can be calculated very precisely; in some circumstances, theoretical predictions and experimental observations can be compared to an astonishing level of precision. The virtual corrections to the magnetic properties of the electron, for example, have been theoretically computed to twelve decimal places, and they agree with experiment to that level of precision.

The second-half of the twentieth century saw a flurry of activity, on both experimental and theoretical fronts. These developments culminated in the 1970s with the construction of the Standard Model of particle physics, a specific quantum field theory that describes all known elementary particles and their interactions down to the smallest distances we have probed so far. There are four basic interactions: gravity and electromagnetism, which were familiar even to the ancients, as well as the weak and strong interactions that reveal themselves only on nuclear scales. Atomic nuclei consist of neutrons and protons. An isolated neutron is unstable, living for about fifteen minutes before disintegrating into a proton, electron, and an anti-neutrino. (This process is also responsible for radioactivity.) Fifteen minutes is enormously long compared to the typical timescales of atoms and nuclei, so the interaction responsible for triggering this decay must be very feeble – hence, *weak* interaction. The earliest incarnation of the strong interaction was noticed in the attraction keeping protons inside nuclei, counterbalancing their huge electrical repulsion.

Some familiar particles, such as electrons and photons, remain as elementary point-like entities in the Standard Model. Others, like the proton, are understood to be bound states, around 10^{-14} cm in diameter made of *quarks*, which are permanently trapped inside the proton through their interaction with *gluons*.

Strong, weak, and electromagnetic interactions seem completely different from each other at long distances, but we now know that these differences are a long-distance illusion. At short scales, these interactions are described in essentially the same quantum-field-theoretic language. Electromagnetism is associated with interactions between electrons and photons of a specific sort. Strong interactions arise from essentially identical interactions between quarks and gluons, while weak interactions connect particles like the electron and the neutrino in the same way, with massive cousins of the photon known as the W and Z particles.

Differences appear at long distances for subtle reasons. The electromagnetic interaction was the first to be detected and understood because the photon is massless and the interaction is long-ranged. The W and Z particles are massive, thus mediating an interaction with a short range of about 10^{-17} cm. The difference with quarks and gluons is more subtle still: the virtual effects of the cloud of gluons surrounding a quark make the "strong charge" of quarks slowly grow stronger at longer distances. At a distance of roughly 10^{-14} cm, the interaction is so strong as to permanently confine quarks inside protons and neutrons.

But from a fundamental short-distance perspective, these are details: the character of the laws is essentially identical. This fact illustrates the central reason why we probe short distances in fundamental physics. It is not so much because we care about the "building blocks of matter" and the associated set of particles we may discover, but because we have learned that the essential unity, simplicity, and beauty of the underlying laws manifest most clearly at short distances.

The Standard Model is one of the triumphs of physics in the twentieth century. It gives us a simple and quantitatively accurate description of everything we know about elementary particles and their interactions. Only one element of the theory has yet to be definitely confirmed by experiment. In the fundamental short-distance theory, where all the interactions are treated on a symmetrical footing, the particles are massless. The mass of particles, such as electrons or the *W* and *Z* particles, arises as a dynamic long-distance effect, known as the Higgs mechanism because of the particles' interactions with the so-called Higgs field. The typical length scale associated with these interactions is around 10^{-17} cm, which is, not coincidentally, also the range of weak interactions. As I discuss at greater length below, it is also fortuitously the distance scale we are now probing with the Large Hadron Collider (LHC), the particle accelerator located at the CERN laboratory just outside Geneva, Switzerland. Collisions at the LHC should put ripples in the Higgs field that manifest as the Higgs particle with very definite properties and experimental signatures. Indeed, last December, the LHC experiments reported preliminary evidence for events consistent with the production of the Higgs particle, with its expected properties. Analysis of the 2012 data should either yield a definitive discovery of the Higgs particle or definitively exclude the simplest realization of the Higgs mechanism within the Standard Model.

Evidence of the Higgs Boson was found

The success of the Standard Model gives us a strong indication that we are headed in the right direction in our understanding of fundamental physics. Yet profound mysteries remain, associated with questions that either lie outside the scope of the Standard Model or are addressed by it, but in a seemingly absurd way. Two of these questions stand out for both their simplicity and urgency, and will drive the development of fundamental physics in the twenty-first century.

The principle of locality – the notion that interactions take place at points in spacetime – is one of the two pillars of quantum field theory. It is therefore unsettling to realize that, due to the effects of both gravity and quantum mechanics, spacetime is necessarily an approximate notion that must emerge from more primitive building blocks.

Because of the uncertainty principle, we have to use high energies to probe short distances. In a world without gravity, we could resolve arbitrarily small distances in this way, but gravity eventually and dramatically changes the picture. At miniscule distances, so much energy has to be concentrated into such a tiny region of space that the region itself collapses into a black hole, making it impossible to extract any information from the experiment. This occurs when we attempt to probe distances around 10^{-33} cm, the so-called Planck length.

The Planck length is a ridiculously tiny distance scale – sixteen orders of magnitude smaller than the tiniest distances we are probing today at the LHC. Its tininess is a direct reflection of the extreme weakness of gravity compared to other forces of nature. The gravitational attraction between a pair of electrons is forty-two orders of magnitude smaller than their electrical repulsion. Classically, both the gravitational and electric forces vary with distance following an inverse-square law; however, at a distance of around 10^{-11} cm, this gets corrected in an important way: again because of the uncertainty principle, simply holding two electrons at shorter distances requires a huge amount of energy. The force of gravity increases with increasing mass, or with equivalently increasing energy, so the attraction between electrons begins to increase relative to the electrical repulsion. At around 10^{-31} cm, gravity surpasses the electric force, and at 10^{-33} cm, it dominates all interactions.

Thus, the combination of gravity and quantum mechanics makes it impossible to operationally probe Planckian distances. Every time we have encountered ideas in physics that cannot even in principle be observed, we have come to see such ideas as approximate notions. However, this instance is particularly disturbing because the notion that emerges as approximate is that of space-time itself.

The description of the situation seems to relegate all the mysteries to tiny distances, and may suggest some sort of granular structure to space-time near the Planck scale. Much as the smooth surface of a table is resolved into discrete units made of molecules and atoms, one might imagine that "atoms of space-time" will replace space-time near the Planck length. This naive idea is very likely wrong. Any sort of granular structure to space-time picks a preferred frame of reference, where the size of the granularity is "small," in sharp conflict with the laws of relativity. But there is a deeper reason to suspect that something much more interesting and subtle than "atoms of space-time" is at play. The problems with space-time are not only localized to small distances; in a precise sense, "inside" regions of space-time cannot appear in any fundamental description of physics at all.

The slogan is that due to quantum mechanics and gravity, there are no "local observables." Indeed, before worrying about what a correct theory combining quantum mechanics and gravity ought to look like, it is worth thinking about what perfectly precise measurements can ever be made by experiments. These (in principle) exact observables provide a target for what the theory should predict.

Imagine trying to perform any sort of local measurement, by which I mean an experiment that can be done in a finite-sized room. To extract a perfectly precise measurement, we need (among other things) to use an infinitely large apparatus in order to avoid inaccuracies arising from the quantum fluctuations of the apparatus. If the apparatus has a large but finite number of components, on a huge but finite timescale, it suffers its own quantum fluctuations, and therefore cannot record the results of the experiment with perfect accuracy. Without gravity, nothing would stop us from conducting the experiment with an infinitely big apparatus to achieve perfect accuracy, but gravity obstructs this. As the apparatus gets bigger, it inevitably also gets heavier. If we are making a local measurement in a finite-sized room, at some large but finite size it becomes so heavy that it collapses the entire room into a black hole.

This means that there is no way, not even in principle, to make perfectly accurate local measurements, and thus local observables cannot have a precise meaning. There is an irreducible error associated with any local measurement that is made in a finite room. While this error is significant close to the Planck scale, it is negligible in ordinary circumstances. But this does not diminish the importance of this observation. The fact that quantum mechanics makes it impossible to determine precisely the position and velocity of a baseball is also irrelevant to a baseball player. However, it is of fundamental importance to physics that we cannot speak precisely of position *and* momentum, but only position *or* momentum. Similarly, the fact that gravity makes it impossible to have precise local observables has the dramatic consequence that the "inside" of any region of space-time does not have a sharp meaning, and is likely an approximate notion that cannot appear in a deeper underlying theory.

If we cannot speak precisely of local observables, what observables can we talk about? Instead of performing observations inside some region of space-time, we can push our detectors out to infinite distances, at the boundary of space-time, where we can make them infinitely big. We can then throw particles into the interior, where they interact and scatter with each other in some way and emerge back out to infinity where they are measured. The results of these scattering experiments can be the perfectly precise observables that one might hope to calculate from a fundamental underlying theory.

String theory is our best attempt to make sense of the mysteries of quantum gravity, and it perfectly exemplifies this basic ideology. In its earliest incarnation, string theory computed the results of scattering processes and was thought of as a generalization of quantum field theory, with point-like particles replaced by extended loops of string. This idea miraculously passed several physical and mathematical consistency checks and spawned a huge amount of theoretical activity. The 1990s brought a steady stream of surprises revealing that string theory is not in fact a theory of strings, but contains both point-like particles as well as higher-dimensional objects as important ingredients.

By the late 1990s, these developments led to an amazing realization, widely considered to be the most important theoretical advance in the field in the past two decades. Early work in string theory focused on understanding scattering processes in flat space-time, where time marches uniformly from the infinite past to the infinite future and where space is not curved. But it is also possible to consider a different kind of geometry on very large scales, known as anti-de Sitter space. Here, time still marches uniformly from the infinite past to the infinite future, but space is curved. While the distance from a point on the interior to the boundary of space is infinite, due to the curvature, a light beam takes a finite amount of time to make it to the boundary. Thus, this geometry can be usefully thought of as the inside of a box.

There is a rich set of observables that we can talk about in this geometry: starting on the walls, we can throw particles into the interior of the box and watch them come back out to the walls at some finite time in the future. Because these experiments start and end on the walls, it is natural to wonder whether there is a way of describing the physics where the interior of the box makes no appearance.

Amazingly, such a description exists, and is given in terms of a completely ordinary quantum field theory living on the walls of the box, made from particles very much like the quarks and gluons of the strong interactions. When the interactions between the "gluons" are made very strong, the physics is completely equivalent to that of string theory living on the inside of the box. In a specific sense, gravity, strings, and an extra direction of space emerge from the strong interactions of a perfectly ordinary quantum field theory in one lower dimension, much like an apparently three-dimensional image can be encoded in a two-dimensional hologram.

At first sight, this holographic equivalence seems impossible. If we had a ball in the middle of the box, how could its position in the interior be encoded only on the walls? The presence of the ball in the interior is represented as some lump of energy in the description on the walls; as the ball moves around the interior, this lump correspondingly shrinks and grows in size. What about the force of gravity between two balls in the interior? The two corresponding lumps of energy modify the virtual cloud of gluons surrounding them, which in turn induces a net attraction between the lumps, precisely reproducing the correct gravitational force. In every physical sense, gravity and the extra direction of space making up the inside of the box do indeed emerge "holographically," from the dynamics of the theory that lives fundamentally on the walls. This correspondence gives us our first concrete clue as to how space-time may emerge from more primitive building blocks.

For the past hundred years, physics has been telling us that there are fewer and fewer observables we can talk about meaningfully. The transition from classical to quantum physics was the most dramatic in this regard: the infinite number of observables in a deterministic universe was reduced to merely computing probabilities. But this loss came with a silver lining: if there are fewer fundamental observables, seemingly disparate phenomena must be more closely related and unified than they appear to be. In this case, the loss of determinism was directly responsible for understanding waves and particles in a unified way. Adding gravity to the mix further eliminates all local observables and pushes the meaningful questions to the boundary of space-time, but this is also what allows gravity and quantum field theory to be holographically equivalent to each other. It is gratifying to see that all the major themes of theoretical physics over the past four decades, in quantum field theory and string theory, have been exploring different aspects of a single underlying structure. But can this theoretical discovery be applied to understanding quantum gravity in the real world? The box in which the gravitational theory lives can be arbitrarily large; indeed, if we did not know about cosmology, we might easily imagine that our universe is a box of this sort, with a size of about ten billion light years. Any questions about gravity and quantum mechanics on shorter scales, from the size of galaxies down to the Planck length, can be asked equally well in this toy box as in our own universe.

But a number of conceptual challenges must be overcome to describe the universe we actually live in, and most of them have to do with a deeper understanding of time. Indeed, the major difference between our universe and the "gravity in a box" toy model we have understood so well is that we do not live in a static universe. Our universe is expanding. Looking back in time, we eventually encounter Planckian space-time curvatures near the "big bang," where all our descriptions of physics break down along with the notion of time itself.

An equally profound set of questions is associated with understanding the universe at late times. Perhaps the most important experimental finding in fundamental physics in the past twenty years has been the discovery that the universe's expansion rate is accelerating and that the universe is growing exponentially, doubling in size every ten billion years or so. Due to this exponential growth, light from regions of space more than ten billion light years away will never make it to us: the finite part of the universe we now see is all we will ever have access to. This simple observation has huge implications. As discussed above, precise observables require a separation of the world into a) an infinitely large measuring apparatus and b) the system being studied. In our accelerating universe, with access to only a finite (though enormous) amount of material, it is impossible to make an infinitely large apparatus. Thus, we appear to have no precise observables to talk about. So what sort of fundamental theory should we be looking for to describe this situation? This is perhaps the deepest conceptual problem we face in physics today. Any progress on this question must involve some essentially new insight into the nature of time.

Having scaled these dizzyingly abstract heights, let us come back down to Earth and ask another set of far simpler seeming questions. One of the most obvious and important properties of the universe is that it is enormous compared to the tiny distance scales of fundamental physics, from atoms and nuclei all the way down to the Planck length. This big universe is also filled with interesting objects that are much larger than atoms. Why is there a macroscopic universe when the basic constituents of matter and all the fundamental distance scales are microscopic?

This question does not at first seem particularly profound: things are big because they are composed of a huge number of atoms. But this is not the whole story. In fact, things are big as a direct consequence of the extreme weakness of gravity relative to other forces in nature. Why is the Earth big? Its size is determined by competition between an attractive gravitational pressure that is counterbalanced by atomic pressures; planets can be so big precisely because gravity is an extremely weak force. Stars are big for a similar reason. If the Planck length were comparable to the scales of atomic and nuclear physics, gravity would be a vastly stronger force, and our planets and stars would all be crushed into black holes. Thus, instead of asking why there is a macroscopic universe, we could ask: why is Planck length so much smaller than all the other scales in physics?

This turns out to be a very deep question. One might think that the scales simply are what they are, and can easily be arranged to be vastly different from each other. But this is not the case. Huge quantum fluctuations near the Planck length seem to make it impossible for macroscopic phenomena to be coherent on larger distance scales.

The most dramatic puzzle arises from the energy carried by quantum fluctuations. Fluctuations in a box of Planckian size should carry Planckian energy, leading us to expect that the vacuum will have some energy density. This vacuum energy density is known as the *cosmological constant*, and we have estimated that it should be set by the Planck scale. Like all other forms of matter and energy, the vacuum energy curves space-time; if the cosmological constant is Planckian, the curvatures should also be Planckian, leading to the absurd conclusion that the universe should be crumpled up near 10^{-33} cm, or should be expanding at an explosive rate, doubling in size every 10^{-43} seconds. Obviously, this looks nothing like the universe we live in. As already mentioned, the expansion rate of our universe is in fact accelerating, but the universe is doubling in size every ten billion years or so. The simplest explanation for this acceleration is a small positive cosmological constant, with a size 120 orders of magnitude smaller than our Planckian estimate. This is the largest disagreement between a “back of the envelope” estimate and reality in the history of physics – all the more disturbing in a subject accustomed to twelve-decimal-place agreements between theory and experiment.

Before addressing more sophisticated questions, our description of nature given by the Standard Model must deal with the extremely basic question of why the universe is big. We have found a huge contribution to the cosmological constant from quantum fluctuations, but there can also be a purely classical part of the cosmological constant, whose size just so happens to delicately cancel the contributions from quantum fluctuations, to an accuracy of 120 decimal places. This is a deeply unsatisfying explanation, and for obvious reasons is referred to as *unnatural fine-tuning* of the parameters of the theory. The fine-tuning needed to understand why we have a big universe is known as the *cosmological constant problem*.

There is an analogous puzzle known as the *hierarchy problem*, related to the question of why atomic scales are so much larger than the Planck length. The relatively large size of the atom is a consequence of the small mass of the electron. As briefly reviewed above, an electron acquires its mass from bumping into the Higgs field, with a typical interaction length near 10^{-17} cm. But the Higgs field itself should have enormous quantum fluctuations growing stronger toward the Planck scale, and so the typical length scale of its interactions with an electron should be closer to 10^{-33} cm. This outcome would make electrons sixteen orders of magnitude heavier than they are observed to be. To avoid this conclusion, we have to invoke another unnatural fine-tuning in the parameters of the theory, this time to an accuracy of one part in 10^{30}.

Unlike the difficulties with the ideas of space-time near the Planck length, these so-called naturalness problems do not represent a direct breakdown of our understanding of the laws of nature. But the extremely delicate adjustment of parameters needed to answer such basic questions seems incredibly implausible, suggesting that we are missing crucial new physical principles to provide a more satisfying explanation for why we have a macroscopic universe. It is as though we see a pencil standing on its tip in the middle of a table. While this scenario is not impossible, if we were confronted with this sight we would seek an explanation, looking for some mechanism that stabilizes the pencil and prevents it from falling over. For instance, we might look to see if the pencil is secretly hanging from a string attached to the ceiling.

The most obvious resolution to these fine-tuning problems would be to find an extension of the Standard Model that somehow removes large vacuum fluctuations. Because these fluctuations are an intrinsic feature of the unification of quantum mechanics and space-time, it stands to reason that any mechanism for removing them must change one of these two pillars of quantum field theory in some essential way; therefore, we can start by asking whether such modifications are even theoretically possible. Quantum mechanics is an extremely rigid theoretical structure, and in the past eight decades, no one has discovered a way to modify its principles even slightly. However, theorists have found an essentially unique theoretical structure – *supersymmetry* – that can extend our notion of space-time in a new way.

Theories with supersymmetry are a special kind of quantum field theory that can be thought of as extending our usual four dimensions of space and time by four additional dimensions. The novelty is that distances in these extra dimensions are not measured by ordinary numbers, but by quantum variables: in a sense, supersymmetry makes space-time more intrinsically quantum-mechanical. Ordinary distances satisfy the basic multiplication law $a \times b = b \times a$, and are said to be *commuting* variables. However, distances in the quantum dimensions satisfy instead $a \times b = -b \times a$, with the crucial minus sign, and are said to be *anti-commuting*. In particular, $a \times a = -a \times a = 0$. Because of this, it is impossible to take more than a single "step" into the quantum dimensions. An electron can move around in our ordinary four dimensions, but it can also take this single step into the quantum dimensions. From the four-dimensional point of view, it will appear to be another particle, the *superpartner* of the electron, with the same mass and charge but different in its magnetic properties. The "symmetry" part of supersymmetry demands that the interactions respect a perfect symmetry between the ordinary and the quantum dimensions.

Supersymmetry is a deep idea that has played a major role in theoretical physics for the past forty years. It is an essential part of string theory, it has helped revolutionize our understanding of quantum field theory, and along the way it has opened up many new connections between physics and mathematics. Among its many remarkable properties, the one relevant to our discussion is that supersymmetry eliminates large vacuum quantum fluctuations in a beautiful way. The inability to take more than a single step into the quantum dimensions means that there can be no wild fluctuations in the quantum dimensions; and because the quantum and ordinary dimensions must be treated symmetrically, there can be no large fluctuations in the ordinary dimensions either. More technically, the large fluctuations from the ordinary particles are perfectly canceled by their superpartners.

Of course, there is a catch: we haven't observed any of the superpartners for the ordinary particles! It is possible, however, that physics at short distances is supersymmetric, but that the perfect symmetry between ordinary and quantum dimensions is hidden by the same kind of long-distance illusion that hides the essential unity of strong, weak, and electromagnetic interactions. This long-distance "breaking" of supersymmetry has the effect of making the superpartners heavier than the ordinary particles we have seen, similar to how the W and Z particles are heavy while the photon is massless.

Can broken supersymmetry still address the fine-tuning problems? If nature becomes supersymmetric at around 10^{-17} cm, then the large quantum fluctuation in the Higgs field will be removed, yielding a completely natural resolution of the hierarchy problem. While there are a few other approaches to the hierarchy problem, supersymmetry is the most compelling, and there are some strong quantitative (though circumstantial) hints that it is on the right track. Whether it is supersymmetry or something else, a natural solution of the hierarchy problem demands *some* sort of new physics at around 10^{-17} cm. If nothing new happens until, say, 10^{-20} cm, then the quantum fluctuation of the Higgs field will be dragged to 10^{-20} cm. In order to make the actual interaction range of 10^{-17} cm natural, something new must show up at just this scale. This is why it is particularly exciting that we are probing exactly these distances at the LHC.

What about the much more severe cosmological constant problem? The cosmological constant is so tiny that its associated length scale is around a millimeter, and nature is clearly not supersymmetric at the millimeter scale. Supersymmetry does improve the fine-tuning problem for the cosmological constant from one part in 10^{120} to one part in 10^{60}, but this is small consolation. The difficulty is not just with supersymmetry: we have not seen any sort of new physics at the millimeter scale, so there is no hint that the cosmological constant problem is solved in a natural way.

This enormous challenge has led some theorists to imagine a different sort of explanation for fine-tuning problems, involving a radical change to our picture of space-time not at short distances, but at huge scales larger than the size of our observable universe. The idea takes some inspiration from developments in string theory over the last decade. String theory is a unique mathematical structure, but it has long been known that it has many different solutions, or *vacua*, each of which corresponds to a different possible long-distance world. The basic laws of nature are the same in all vacua, but the menu of particles and interaction strengths changes from vacuum to vacuum. The new realization is that the number of vacua with broken supersymmetry – the ones that might roughly resemble our world – is gargantuan: a rough estimate is that 10^{500} such vacua may exist. Furthermore, an important idea in cosmology, known as *eternal inflation*, makes it possible that all these vacua are actually realized somewhere in space-time. Many of these vacua have positive cosmological constants and are undergoing exponential expansion. Quantum mechanics enables bubbles of a new vacuum to form in this cosmology.

The bubble containing this "daughter" vacuum grows at nearly the speed of light and would naively appear to consume the "parent" vacuum. But this does not happen: because the parent is growing exponentially, it is never completely swallowed up, and it continues its exponential expansion forever. Thus, all possible daughter vacua are produced, giving rise to the picture of an infinite *multiverse* where all vacua are produced, infinitely often, somewhere in space-time.

In most of these vacua, the cosmological constant is enormous; but these vacua also undergo explosive accelerated expansion that would rip apart all structures, so in these regions the universe would be empty. However, there are so many vacua that, statistically speaking, some of them will have a small cosmological constant. It is only in those regions that the universe is not empty, and so it is not surprising that we should find ourselves there.

This picture is currently the only reasonable explanation that we have for the smallness of the cosmological constant, and it is not impossible that similar considerations may also be relevant for the hierarchy problem. So, is our universe just a tiny part of a vast and mostly lethal multiverse? If this picture is correct, it would be a further extension of the Copernican revolution. However, a number of major conceptual challenges must be overcome to determine whether these ideas make coherent sense, even on purely theoretical grounds. Because our own universe is accelerating, we can never see the other regions in the multiverse, and so it is not obvious that we can talk about these regions in a physically and mathematically meaningful way. But it is also not impossible to make proper sense of this picture. This has been an active area of research in the last decade, although serious theoretical progress on these problems still seems rather distant. Once again, the thorniest questions lie at the intersection of quantum mechanics, gravity, and cosmology.

What might we expect to learn from experiments in the coming decade? The Large Hadron Collider is perhaps the most important experiment today, pushing the frontiers of fundamental physics. The accelerator itself is housed in a tunnel a hundred meters underground, with a circumference of twenty-seven kilometers. The tunnel contains a ring, where two sets of protons, moving in opposite directions, are accelerated to a speed 0.9999999 times the speed of light. The protons are made to collide head-on at two points around the ring, which are surrounded by enormous detectors. Two teams, each consisting of three thousand physicists, study the debris from these collisions, which give us a direct window into the laws of nature at distances of order 10^{-17} cm, an order of magnitude better than we have probed before.

As mentioned, a proton is not a point-like particle, but is a messy 10^{-14} cm bag containing quarks that are permanently trapped inside by gluons. When two of these messy bags collide at enormous energies, they usually break up into other messy collections of strongly interacting particles, zooming along in the initial direction of the beams. These typical interactions are not our main interest in probing short-distance physics. Rather, we are after the head-on collisions between the quarks and gluons in one proton and the quarks and gluons in the other. The telltale sign that a head-on collision has occurred is that particles scatter off at large angles relative to the initial direction of the beams. The collision can also produce energy enough to create new heavy particles and anti-particles.

Any new particles will typically be unstable, decaying on a timescale of order 10^{-27}

seconds into ordinary particles like electrons and positrons, quarks and antiquarks, and so on. These decay products will also spray off at large angles relative to the initial direction of the beam. Thus, studying all the debris from the high-energy collisions that come off at large angles is, in general, the best probe we have for studying short-distance physics. Having this rough means to discriminate "typical" and "interesting" events is crucial because the interesting events are exceedingly rare relative to the typical ones. There are about a billion typical collisions per second, whereas the timescale for producing, say, supersymmetric particles is expected to be in the range of a few per minute to a few per hour. The debris from these collisions are collected by the huge detectors and studied in great detail to look for the proverbial needle in the haystack.

The first order of business at the LHC is the search for the Higgs particle. As noted, analysis of the 2012 data should either definitively confirm or definitively exclude its existence. (Most physicists expect the former, especially following the solid hint reported in December 2011.) Assuming that the existence of the Higgs particle is confirmed, an accurate measurement of the rate at which it is produced, and the way it interacts with other particles, will shed light on whether it behaves as expected in the Standard Model, or whether it has modified properties that would indicate new physics.

The search for supersymmetry, or some other natural mechanism that would solve the hierarchy problem, is another central goal of the LHC program. The collision between quarks can have sufficiently high energy to pop the quarks into quantum dimensions and produce *squarks*, which rapidly decay to ordinary particles and other superpartners. In the simplest versions of the theory, the lightest of all the superpartners is a stable, electrically neutral particle that is so weakly interacting it sails through the detectors without leaving a trace. Thus, supersymmetric events should have the distinctive feature of seeming to have "missing" energy and momentum. No evidence for superpartners has yet emerged in the data, and the searches are beginning to encroach on the territory where superpartners must show up, if the supersymmetry indeed naturally solves the hierarchy problem.

After running through 2012, the LHC will stop and restart operations in 2014–2015 with twice its current energy. What might we know by 2020? The discovery of supersymmetry would represent the first extension of our notion of spacetime since Einstein and would confirm one of the most important theoretical ideas of the past forty years. We would also find a completely satisfactory understanding of the question, why is gravity weak? On the other hand, if neither supersymmetry nor any other sort of natural solution to the hierarchy problem appears in the data, the situation will be much more confusing. We will have solid experimental evidence for fine-tuning in the parameters that determine elementary particle masses, something we have never seen in such dramatic fashion. This would strongly resonate with the apparently enormous fine-tuning problems associated with the cosmological constant, and would give theorists a strong incentive to take the ideas of the multiverse more seriously.

It should be clear that we have arrived at a bifurcatory moment in the history of fundamental physics, a moment that has enormous implications for the future of the subject. With many theoretical speculations pointing in radically different directions, it is now up to experiment to render its verdict!

The twentieth century was dominated by the ideas of relativity and quantum mechanics, and their synthesis is quantum field theory. As I have discussed, there are strong reasons to think that some essentially new ideas are needed in the twenty-first century. The LHC is poised to shed significant light on the question of why a macroscopic universe exists, but the questions having to do with the deeper origin of space-time seem tied to the Planck scale, offering little hope for direct clues from experiment in the near future. Even so, the requirements of physical and mathematical consistency have provided a strong guide to the theoretical investigation of these questions. Indeed, the spectacular progress in string theory over the last four decades, which has time and again surprised us with unanticipated connections between disparate parts of physics and mathematics, has been driven in this way. Today, however, we confront even deeper mysteries, such as coming to grips with emergent time and the application of quantum mechanics to the entire universe. These challenges call for a bigger shift in perspective. Is there any hope for taking such large steps without direct input from experiment?

We can take some inspiration by looking at the path that led from classical physics to relativity and quantum mechanics. Some of the crucial clues to future developments were lying in plain sight, in the structure of existing theories. Einstein's motivations for developing both special and general relativity were rooted in "obvious" properties of classical physics. Newton's laws already had a notion of Galilean relativity. However, Galilean relativity allowed for arbitrarily large signal velocities and thus action at a distance. This was in conflict with Maxwell's laws of electromagnetism, in which the interactions involving electromagnetic fields were local. Einstein resolved this purely theoretical conflict between the two pillars of classical physics by realizing that the Galilean notion of relativity had to be deformed to one that was compatible with a maximal speed for signal propagation and thus with locality.

The loss of determinism in passing from classical to quantum mechanics was a much more radical change in our picture of the world, and yet even this transition was presaged in classical physics. Newton's laws are manifestly deterministic; given the initial position and velocity of a particle, together with all the forces acting on it, the laws of motion tell us where the particle goes in the next instant of time. However, in the century after Newton, physicists and mathematicians discovered a reformulation of Newton's laws that led to exactly the same equations, but from a completely different philosophical starting point. Of all the possible trajectories a particle can take from A to B, it chooses the one that minimizes the average value of difference between the kinetic and potential energies along the path, a quantity known as the *action* of the path. This law does not look deterministic: the particle seems to sniff out all possible paths it could take from A to B and then chooses the one that minimizes the action. But it turns out that the paths that minimize the action are precisely the ones that satisfy Newton's laws.

Why should it be possible to talk about Newton's laws in such a different way, which seems to hide their most essential feature of deterministic evolution in time? We now know the deep answer to this question is that the world is quantum-mechanical. As Richard Feynman pointed out in the mid-1940s, a quantum-mechanical particle takes all possible paths from A to B; in the classical limit, the dominant contributions to the probability are peaked on the trajectories that minimize the action, which are, secondarily, the ones that satisfy Newton's laws. Since quantum mechanics is not deterministic, the classical limit of the theory could not land on Newton's laws, but instead lands on a different formulation of classical physics in which determinism is not manifest but rather is a secondary, derived notion.

If there are any clues hiding in plain sight today, they are lurking in the many astonishing properties of quantum field theory and string theory that have been uncovered over the past two decades. The founders of quantum field theory could never have imagined that it might describe a theory of gravity in a higher-dimensional curved space, and yet it does. We have learned that theories that seem completely different from the classical perspective are secretly identical at the quantum-mechanical level. Many of these developments have uncovered deep connections between physics and modern mathematics. Even "bread and butter" calculations in field theory, needed to understand the strong interaction processes at the LHC, have revealed major surprises. Textbook calculations for the rates of these processes quickly lead to hundreds of pages of algebra, yet in recent years we have understood that the final expressions can fit on a single line. These simplifications are associated with a new set of completely hidden symmetries enjoyed by ordinary quantum field theories. They have been sitting under our noses undetected for sixty years, and now they are exposing connections to yet another set of new structures in mathematics.

Thus, while we may not have experimental data to tell us about physics near the Planck scale, we do have an ocean of "theoretical data" in the wonderful mathematical structures hidden in quantum field theory and string theory. These structures beg for a deeper explanation. The standard formulation of field theory hides these amazing features as a direct consequence of its deference to space-time locality. There must be a new way of thinking about quantum field theories, in which space-time locality is not the star of the show and these remarkable hidden structures are made manifest. Finding this reformulation might be analogous to discovering the least-action formulation of classical physics; by removing space-time from its primary place in our description of standard physics, we may be in a better position to make the leap to the next theory, where space-time finally ceases to exist.