12

The World on a String

THE FABRIC ACCORDING TO STRING THEORY

Imagine a universe in which to understand anything you'd need to understand everything. A universe in which to say anything about why a planet orbits a star, about why a baseball flies along a particular trajectory, about how a magnet or a battery works, about how light and gravity operate—a universe in which to say anything about anything—you would need to uncover the most fundamental laws and determine how they act on the finest constituents of matter. Thankfully, this universe is not our universe.

If it were, it's hard to see how science would have made any progress at all. Over the centuries, the reason we've been able to make headway is that we've been able to work piecemeal; we've been able to unravel mysteries step by step, with each new discovery going a bit deeper than the previous. Newton didn't need to know about atoms to make great strides in understanding motion and gravity. Maxwell didn't need to know about electrons and other charged particles to develop a powerful theory of electromagnetism. Einstein didn't need to address the primordial incarnation of space and time to formulate a theory of how they curve in the service of the gravitational force. Instead, each of these discoveries, as well as the many others that underlie our current conception of the cosmos, proceeded within a limited context that unabashedly left many basic questions unanswered. Each discovery was able to contribute its own piece to the puzzle, even though no one knew—and we still don't know—what grand synthesizing picture comprises all the puzzle's pieces.

A closely related observation is that although today's science differs sharply from that of even fifty years ago, it would be simplistic to summarize scientific progress in terms of new theories overthrowing their predecessors. A more correct description is that each new theory refines its predecessor by providing a more accurate and more wide-reaching framework. Newton's theory of gravity has been superseded by Einstein's general relativity, but it would be naïve to say that Newton's theory was wrong. In the domain of objects that don't move anywhere near as fast as light and don't produce gravitational fields anywhere near as strong as those of black holes, Newton's theory is fantastically accurate. Yet this is not to say that Einstein's theory is a minor variant on Newton's; in the course of improving Newton's approach to gravity, Einstein invoked a whole new conceptual schema, one that radically altered our understanding of space and time. But the power of Newton's discovery within the domain he intended it for (planetary motion, commonplace terrestrial motion, and so on) is unassailable.

We envision each new theory taking us closer to the elusive goal of truth, but whether there is an ultimate theory—a theory that cannot be refined further, because it has finally revealed the workings of the universe at the deepest possible level—is a question no one can answer. Even so, the pattern traced out during the last three hundred years of discovery gives tantalizing evidence that such a theory can be developed. Broadly speaking, each new breakthrough has gathered a wider range of physical phenomena under fewer theoretical umbrellas. Newton's discoveries showed that the forces governing planetary motion are the same as those governing the motion of falling objects here on earth. Maxwell's discoveries showed that electricity and magnetism are two sides of the same coin. Einstein's discoveries showed that space and time are as inseparable as Midas' touch and gold. The discoveries of a generation of physicists in the early twentieth century established that myriad mysteries of microphysics could be explained with precision using quantum mechanics. More recently, the discoveries of Glashow, Salam, and Weinberg showed that electromagnetism and the weak nuclear force are two manifestations of a single force—the electroweak force—and there is even tentative, circumstantial evidence that the strong nuclear force may join the electroweak force in a yet grander synthesis.[1] Taking all this together, we see a pattern that goes from complexity to simplicity, a pattern that goes from diversity to unity. The explanatory arrows seem to be converging on a powerful, yet-to-be discovered framework that would unify all of nature's forces and all of matter within a single theory capable of describing all physical phenomena.

Albert Einstein, who for more than three decades sought to combine electromagnetism and general relativity in a single theory, is rightly credited with initiating the modern search for a unified theory. For long stretches during those decades, he was the sole searcher for such a unified theory, and his passionate yet solitary quest alienated him from the mainstream physics community. During the last twenty years, though, there has been a resurgence in the quest for a unified theory; Einstein's lonely dream has become the driving force for a whole generation of physicists. But with the discoveries since Einstein's time has come a shift in focus. Even though we don't yet have a successful theory combining the strong nuclear force and the electroweak force, all three of these forces (electromagnetic, weak, strong) have been described by a single uniform language based on quantum mechanics. But general relativity, our most refined theory of the fourth force, stands outside this framework. General relativity is a classical theory: it does not incorporate any of the probabilistic concepts of quantum theory. A primary goal of the modern unification program is therefore to combine general relativity and quantum mechanics, and to describe all four forces within the same quantum mechanical framework. This has proven to be one of the most difficult problems theoretical physics has ever encountered.

Let's see why.

Quantum Jitters and Empty Space

If I had to select the single most evocative feature of quantum mechanics, I'd choose the uncertainty principle. Probabilities and wavefunctions certainly provide a radically new framework, but it's the uncertainty principle that encapsulates the break from classical physics. Remember, in the seventeenth and eighteenth centuries, scientists believed that a complete description of physical reality amounted to specifying the positions and velocities of every constituent of matter making up the cosmos. And with the advent of the field concept in the nineteenth century, and its subsequent application to the electromagnetic and gravitational forces, this view was augmented to include the value of each field—the strength of each field, that is—and the rate of change of each field's value, at every location in space. But by the 1930s, the uncertainty principle dismantled this conception of reality by showing that you can't ever know both the position and the velocity of a particle; you can't ever know both the value of a field at some location in space and how quickly the field value is changing. Quantum uncertainty forbids it.

As we discussed in the last chapter, this quantum uncertainty ensures that the microworld is a turbulent and jittery realm. Earlier, we focused on uncertainty-induced quantum jitters for the inflaton field, but quantum uncertainty applies to all fields. The electromagnetic field, the strong and weak nuclear force fields, and the gravitational field are all subject to frenzied quantum jitters on microscopic scales. In fact, these field jitters exist even in space you'd normally think of as empty, space that would seem to contain no matter and no fields. This is an idea of critical importance, but if you haven't encountered it previously, it's natural to be puzzled. If a region of space contains nothing—if it's a vacuum—doesn't that mean there's nothing to jitter? Well, we've already learned that the concept of nothingness is subtle. Just think of the Higgs ocean that modern theory claims to permeate empty space. The quantum jitters I'm now referring to serve only to make the notion of "nothing" subtler still. Here's what I mean.

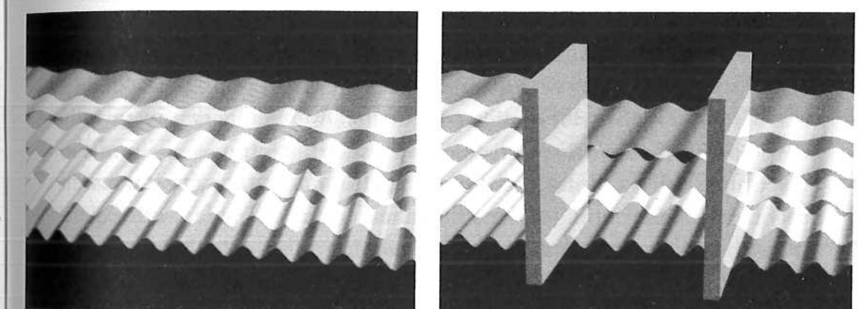
In prequantum (and pre-Higgs) physics, we'd declare a region of space completely empty if it contained no particles and the value of every field was uniformly zero.* Let's now think about this classical notion of emptiness in light of the quantum uncertainty principle. If a field were to have and maintain a vanishing value, we would know its value—zero—and also the rate of change of its value—zero, too. But according to the uncertainty principle, it's impossible for both these properties to be definite. Instead, if a field has a definite value at some moment, zero in the case at hand, the uncertainty principle tells us that its rate of change is completely random. And a random rate of change means that in subsequent moments the field's value will randomly jitter up and down, even in what we normally think of as completely empty space. So the intuitive notion of emptiness, one in which all fields have and maintain the value

*For ease of writing, we'll consider only fields that reach their lowest energy when their values are zero. The discussion for other fields—Higgs fields—is identical, except the jitters fluctuate about the field's *nonzero*, lowest-energy value. If you are tempted to say that a region of space is empty only if there is no matter present and all fields are *absent*, not just that they have the value zero, see notes section.[2]

zero, is incompatible with quantum mechanics. *A field's value can jitter around the value zero but it can't be uniformly zero throughout a region for more than a brief moment.*[3] In technical language, physicists say that fields undergo *vacuum fluctuations*.

The random nature of vacuum field fluctuations ensures that in all but the most microscopic of regions, there are as many "up" jitters as "down" and hence they average out to zero, much as a marble surface appears perfectly smooth to the naked eye even though an electron microscope reveals that it's jagged on minuscule scales. Nevertheless, even though we can't see them directly, more than half a century ago the reality of quantum field jitters, even in empty space, was conclusively established through a simple yet profound discovery.

In 1948, the Dutch physicist Hendrik Casimir figured out how vacuum fluctuations of the electromagnetic field could be experimentally detected. Quantum theory says that the jitters of the electromagnetic field in empty space will take on a variety of shapes, as illustrated in Figure 12.1a. Casimir's breakthrough was to realize that by placing two ordinary metal plates in an otherwise empty region, as in Figure 12.1b, he could induce a subtle modification to these vacuum field jitters. Namely, the quantum equations show that in the region between the plates there will be fewer fluctuations (only those electromagnetic field fluctuations whose values vanish at the location of each plate are allowed). Casimir analyzed the implications of this reduction in field jitters and found



(a) (b)

Figure 12.1 (**a**) Vacuum fluctuations of the electromagnetic field. (**b**) Vacuum fluctuations between two metal plates and those outside the plates.

something extraordinary. Much as a reduction in the amount of air in a region creates a pressure imbalance (for example, at high altitude you can feel the thinner air exerting less pressure on the outside of your eardrums), the reduction in quantum field jitters between the plates also yields a pressure imbalance: the quantum field jitters between the plates become a bit weaker than those outside the plates, and this imbalance *drives the plates toward each other.*

Think about how thoroughly odd this is. You place two plain, ordinary, uncharged metal plates into an *empty* region of space, facing one another. As their masses are tiny, the gravitational attraction between them is so small that it can be completely ignored. Since there is nothing else around, you naturally conclude that the plates will stay put. But this is *not* what Casimir's calculations predicted would happen. He concluded that the plates would be gently guided by the ghostly grip of quantum vacuum fluctuations to move toward one another.

When Casimir first announced these theoretical results, equipment sensitive enough to test his predictions didn't exist. Yet, within about a decade, another Dutch physicist, Marcus Spaarnay, was able to initiate the first rudimentary tests of this *Casimir force,* and increasingly precise experiments have been carried out ever since. In 1997, for example, Steve Lamoreaux, then at the University of Washington, confirmed Casimir's predictions to an accuracy of 5 percent.[4] (For plates roughly the size of playing cards and placed one ten-thousandth of a centimeter apart, the force between them is about equal to the weight of a single teardrop; this shows how challenging it is to measure the Casimir force.) There is now little doubt that the intuitive notion of empty space as a static, calm, eventless arena is thoroughly off base. Because of quantum uncertainty, empty space is teeming with quantum activity.

It took scientists the better part of the twentieth century to fully develop the mathematics for describing such quantum activity of the electromagnetic, and strong and weak nuclear forces. The effort was well spent: calculations using this mathematical framework agree with experimental findings to an unparalleled precision (e.g., calculations of the effect of vacuum fluctuations on the magnetic properties of electrons agree with experimental results to one part in a billion).[5]

Yet despite all this success, for many decades physicists have been aware that quantum jitters have been fomenting discontent within the laws of physics.

Jitters and Their Discontent[6]

So far, we've discussed only quantum jitters for fields that exist *within* space. What about the quantum jitters of space itself? While this might sound mysterious, it's actually just another example of quantum field jitters—an example, however, that proves particularly troublesome. In the general theory of relativity, Einstein established that the gravitational force can be described by warps and curves in the fabric of space; he showed that gravitational fields manifest themselves through the shape or geometry of space (and of spacetime, more generally). Now, just like any other field, the gravitational field is subject to quantum jitters: the uncertainty principle ensures that over tiny distance scales, the gravitational field fluctuates up and down. And since the gravitational field is synonymous with the shape of space, such quantum jitters mean that the shape of space fluctuates randomly. Again, as with all examples of quantum uncertainty, on everyday distance scales the jitters are too small to be sensed directly, and the surrounding environment appears smooth, placid, and predictable. But the smaller the scale of observation the larger the uncertainty, and the more tumultuous the quantum fluctuations become.

This is illustrated in Figure 12.2, in which we sequentially magnify the fabric of space to reveal its structure at ever smaller distances. The lowermost level of the figure shows the quantum undulations of space on familiar scales and, as you can see, there's nothing to see—the undulations are unobservably small, so space appears calm and flat. But as we home in by sequentially magnifying the region, we see that the undulations of space get increasingly frenetic. By the highest level in the figure, which shows the fabric of space on scales smaller than the *Planck length*—a millionth of a billionth of a billionth of a billionth (10^{-33}) of a centimeter—space becomes a seething, boiling cauldron of frenzied fluctuations. As the illustration makes clear, the usual notions of left/right, back/forth, and up/down become so jumbled by the ultramicroscopic tumult that they lose all meaning. Even the usual notion of before/after, which we've been illustrating by sequential slices in the spacetime loaf, is rendered meaningless by quantum fluctuations on time scales shorter than the *Planck time*, about a tenth of a millionth of a trillionth of a trillionth of a trillionth (10^{-43}) of a second (which is roughly the time it takes light to travel a Planck length). Like a blurry photograph, the wild undu-

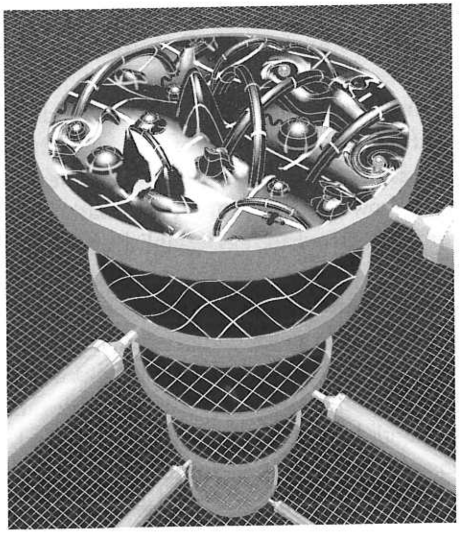


Figure 12.2 Successive magnifications of space reveal that below the Planck length, space becomes unrecognizably tumultuous due to quantum jitters. (These are imaginary magnifying glasses, each of which magnifies between 10 million and 100 million times.)

lations in Figure 12.2 make it impossible to distinguish one time slice from another unambiguously when the time interval between them becomes shorter than the Planck time. The upshot is that on scales shorter than Planck distances and durations, quantum uncertainty renders the fabric of the cosmos so twisted and distorted that the usual conceptions of space and time are no longer applicable.

While exotic in detail, the broad-brush lesson illustrated by Figure 12.2 is one with which we are already familiar: concepts and conclusions relevant on one scale may not be applicable on all scales. This is a key principle in physics, and one that we encounter repeatedly even in far more prosaic contexts. Take a glass of water. Describing the water as a smooth, uniform liquid is both useful and relevant on everyday scales, but it's an approximation that breaks down if we analyze the water with submicroscopic precision. On tiny scales, the smooth image gives way to a completely different framework of widely separated molecules and atoms. Similarly, Figure 12.2 shows that Einstein's conception of a smooth, gently curving, geometrical space and time, although powerful and accurate

for describing the universe on large scales, breaks down if we analyze the universe at extremely short distance and time scales. Physicists believe that, as with water, the smooth portrayal of space and time is an approximation that gives way to another, more fundamental framework when considered on ultramicroscopic scales. What that framework is—what constitutes the "molecules" and "atoms" of space and time—is a question currently being pursued with great vigor. It has yet to be resolved.

Even so, what is thoroughly clear from Figure 12.2 is that on tiny scales the smooth character of space and time envisioned by general relativity locks horns with the frantic, jittery character of quantum mechanics. The core principle of Einstein's general relativity, that space and time form a gently curving geometrical shape, runs smack into the core principle of quantum mechanics, the uncertainty principle, which implies a wild, tumultuous, turbulent environment on the tiniest of scales. The violent clash between the central ideas of general relativity and quantum mechanics has made meshing the two theories one of the most difficult challenges physicists have encountered during the last eighty years.

Does It Matter?

In practice, the incompatibility between general relativity and quantum mechanics rears its head in a very specific way. If you use the combined equations of general relativity and quantum mechanics, they almost always yield one answer: infinity. And that's a problem. It's nonsense. Experimenters never measure an infinite amount of anything. Dials never spin around to infinity. Meters never reach infinity. Calculators never register infinity. Almost always, an infinite answer is meaningless. All it tells us is that the equations of general relativity and quantum mechanics, when merged, go haywire.

Notice that this is quite unlike the tension between *special* relativity and quantum mechanics that came up in our discussion of quantum nonlocality in Chapter 4. There we found that reconciling the tenets of special relativity (in particular, the symmetry among all constant velocity observers) with the behavior of entangled particles requires a more complete understanding of the quantum measurement problem than has so far been attained (see pages 117–120). But this incompletely resolved issue does not result in mathematical inconsistencies or in equations that yield nonsensical answers. To the contrary, the combined equations of special relativity and quantum mechanics have been used to make the most precisely confirmed predictions in the history of science. The quiet tension between special relativity and quantum mechanics points to an area in need of further theoretical development, but it has hardly any impact on their combined predictive power. Not so with the explosive union between general relativity and quantum mechanics, in which all predictive power is lost.

Nevertheless, you can still ask whether the incompatibility between general relativity and quantum mechanics really matters. Sure, the combined equations may result in nonsense, but when do you ever really need to use them together? Years of astronomical observations have shown that general relativity describes the macro world of stars, galaxies, and even the entire expanse of the cosmos with impressive accuracy; decades of experiments have confirmed that quantum mechanics does the same for the micro world of molecules, atoms, and subatomic particles. Since each theory works wonders in its own domain, why worry about combining them? Why not keep them separate? Why not use general relativity for things that are large and massive, quantum mechanics for things that are tiny and light, and celebrate humankind's impressive achievement of successfully understanding such a wide range of physical phenomena?

As a matter of fact, this *is* what most physicists have done since the early decades of the twentieth century, and there's no denying that it's been a distinctly fruitful approach. The progress science has made by working in this disjointed framework is impressive. All the same, there are a number of reasons why the antagonism between general relativity and quantum mechanics must be reconciled. Here are two.

First, at a gut level, it is hard to believe that the deepest understanding of the universe consists of an uneasy union between two powerful theoretical frameworks that are mutually incompatible. It's not as though the universe comes equipped with a line in the sand separating things that are properly described by quantum mechanics from things properly described by general relativity. Dividing the universe into two separate realms seems both artificial and clumsy. To many, this is evidence that there must be a deeper, unified truth that overcomes the rift between general relativity and quantum mechanics and that can be applied to *everything*. We have one universe and therefore, many strongly believe, we should have one theory.

Second, although most things are either big and heavy or small and light, and therefore, as a practical matter, can be described using general

relativity *or* quantum mechanics, this is not true of all things. Black holes provide a good example. According to general relativity, all the matter that makes up a black hole is crushed together at a single minuscule point at the black hole's center.[7] This makes the center of a black hole both enormously massive and incredibly tiny, and hence it falls on both sides of the purported divide: we need to use general relativity because the large mass creates a substantial gravitational field, and we also need to use quantum mechanics because all the mass is squeezed to a tiny size. But in combination, the equations break down, so no one has been able to determine what happens right at the center of a black hole.

That's a good example, but if you're a real skeptic, you might still wonder whether this is something that should keep anyone up at night. Since we can't see inside a black hole unless we jump in, and, moreover, were we to jump in we wouldn't be able to report our observations back to the outside world, our incomplete understanding of the black hole's interior may not strike you as particularly worrisome. For physicists, though, the existence of a realm in which the known laws of physics break down—no matter how esoteric the realm might seem—throws up red flags. If the known laws of physics break down under any circumstances, it is a clear signal that we have not reached the deepest possible understanding. After all, the universe works; as far as we can tell, the universe does not break down. The correct theory of the universe should, at the very least, meet the same standard.

Well, that surely seems reasonable. But for my money, the full urgency of the conflict between general relativity and quantum mechanics is revealed only through another example. Look back at Figure 10.6. As you can see, we have made great strides in piecing together a consistent and predictive story of cosmic evolution, but the picture remains incomplete because of the fuzzy patch near the inception of the universe. And within the foggy haze of those earliest moments lies insight into the most tantalizing of mysteries: the origin and fundamental nature of space and time. So what has prevented us from penetrating the haze? The blame rests squarely on the conflict between general relativity and quantum mechanics. The antagonism between the laws of the large and those of the small is the reason the fuzzy patch remains obscure and we still have no insight into what happened at the very beginning of the universe.

To understand why, imagine, as in Chapter 10, running a film of the expanding cosmos in reverse, heading back toward the big bang. In reverse, everything that is now rushing apart comes together, and so as we run the film farther back, the universe gets ever smaller, hotter, and denser. As we close in on time zero itself, the *entire* observable universe is compressed to the size of the sun, then further squeezed to the size of the earth, then crushed to the size of a bowling ball, a pea, a grain of sand—smaller and smaller the universe shrinks as the film rewinds toward its initial frames. There comes a moment in this reverse-run film when the entire known universe has a size close to the Planck length—the millionth of a billionth of a billionth of a billionth of a centimeter at which general relativity and quantum mechanics find themselves at loggerheads. At this moment, all the mass and energy responsible for spawning the observable universe is contained in a speck that's less than a hundredth of a billionth of a billionth of the size of a single atom.[8]

Thus, just as in the case of a black hole's center, the early universe falls on both sides of the divide: The enormous density of the early universe requires the use of general relativity. The tiny size of the early universe requires the use of quantum mechanics. But once again, in combination the laws break down. The projector jams, the cosmic film burns up, and we are unable to access the universe's earliest moments. Because of the conflict between general relativity and quantum mechanics, we remain ignorant about what happened at the beginning and are reduced to drawing a fuzzy patch in Figure 10.6.

If we ever hope to understand the origin of the universe—one of the deepest questions in all of science—the conflict between general relativity and quantum mechanics *must* be resolved. We must settle the differences between the laws of the large and the laws of the small and merge them into a single harmonious theory.

The Unlikely Road to a Solution*

As the work of Newton and Einstein exemplifies, scientific breakthroughs are sometimes born of a single scientist's staggering genius, pure and simple. But that's rare. Much more frequently, great breakthroughs represent the collective effort of many scientists, each building on the insights of others to accomplish what no individual could have achieved in isolation. One scientist might contribute an idea that sets a colleague thinking, which leads to an observation that reveals an unexpected relationship that inspires an important advance, which starts anew the cycle of discovery. Broad knowledge, technical facility, flexibility of thought, openness to unanticipated connections, immersion in the free flow of ideas worldwide, hard work, and significant luck are all critical parts of scientific discovery. In recent times, there is perhaps no major breakthrough that better exemplifies this than the development of *superstring theory*.

*The remainder of this chapter recounts the discovery of superstring theory and discusses the theory's essential ideas regarding unification and the structure of spacetime. Readers of *The Elegant Universe* (especially Chapters 6 through 8) will be familiar with much of this material, and should feel free to skim this chapter and move on to the next.

Superstring theory is an approach that many scientists believe successfully merges general relativity and quantum mechanics. And as we will see, there is reason to hope for even more. Although it is still very much a work in progress, superstring theory may well be a fully unified theory of all forces and all matter, a theory that reaches Einstein's dream and beyond—a theory, I and many others believe, that is blazing the beginnings of a trail which will one day lead us to the deepest laws of the universe. Truthfully, though, superstring theory was not conceived as an ingenious means to reach these noble and long-standing goals. Instead, the history of superstring theory is full of accidental discoveries, false starts, missed opportunities, and nearly ruined careers. It is also, in a precise sense, the story of the discovery of the right solution for the wrong problem.

In 1968, Gabriele Veneziano, a young postdoctoral research fellow working at CERN, was one of many physicists trying to understand the strong nuclear force by studying the results of high-energy particle collisions produced in atom smashers around the world. After months of analyzing patterns and regularities in the data, Veneziano recognized a surprising and unexpected connection to an esoteric area of mathematics. He realized that a two-hundred-year-old formula discovered by the famous Swiss mathematician Leonhard Euler (the *Euler beta function*) seemed to match data on the strong nuclear force with precision. While this might not sound particularly unusual—theoretical physicists deal with arcane formulae all the time—it was a striking case of the cart's rolling miles ahead of the horse. More often than not, physicists first develop an intuition, a mental picture, a broad understanding of the physical principles underlying whatever they are studying and only then seek the equations necessary to ground their intuition in rigorous mathematics. Veneziano, to the contrary, jumped right to the equation; his brilliance was to recognize unusual patterns in the data and to make the unanticipated link to a formula devised centuries earlier for purely mathematical interest.

But although Veneziano had the formula in hand, he had no explanation for *why* it worked. He lacked a physical picture of why Euler's beta function should be relevant to particles influencing each other through the strong nuclear force. Within two years the situation completely changed. In 1970, papers by Leonard Susskind of Stanford, Holger Nielsen of the Niels Bohr Institute, and Yoichiro Nambu of the University of Chicago revealed the physical underpinnings of Veneziano's discovery. These physicists showed that if the strong force between two particles were due to a tiny, extremely thin, almost rubber-band-like strand that connected the particles, then the quantum processes that Veneziano and others had been poring over would be mathematically described using Euler's formula. The little elastic strands were christened *strings* and now, with the horse properly before the cart, string theory was officially born.

But hold the bubbly. To those involved in this research, it was gratifying to understand the physical origin of Veneziano's insight, since it suggested that physicists were on their way to unmasking the strong nuclear force. Yet the discovery was not greeted with universal enthusiasm; far from it. Very far. In fact, Susskind's paper was returned by the journal to which he submitted it with the comment that the work was of minimal interest, an evaluation Susskind recalls well: "I was stunned, I was knocked off my chair, I was depressed, so I went home and got drunk."[9] Eventually, his paper and the others that announced the string concept were all published, but it was not long before the theory suffered two more devastating setbacks. Close scrutiny of more refined data on the strong nuclear force, collected during the early 1970s, showed that the string approach failed to describe the newer results accurately. Moreover, a new proposal called *quantum chromodynamics*, which was firmly rooted in the traditional ingredients of particles and fields—no strings at all—*was* able to describe all the data convincingly. And so by 1974, string theory had been dealt a one-two knockout punch. Or so it seemed.

John Schwarz was one of the earliest string enthusiasts. He once told me that from the start, he had a gut feeling that the theory was deep and important. Schwarz spent a number of years analyzing its various mathematical aspects; among other things, this led to the discovery of *superstring theory*—as we shall see, an important refinement of the original string proposal. But with the rise of quantum chromodynamics and the failure of the string framework to describe the strong force, the justification for continuing to work on string theory began to run thin. Nevertheless, there was one particular mismatch between string theory and the strong nuclear force that kept nagging at Schwarz, and he found that he just couldn't let it go. The quantum mechanical equations of string theory predicted that a particular, rather unusual, particle should be copiously produced in the high-energy collisions taking place in atom smashers. The particle would have zero mass, like a photon, but string theory predicted it would have *spin-two*, meaning, roughly speaking, that it would spin twice as fast as a photon. None of the experiments had ever found such a particle, so this appeared to be among the erroneous predictions made by string theory.

Schwarz and his collaborator Joël Scherk puzzled over this case of a missing particle, until in a magnificent leap they made a connection to a completely different problem. Although no one had been able to combine general relativity and quantum mechanics, physicists had determined certain features that would emerge from any successful union. And, as indicated in Chapter 9, one feature they found was that just as the electromagnetic force is transmitted microscopically by photons, the gravitational force should be microscopically transmitted by another class of particles, gravitons (the most elementary, quantum bundles of gravity). Although gravitons have yet to be detected experimentally, the theoretical analyses all agreed that gravitons must have two properties: they must be massless and have spin-two. For Schwarz and Scherk this rang a loud bell—these were just the properties of the rogue particle predicted by string theory—and inspired them to make a bold move, one that would transform a failing of string theory into a striking success.

They proposed that string theory shouldn't be thought of as a quantum mechanical theory of the strong nuclear force. They argued that even though the theory had been discovered in an attempt to understand the strong force, it was actually the solution to a different problem. It was actually the first ever quantum mechanical theory of the *gravitational* force. They claimed that the massless spin-two particle predicted by string theory was the graviton, and that the equations of string theory necessarily embodied a quantum mechanical description of gravity.

Schwarz and Scherk published their proposal in 1974 and expected a major reaction from the physics community. Instead, their work was ignored. In retrospect, it's not hard to understand why. It seemed to some that the string concept had become a theory in search of an application.

After the attempt to use string theory to explain the strong nuclear force had failed, it seemed as though its proponents wouldn't accept defeat and, instead, were flat out determined to find relevance for the theory elsewhere. Fuel was added to this view's fire when it became clear that Schwarz and Scherk needed to change the size of strings in their theory radically so that the force transmitted by the candidate gravitons would have the familiar, known strength of gravity. Since gravity is an extremely weak force* and since, it turns out, the longer the string the *stronger* the force transmitted, Schwarz and Scherk found that strings needed to be extremely tiny to transmit a force with gravity's feeble strength; they needed to be about the Planck length in size, a hundred billion billion times smaller than previously envisioned. So small, doubters wryly noted, that there was no equipment that would be able to see them, which meant that the theory could not be tested experimentally.[10]

By contrast, the 1970s witnessed one success after another for the more conventional, non-string-based theories, formulated with point particles and fields. Theorists and experimenters alike had their heads and hands full of concrete ideas to investigate and predictions to test. Why turn to speculative string theory when there was so much exciting work to be done within a tried-and-true framework? In much the same vein, although physicists knew in the backs of their minds that the problem of merging gravity and quantum mechanics remained unsolved using conventional methods, it was not a problem that commanded attention. Almost everyone acknowledged that it was an important issue and would need to be addressed one day, but with the wealth of work still to be done on the nongravitational forces, the problem of quantizing gravity was pushed to a barely burning back burner. And, finally, in the mid to late 1970s, string theory was far from having been completely worked out. Containing a candidate for the graviton was a success, but many conceptual and technical issues had yet to be addressed. It seemed thoroughly plausible that the theory would be unable to surmount one or more of these issues, so working on string theory meant taking a considerable risk. Within a few years, the theory might be dead.

Schwarz remained resolute. He believed that the discovery of string theory, the first plausible approach for describing gravity in the language

*Remember, as noted in Chapter 9, even a puny magnet can overpower the pull of the entire earth's gravity and pick up a paper clip. Numerically, the gravitational force has about 10^{-42} times the strength of the electromagnetic force.

of quantum mechanics, was a major breakthrough. If no one wanted to listen, fine. He would press on and develop the theory, so that when people were ready to pay attention, string theory would be that much further along. His determination proved prescient.

In the late 1970s and early 1980s, Schwarz teamed up with Michael Green, then of Queen Mary College in London, and set to work on some of the technical hurdles facing string theory. Primary among these was the problem of *anomalies*. The details are not of the essence, but, roughly speaking, an anomaly is a pernicious quantum effect that spells doom for a theory by implying that it violates certain sacred principles, such as energy conservation. To be viable, a theory must be free of all anomalies. Initial investigations had revealed that string theory was plagued by anomalies, which was one of the main technical reasons it had failed to generate much enthusiasm. The anomalies signified that although string theory appeared to provide a quantum theory of gravity, since it contained gravitons, on closer inspection the theory suffered from its own subtle mathematical inconsistencies.

Schwarz realized, however, that the situation was not clear-cut. There was a chance—it was a long shot—that a complete calculation would reveal that the various quantum contributions to the anomalies afflicting string theory, when combined correctly, cancelled each other out. Together with Green, Schwarz undertook the arduous task of calculating these anomalies, and by the summer of 1984 the two hit pay dirt. One stormy night, while working late at the Aspen Center for Physics in Colorado, they completed one of the field's most important calculations—a calculation proving that all of the potential anomalies, in a way that seemed almost miraculous, *did* cancel each other out. String theory, they revealed, was free of anomalies and hence suffered from no mathematical inconsistencies. String theory, they demonstrated convincingly, was quantum mechanically viable.

This time physicists listened. It was the mid-1980s, and the climate in physics had shifted considerably. Many of the essential features of the three nongravitational forces had been worked out theoretically and confirmed experimentally. Although important details remained unresolved—and some still do—the community was ready to tackle the next major problem: the merging of general relativity and quantum mechanics. Then, out of a little-known corner of physics, Green and Schwarz burst on the scene with a definite, mathematically consistent, and aesthetically pleasing proposal for how to proceed. Almost overnight, the number of researchers working on string theory leaped from two to over a thousand. The first string revolution was under way.

The First Revolution

I began graduate school at Oxford University in the fall of 1984, and within a few months the corridors were abuzz with talk of a revolution in physics. As the Internet was yet to be widely used, rumor was a dominant channel for the rapid spread of information, and every day brought word of new breakthroughs. Researchers far and wide commented that the atmosphere was charged in a way unseen since the early days of quantum mechanics, and there was serious talk that the end of theoretical physics was within reach.

String theory was new to almost everyone, so in those early days its details were not common knowledge. We were particularly fortunate at Oxford: Michael Green had recently visited to lecture on string theory, so many of us became familiar with the theory's basic ideas and essential claims. And impressive claims they were. In a nutshell, here is what the theory said:

Take any piece of matter—a block of ice, a chunk of rock, a slab of iron—and imagine cutting it in half, then cutting one of the pieces in half again, and on and on; imagine continually cutting the material into ever smaller pieces. Some 2,500 years ago, the ancient Greeks had posed the problem of determining the finest, uncuttable, indivisible ingredient that would be the end product of such a procedure. In our age we have learned that sooner or later you come to atoms, but atoms are not the answer to the Greeks' question, because they can be cut into finer constituents. Atoms can be split. We have learned that they consist of electrons that swarm around a central nucleus that is composed of yet finer particles—protons and neutrons. And in the late 1960s, experiments at the Stanford Linear Accelerator revealed that even neutrons and protons themselves are made up of more fundamental constituents: each proton and each neutron consists of three particles known as quarks, as mentioned in Chapter 9 and illustrated in Figure 12.3a.

Conventional theory, supported by state-of-the-art experiments, envisions electrons and quarks as dots with no spatial extent whatsoever; in this view, therefore, they mark the end of the line—the last of nature's matryoshka dolls to be found in the microscopic makeup of matter. Here

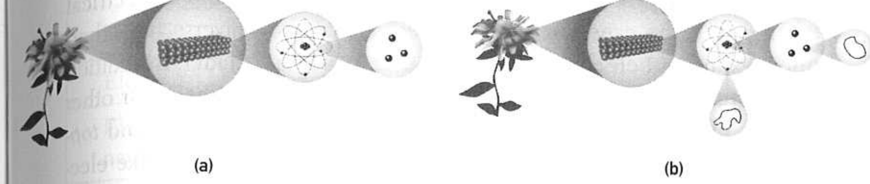


Figure 12.3 (**a**) Conventional theory is based on electrons and quarks as the basic constituents of matter. (**b**) String theory suggests that each particle is actually a vibrating string.

is where string theory makes its appearance. String theory challenges the conventional picture by proposing that electrons and quarks are *not* zero-sized particles. Instead, the conventional particle-as-dot model, according to string theory, is an approximation of a more refined portrayal in which each particle is actually a tiny, vibrating filament of energy, called a *string*, as you can see in Figure 12.3b. These strands of vibrating energy are envisioned to have no thickness, only length, and so strings are one-dimensional entities. Yet, because the strings are so small, some hundred billion billion times smaller than a single atomic nucleus (10^{-33} centimeters), they appear to be points even when examined with our most advanced atom smashers.

Because our understanding of string theory is far from complete, no one knows for sure whether the story ends here—whether, assuming the theory is correct, strings are truly the final Russian doll, or whether strings themselves might be composed of yet finer ingredients. We will come back to this issue, but for now we follow the historical development of the subject and imagine that strings are truly where the buck stops; we imagine that strings are *the* most elementary ingredient in the universe.

String Theory and Unification

That's string theory in brief, but to convey the power of this new approach, I need to describe conventional particle physics a little more fully. Over the past hundred years, physicists have prodded, pummeled, and pulverized matter in search of the universe's elementary constituents. And, indeed, they have found that in almost everything anyone has ever encountered, the fundamental ingredients are the electrons and quarks just mentioned—more precisely, as in Chapter 9, electrons and two kinds of quarks, up-quarks and down-quarks, that differ in mass and in electrical charge. But the experiments also revealed that the universe has other, more exotic particle species that don't arise in ordinary matter. In addition to up-quarks and down-quarks, experimenters have identified four other species of quarks (*charm-quarks, strange-quarks, bottom-quarks,* and *top-quarks*) and two other species of particles that are very much like electrons, only heavier (*muons* and *taus*). It is likely that these particles were plentiful just after the big bang, but today they are produced only as the ephemeral debris from high-energy collisions between the more familiar particle species. Finally, experimenters have also discovered three species of ghostly particles called *neutrinos* (*electron-neutrinos, muon-neutrinos,* and *tau-neutrinos*) that can pass through trillions of miles of lead as easily as we pass through air. These particles—the electron and its two heavier cousins, the six kinds of quarks, and the three kinds of neutrinos—constitute a modern-day particle physicist's answer to the ancient Greek question about the makeup of matter.[11]

The laundry list of particle species can be organized into three "families" or "generations" of particles, as in Table 12.1. Each family has two of the quarks, one of the neutrinos, and one of the electronlike particles; the only difference between corresponding particles in each family is that their masses increase in each successive family. The division into families certainly suggests an underlying pattern, but the barrage of particles can easily make your head spin (or, worse, make your eyes glaze over). Hang on, though, because one of the most beautiful features of string theory is that it provides a means for taming this apparent complexity.

According to string theory, there is only *one* fundamental ingredient—the string—and the wealth of particle species simply reflects the different vibrational patterns that a string can execute. It's just like what happens with more familiar strings like those on a violin or cello. A cello string can vibrate in many different ways, and we hear each pattern as a different musical note. In this way, one cello string can produce a range of different sounds. The strings in string theory behave similarly: they too can vibrate in different patterns. But instead of yielding different musical tones, *the different vibrational patterns in string theory correspond to different kinds of particles.* The key realization is that the detailed pattern of vibration executed by a string produces a specific mass, a specific electric charge, a specific spin, and so on—the specific list of properties, that is, which distinguish one kind of particle from another. A string vibrating in one particular pattern might have the properties of an electron, while a string

Family 1		Family 2		Family 3	
Particle	Mass	Particle	Mass	Particle	Mass
Electron	.00054	Muon	.11	Tau	1.9
Electron-neutrino	$<10^{-9}$	Muon-neutrino	$<10^{-4}$	Tau-neutrino	$<10^{-3}$
Up-quark	.0047	Charm-quark	1.6	Top-quark	189
Down-quark	.0074	Strange-quark	.16	Bottom-quark	5.2

Table 12.1 The three families of fundamental particles and their masses (in multiples of the proton mass). The values of the neutrino masses are known to be nonzero but their exact values have so far eluded experimental determination.

vibrating in a different pattern might have the properties of an up-quark, a down-quark, or any of the other particle species in Table 12.1. It is not that an "electron string" makes up an electron, or an "up-quark string" makes up an up-quark, or a "down-quark string" makes up a down-quark. Instead, the *single* species of string can account for a great variety of particles because the string can execute a great variety of vibrational patterns.

As you can see, this represents a potentially giant step toward unification. If string theory is correct, the head-spinning, eye-glazing list of particles in Table 12.1 manifests the vibrational repertoire of a single basic ingredient. Metaphorically, the different notes that can be played by a single species of string would account for all of the different particles that have been detected. At the ultramicroscopic level, the universe would be akin to a string symphony vibrating matter into existence.

This is a delightfully elegant framework for explaining the particles in Table 12.1, yet string theory's proposed unification goes even further. In Chapter 9 and in our discussion above, we discussed how the forces of nature are transmitted at the quantum level by other particles, the messenger particles, which are summarized in Table 12.2. String theory accounts for the messenger particles exactly as it accounts for the matter particles. Namely, each messenger particle is a string that's executing a particular vibrational pattern. A photon is a string vibrating in one particular pattern, a W particle is a string vibrating in a different pattern, a gluon is a string vibrating in yet another pattern. And, of prime importance, what Schwarz and Scherk showed in 1974 is that there is a parti-

Force	Force particle	Mass
Strong	Gluon	0
Electromagnetic	Photon	0
Weak	W, Z	86, 97
Gravity	Graviton	0

Table 12.2 The four forces of nature, together with their associated force particles and their masses in multiples of the proton mass. (There are actually two W particles—one with charge +1 and one with charge −1—that have the same mass; for simplicity we ignore this detail and refer to each as a W particle.

cular vibrational pattern that has all the properties of a graviton, so that the gravitational force is included in string theory's quantum mechanical framework. Thus, not only do matter particles arise from vibrating strings, but so do the messenger particles—even the messenger particle for gravity.

And so, beyond providing the first successful approach for merging gravity and quantum mechanics, string theory revealed its capacity to provide a unified description of all matter and all forces. That's the claim that knocked thousands of physicists off their chairs in the mid-1980s; by the time they got up and dusted themselves off, many were converts.

Why Does String Theory Work?

Before the development of string theory, the path of scientific progress was strewn with unsuccessful attempts to merge gravity and quantum mechanics. So what is it about string theory that has allowed it to succeed thus far? We've described how Schwarz and Scherk realized, much to their surprise, that one particular string vibrational pattern had just the right properties to be the graviton particle, and how they then concluded that string theory provided a ready-made framework for merging the two theories. Historically, that is indeed how the power and promise of string theory was fortuitously realized, but as an explanation for why the string approach succeeded where all other attempts failed, it leaves us wanting. Figure 12.2 encapsulates the conflict between general relativity and quantum

mechanics—on ultrashort distance (and time) scales, the frenzy of quantum uncertainty becomes so violent that the smooth geometrical model of spacetime underlying general relativity is destroyed—so the question is, How does string theory solve the problem? How does string theory calm the tumultuous fluctuations of spacetime at ultramicroscopic distances?

The main new feature of string theory is that its basic ingredient is not a point particle—a dot of no size—but instead is an object that has spatial extent. This difference is the key to string theory's success in merging gravity and quantum mechanics.

The wild frenzy depicted in Figure 12.2 arises from applying the uncertainty principle to the gravitational field; on smaller and smaller scales, the uncertainty principle implies that fluctuations in the gravitational field get larger and larger. On such extremely tiny distance scales, though, we should describe the gravitational field in terms of its fundamental constituents, gravitons, much as on molecular scales we should describe water in terms of H_2O molecules. In this language, the frenzied gravitational field undulations should be thought of as large numbers of gravitons wildly flitting this way and that, like bits of dirt and dust caught up in a ferocious tornado. Now, if gravitons were point particles (as envisioned in all earlier, failed attempts to merge general relativity and quantum mechanics), Figure 12.2 would accurately reflect their collective behavior: ever shorter distance scales, ever greater agitation. But string theory changes this conclusion.

In string theory, each graviton is a vibrating string—something that is not a point, but instead is roughly a Planck length (10^{-33} centimeters) in size.[12] And since the gravitons are the finest, most elementary constituents of a gravitational field, it makes no sense to talk about the behavior of gravitational fields on sub–Planck length scales. Just as resolution on your TV screen is limited by the size of individual pixels, resolution of the gravitational field in string theory is limited by the size of gravitons. Thus, the nonzero size of gravitons (and everything else) in string theory sets a limit, at roughly the Planck scale, to how finely a gravitational field can be resolved.

That is the vital realization. The uncontrollable quantum fluctuations illustrated in Figure 12.2 arise only when we consider quantum uncertainty on arbitrarily short distance scales—scales shorter than the Planck length. In a theory based on zero-sized point particles, such an application of the uncertainty principle is warranted and, as we see in the figure, this leads us to a wild terrain beyond the reach of Einstein's general relativity. A theory based on strings, however, includes a built-in fail-safe. In string theory, strings are the smallest ingredient, so our journey into the ultramicroscopic comes to an end when we reach the Planck length—the size of strings themselves. In Figure 12.2, the Planck scale is represented by the second highest level; as you can see, on such scales there are still undulations in the spatial fabric because the gravitational field is still subject to quantum jitters. But the jitters are mild enough to avoid irreparable conflict with general relativity. The precise mathematics underlying general relativity must be modified to incorporate these quantum undulations, but this can be done and the math remains sensible.

Thus, by limiting how small you can get, string theory limits how violent the jitters of the gravitational field become—and the limit is just big enough to avoid the catastrophic clash between quantum mechanics and general relativity. In this way, string theory quells the antagonism between the two frameworks and is able, for the first time, to join them.

Cosmic Fabric in the Realm of the Small

What does this mean for the ultramicroscopic nature of space and spacetime more generally? For one thing, it forcefully challenges the conventional notion that the fabric of space and time is continuous—that you can always divide the distance between here and there or the duration between now and then in half and in half again, endlessly partitioning space and time into ever smaller units. Instead, when you get down to the Planck length (the length of a string) and Planck time (the time it would take light to travel the length of a string) and try to partition space and time more finely, you find you can't. The concept of "going smaller" ceases to have meaning once you reach the size of the *smallest* constituent of the cosmos. For zero-sized point particles this introduces no constraint, but since strings have size, it does. If string theory is correct, the usual concepts of space and time, the framework within which all of our daily experiences take place, simply don't apply on scales finer than the Planck scale—the scale of strings themselves.

As for what concepts take over, there is as yet no consensus. One possibility that jibes with the explanation above for how string theory meshes quantum mechanics and general relativity is that the fabric of space on the Planck scale resembles a lattice or a grid, with the "space" between

the grid lines being outside the bounds of physical reality. Just as a microscopic ant walking on an ordinary piece of fabric would have to leap from thread to thread, perhaps motion through space on ultramicroscopic scales similarly requires discrete leaps from one "strand" of space to another. Time, too, could have a grainy structure, with individual moments being packed closely together but not melding into a seamless continuum. In this way of thinking, the concepts of ever smaller space and time intervals would sharply come to an end at the Planck scale. Just as there is no such thing as an American coin value smaller than a penny, if ultramicroscopic spacetime has a grid structure, there would be no such thing as a distance shorter than the Planck length or a duration shorter than the Planck time.

Another possibility is that space and time do not abruptly cease to have meaning on extremely small scales, but instead gradually morph into other, more fundamental concepts. Shrinking smaller than the Planck scale would be off limits not because you run into a fundamental grid, but because the concepts of space and time segue into notions for which "shrinking smaller" is as meaningless as asking whether the number nine is happy. That is, we can envision that as familiar, macroscopic space and time gradually transform into their unfamiliar ultramicroscopic counterparts, many of their usual properties—such as length and duration—become irrelevant or meaningless. Just as you can sensibly study the temperature and viscosity of liquid water—concepts that apply to the macroscopic properties of a fluid—but when you get down to the scale of individual H_2O molecules, these concepts cease to be meaningful, so, perhaps, although you can divide regions of space and durations of time in half and in half again on everyday scales, as you pass the Planck scale they undergo a transformation that renders such division meaningless.

Many string theorists, including me, strongly suspect that something along these lines actually happens, but to go further we need to figure out the more fundamental concepts into which space and time transform.* To date, this is an unanswered question, but cutting-edge research (described in the final chapter) has suggested some possibilities with far-reaching implications.

*I might note that the proponents of another approach for merging general relativity and quantum mechanics, *loop quantum gravity*, to be briefly discussed in Chapter 16, take a viewpoint that is closer to the former conjecture—that spacetime has a discrete structure on the smallest of scales.

The Finer Points

With the description I've given so far, it might seem baffling that any physicist would resist the allure of string theory. Here, finally, is a theory that promises to realize Einstein's dream and more; a theory that could quell the hostility between quantum mechanics and general relativity; a theory with the capacity to unify all matter and all forces by describing everything in terms of vibrating strings; a theory that suggests an ultramicroscopic realm in which familiar space and time might be as quaint as a rotary telephone; a theory, in short, that promises to take our understanding of the universe to a whole new level. But bear in mind that no one has ever seen a string and, except for some maverick ideas discussed in the next chapter, it is likely that even if string theory is right, no one ever will. Strings are so small that a direct observation would be tantamount to reading the text on this page from a distance of 100 light-years: it would require resolving power nearly a billion billion times finer than our current technology allows. Some scientists argue vociferously that a theory so removed from direct empirical testing lies in the realm of philosophy or theology, but not physics.

I find this view shortsighted, or, at the very least, premature. While we may never have technology capable of seeing strings directly, the history of science is replete with theories that were tested experimentally through indirect means.[13] String theory isn't modest. Its goals and promises are big. And that's exciting and useful, because if a theory is to be *the* theory of our universe, it must match the real world not just in the broad-brush outline discussed so far, but also in minute detail. As we'll now discuss, therein lie potential tests.

During the 1960s and 1970s, particle physicists made great strides in understanding the quantum structure of matter and the nongravitational forces that govern its behavior. The framework to which they were finally led by experimental results and theoretical insights is called the *standard model* of particle physics and is based on quantum mechanics, the matter particles in Table 12.1, and the force particles in Table 12.2 (ignoring the graviton, since the standard model does not incorporate gravity, and including the Higgs particle, which is not listed in the tables), all viewed as point particles. The standard model is able to explain essentially all data produced by the world's atom smashers, and over the years its inventors have been deservedly lauded with the highest of honors. Even so, the standard model has significant limitations. We've already discussed how it, and every other approach prior to string theory, failed to merge gravity and quantum mechanics. But there are other shortcomings as well.

The standard model failed to explain *why* the forces are transmitted by the precise list of particles in Table 12.2 and *why* matter is composed of the precise list of particles in Table 12.1. Why are there three families of matter particles and why does each family have the particles it does? Why not two families or just one? Why does the electron have three times the electric charge of the down-quark? Why does the muon weigh 23.4 times as much as the up-quark, and why does the top-quark weigh about 350,000 times as much as an electron? Why is the universe constructed with this range of seemingly random numbers? The standard model takes the particles in Tables 12.1 and 12.2 (again, ignoring the graviton) as *input*, then makes impressively accurate predictions for how the particles will interact and influence each other. But the standard model can't explain the input—the particles and their properties—any more than your calculator can explain the numbers you input the last time you used it.

Puzzling over the properties of these particles is not an academic question of why this or that esoteric detail happens to be one way or another. Over the last century, scientists have realized that the universe has the familiar features of common experience only because the particles in Tables 12.1 and 12.2 have precisely the properties they do. Even fairly minor changes to the masses or electric charges of some of the particles would, for example, make them unable to engage in the nuclear processes that power stars. And without stars, the universe would be a completely different place. Thus, the detailed features of the elementary particles are entwined with what many view as the deepest question in all of science: *Why do the elementary particles have just the right properties to allow nuclear processes to happen, stars to light up, planets to form around stars, and on at least one such planet, life to exist?*

The standard model can't offer any insight into this question since the particle properties are part of its required input. The theory won't start to chug along and produce results until the particle properties are specified. But string theory is different. In string theory, particle properties are *determined* by string vibrational patterns and so the theory holds the promise of providing an explanation.

Particle Properties in String Theory

To understand string theory's new explanatory framework, we need to have a better feel for how string vibrations produce particle properties, so let's consider the simplest property of a particle, its mass.

From $E = mc^2$, we know that mass and energy are interchangeable; like dollars and euros, they are convertible currencies (but unlike monetary currencies, they have a fixed exchange rate, given by the speed of light times itself, c^2). Our survival depends on Einstein's equation, since the sun's life-sustaining heat and light are generated by the conversion of 4.3 million tons of matter into energy every second; one day, nuclear reactors on earth may emulate the sun by safely harnessing Einstein's equation to provide humanity with an essentially limitless supply of energy.

In these examples, energy is produced from mass. But Einstein's equation works perfectly well in reverse—the direction in which mass is produced from energy—and that's the direction in which string theory uses Einstein's equation. The *mass* of a particle in string theory is nothing but the *energy* of its vibrating string. For instance, the explanation string theory offers for why one particle is heavier than another is that the string constituting the heavier particle is vibrating faster and more furiously than the string constituting the lighter particle. Faster and more furious vibration means higher energy, and higher energy translates, via Einstein's equation, into greater mass. Conversely, the lighter a particle is, the slower and less frenetic is the corresponding string vibration; a massless particle like a photon or a graviton corresponds to a string executing the most placid and gentle vibrational pattern that it possibly can.*[14]

Other properties of a particle, such as its electric charge and its spin, are encoded through more subtle features of the string's vibrations. Compared with mass, these features are harder to describe nonmathematically, but they follow the same basic idea: the vibrational pattern is the particle's fingerprint; all the properties that we use to distinguish one particle from another are determined by the vibrational pattern of the particle's string.

In the early 1970s, when physicists analyzed the vibrational patterns arising in the first incarnation of string theory—the *bosonic string theory*—

*The relationship to mass arising from a Higgs ocean will be discussed later in the chapter.

to determine the kinds of particle properties the theory predicted, they hit a snag. Every vibrational pattern in the bosonic string theory had a whole-number amount of spin: spin-0, spin-1, spin-2, and so on. This was a problem, because although the messenger particles have spin values of this sort, particles of matter (like electrons and quarks) don't. They have a fractional amount of spin, spin-½. In 1971, Pierre Ramond of the University of Florida set out to remedy this deficiency; in short order, he found a way to modify the equations of the bosonic string theory to allow for half-integer vibrational patterns as well.

In fact, on closer inspection, Ramond's research, together with results found by Schwarz and his collaborator André Neveu and later insights of Ferdinando Gliozzi, Joël Scherk, and David Olive, revealed a perfect balance—a novel symmetry—between the vibrational patterns with different spins in the modified string theory. These researchers found that the new vibrational patterns arose in pairs whose spin values differed by half a unit. For every vibrational pattern with spin-½ there was an associated vibrational pattern with spin-0. For every vibrational pattern of spin-1, there was an associated vibrational pattern of spin-½, and so on. The relationship between integer and half-integer spin values was named *supersymmetry*, and with these results the *supersymmetric string theory*, or *superstring theory*, was born. Nearly a decade later, when Schwarz and Green showed that all the potential anomalies that threatened string theory canceled each other out, they were actually working in the framework of superstring theory, and so the revolution their paper ignited in 1984 is more appropriately called the first *superstring* revolution. (In what follows, we will often refer to strings and to string theory, but that's just a shorthand; we always mean superstrings and superstring theory.)

With this background, we can now state what it would mean for string theory to reach beyond broad-brush features and explain the universe in detail. It comes down to this: among the vibrational patterns that strings can execute, there must be patterns whose properties agree with those of the known particle species. The theory has vibrational patterns with spin-½, but it must have spin-½ vibrational patterns that match *precisely* the known matter particles, as summarized in Table 12.1. The theory has spin-1 vibrational patterns, but it must have spin-1 vibrational patterns that match *precisely* the known messenger particles, as summarized in Table 12.2. Finally, if experiments do indeed discover spin-0 particles, such as are predicted for Higgs fields, string theory must yield vibrational patterns that match *precisely* the properties of these particles as well. In short, for string theory to be viable, its vibrational patterns must yield and explain the particles of the standard model.

Here, then, is string theory's grand opportunity. If string theory is right, there *is* an explanation for the particle properties that experimenters have measured, and it's to be found in the resonant vibrational patterns that strings can execute. If the properties of these vibrational patterns match the particle properties in Tables 12.1 and 12.2, I think that would convince even the diehard skeptics of string theory's veracity, whether or not anyone had directly seen the extended structure of a string itself. And beyond establishing itself as the long-sought unified theory, with such a match between theory and experimental data, string theory would provide the first fundamental explanation for why the universe is the way it is.

So how does string theory fare on this critical test?

Too Many Vibrations

Well, at first blush, string theory fails. For starters, there are an infinite number of different string vibrational patterns, with the first few of an endless series schematically illustrated in Figure 12.4. Yet Tables 12.1 and 12.2 contain only a finite list of particles, and so from the get-go we appear to have a vast mismatch between string theory and the real world. What's more, when we analyze mathematically the possible energies—and hence masses—of these vibrational patterns, we come upon another significant mismatch between theory and observation. The masses of the permissible string vibrational patterns bear no resemblance to the experimentally measured particle masses recorded in Tables 12.1 and 12.2. It's not hard to see why.

Since the early days of string theory, researchers have realized that the stiffness of a string is inversely proportional to its length (its length squared, to be more precise): while long strings are easy to bend, the shorter the string the more rigid it becomes. In 1974, when Schwarz and Scherk proposed decreasing the size of strings so that they'd embody a gravitational force of the right strength, they therefore also proposed increasing the tension of the strings—all the way, it turns out, to about a thousand trillion trillion trillion (10^{39}) tons, about 100000000000000000000000000000000000000000 (10^{41}) times the tension on an average piano string. Now, if you imagine bending a tiny, extremely stiff string into one of the increasingly elaborate patterns in Figure 12.4, you'll realize that the more peaks and troughs there are, the more energy you'll have to exert. Conversely, once a string is vibrating in such an elaborate pattern, it embodies a huge amount of energy. Thus, all but the simplest string vibrational patterns are highly energetic and hence, via $E = mc^2$, correspond to particles with huge masses.

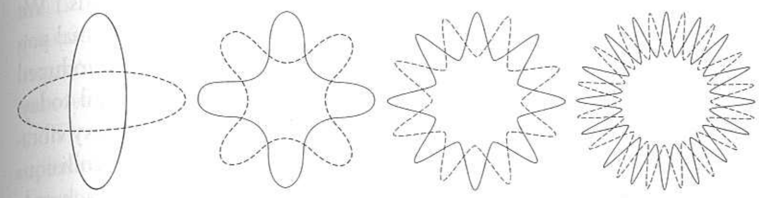


Figure 12.4 The first few examples of string vibrational patterns.

And by huge, I really mean huge. Calculations show that the masses of the string vibrations follow a series analogous to musical harmonics: they are all multiples of a fundamental mass, the *Planck mass*, much as overtones are all multiples of a fundamental frequency or tone. By the standards of particle physics, the Planck mass is colossal—it is some 10 billion billion (10^{19}) times the mass of a proton, roughly the mass of a dust mote or a bacterium. Thus, the possible masses of string vibrations are 0 times the Planck mass, 1 times the Planck mass, 2 times the Planck mass, 3 times the Planck mass, and so on, showing that the masses of all but the 0-mass string vibrations are gargantuan.[15]

As you can see, some of the particles in Tables 12.1 and 12.2 are indeed massless, but most aren't. And the nonzero masses in the tables are farther from the Planck mass than the Sultan of Brunei is from needing a loan. Thus, we see clearly that the known particle masses do not fit the pattern advanced by string theory. Does this mean that string theory is ruled out? You might think so, but it doesn't. Having an endless list of vibrational patterns whose masses become ever more remote from those of known particles is a challenge the theory must overcome. Years of research have revealed promising strategies for doing so.

As a start, note that experiments with the known particle species have taught us that heavy particles tend to be unstable; typically, heavy particles disintegrate quickly into a shower of lower-mass particles, ultimately generating the lightest and most familiar species in Tables 12.1 and 12.2.

(For instance, the top-quark disintegrates in about 10^{-24} seconds.) We expect this lesson to hold true for the "superheavy" string vibrational patterns, and that would explain why, even if they were copiously produced in the hot, early universe, few if any would have survived until today. Even if string theory is right, our only chance to see the superheavy vibrational patterns would be to produce them through high-energy collisions in particle accelerators. However, as current accelerators can reach only energies equivalent to roughly 1,000 times the mass of a proton, they are far too feeble to excite any but string theory's most placid vibrational patterns. Thus, string theory's prediction of a tower of particles with masses starting some million billion times greater than that achievable with today's technology is not in conflict with observations.

This explanation also makes clear that contact between string theory and particle physics will involve only the lowest-energy—the massless—string vibrations, since the others are way beyond what we can reach with today's technology. But what of the fact that most of the particles in Tables 12.1 and 12.2 are not massless? It's an important issue, but less troubling than it might at first appear. Since the Planck mass is huge, even the most massive particle known, the top-quark, weighs in at only .0000000000000000116 (about 10^{-17}) times the Planck mass. As for the electron, it weighs in at .0000000000000000000000034 (about 10^{-23}) times the Planck mass. So, to a first approximation—*valid to better than 1 part in 10^{17}*—all the particles in Tables 12.1 and 12.2 *do* have masses equal to zero times the Planck mass (much as most earthlings' wealth, to a first approximation, is 0 times that of the Sultan of Brunei), just as "predicted" by string theory. Our goal is to better this approximation and show that string theory explains the tiny deviations from 0 times the Planck mass characteristic of the particles in Tables 12.1 and 12.2. But massless vibrational patterns are not as grossly at odds with the data as you might have initially thought.

This is encouraging, but detailed scrutiny reveals yet further challenges. Using the equations of superstring theory, physicists have listed every massless string vibrational pattern. One entry is the spin-2 graviton, and that's the great success which launched the whole subject; it ensures that gravity is a part of quantum string theory. But the calculations also show that there are *many* more massless spin-1 vibrational patterns than there are particles in Table 12.2, and there are *many* more massless spin-½ vibrational patterns than there are particles in Table 12.1. Moreover, the list of spin-½ vibrational patterns shows no trace of any repetitive groupings like the family structure of Table 12.1. With a less cursory inspection, then, it seems increasingly difficult to see how string vibrations will align with the known particle species.

Thus, by the mid-1980s, while there were reasons to be excited about superstring theory, there were also reasons to be skeptical. Undeniably, superstring theory presented a bold step toward unification. By providing the first consistent approach for merging gravity and quantum mechanics, it did for physics what Roger Bannister did for the four-minute mile: it showed the seemingly impossible to be possible. Superstring theory established definitively that we could break through the seemingly impenetrable barrier separating the two pillars of twentieth-century physics.

Yet, in trying to go further and show that superstring theory could explain the detailed features of matter and nature's forces, physicists encountered difficulties. This led the skeptics to proclaim that superstring theory, despite all its potential for unification, was merely a mathematical structure with no direct relevance for the physical universe.

Even with the problems just discussed, at the top of the skeptics' list of superstring theory's shortcomings was a feature I've yet to introduce. Superstring theory does indeed provide a successful merger of gravity and quantum mechanics, one that is free of the mathematical inconsistencies that plagued all previous attempts. However, strange as it may sound, in the early years after its discovery, physicists found that the equations of superstring theory do *not* have these enviable properties if the universe has three spatial dimensions. Instead, the equations of superstring theory are mathematically consistent only if the universe has *nine* spatial dimensions, or, including the time dimension, they work only in a universe with *ten* spacetime dimensions!

In comparison to this bizarre-sounding claim, the difficulty in making a detailed alignment between string vibrational patterns and known particle species seems like a secondary issue. Superstring theory requires the existence of *six* dimensions of space that no one has ever seen. That's not a fine point—*that's* a problem.

Or is it?

Theoretical discoveries made during the early decades of the twentieth century, long before string theory came on the scene, suggested that extra dimensions need not be a problem at all. And, with a late-twentieth-century updating, physicists showed that these extra dimensions have the capacity to bridge the gap between string theory's vibrational patterns and the elementary particles experimenters have discovered.

This is one of the theory's most gratifying developments; let's see how it works.

Unification in Higher Dimensions

In 1919, Einstein received a paper that could easily have been dismissed as the ravings of a crank. It was written by a little-known German mathematician named Theodor Kaluza, and in a few brief pages it laid out an approach for unifying the two forces known at the time, gravity and electromagnetism. To achieve this goal, Kaluza proposed a radical departure from something so basic, so completely taken for granted, that it seemed beyond questioning. He proposed that the universe does not have three space dimensions. Instead, Kaluza asked Einstein and the rest of the physics community to entertain the possibility that the universe has *four* space dimensions so that, together with time, it has a total of five spacetime dimensions.

First off, what in the world does that mean? Well, when we say that there are three space dimensions we mean that there are three independent directions or axes along which you can move. From your current position you can delineate these as left/right, back/forth, and up/down; in a universe with three space dimensions, any motion you undertake is some combination of motion along these three directions. Equivalently, in a universe with three space dimensions you need precisely three pieces of information to specify a location. In a city, for example, you need a building's street, its cross street, and a floor number to specify the whereabouts of a dinner party. And if you want people to show up while the food is still hot, you also need to specify a fourth piece of data: a time. That's what we mean by spacetime's being four-dimensional.

Kaluza proposed that in addition to left/right, back/forth, and up/down, *the universe actually has one more spatial dimension that, for some reason, no one has ever seen.* If correct, this would mean that there is another independent direction in which things can move, and therefore that we need to give four pieces of information to specify a precise location in space, and a total of five pieces of information if we also specify a time.

Okay; that's what the paper Einstein received in April 1919 proposed.

The question is, Why didn't Einstein throw it away? We don't see another space dimension—we never find ourselves wandering aimlessly because a street, a cross street, and a floor number are somehow insufficient to specify an address—so why contemplate such a bizarre idea? Well, here's why. Kaluza realized that the equations of Einstein's general theory of relativity could fairly easily be extended mathematically to a universe that had one more space dimension. Kaluza undertook this extension and found, naturally enough, that the higher-dimensional version of general relativity not only included Einstein's original gravity equations but, because of the extra space dimension, also had extra equations. When Kaluza studied these extra equations, he discovered something extraordinary: the extra equations were none other than the equations Maxwell had discovered in the nineteenth century for describing the electromagnetic field! By imagining a universe with one new space dimension, Kaluza had proposed a solution to what Einstein viewed as one of the most important problems in all of physics. *Kaluza had found a framework that combined Einstein's original equations of general relativity with those of Maxwell's equations of electromagnetism.* That's why Einstein didn't throw Kaluza's paper away.

Intuitively, you can think of Kaluza's proposal like this. In general relativity, Einstein awakened space and time. As they flexed and stretched, Einstein realized that he'd found the geometrical embodiment of the gravitational force. Kaluza's paper suggested that the geometrical reach of space and time was greater still. Whereas Einstein realized that gravitational fields can be described as warps and ripples in the usual three space and one time dimensions, Kaluza realized that in a universe with an additional space dimension there would be additional warps and ripples. And those warps and ripples, his analysis showed, would be just right to describe electromagnetic fields. In Kaluza's hands, Einstein's own geometrical approach to the universe proved powerful enough to unite gravity and electromagnetism.

Of course, there was still a problem. Although the mathematics worked, there was—and still is—no evidence of a spatial dimension beyond the three we all know about. So was Kaluza's discovery a mere curiosity, or was it somehow relevant to our universe? Kaluza had a powerful trust in theory—he had, for example, learned to swim by studying a treatise on swimming and then diving into the sea—but the idea of an invisible space dimension, no matter how compelling the theory, still sounded outrageous. Then, in 1926, the Swedish physicist Oskar Klein injected a new twist into Kaluza's idea, one that suggested where the extra dimension might be hiding.

The Hidden Dimensions

To understand Klein's idea, picture Philippe Petit walking on a long, rubber-coated tightrope stretched between Mount Everest and Lhotse. Viewed from a distance of many miles, as in Figure 12.5, the tightrope appears to be a one-dimensional object like a line—an object that has extension only along its length. If we were told that a tiny worm was slithering along the tightrope in front of Philippe, we'd wildly cheer it on because it needs to stay ahead of Philippe's step to avoid disaster. Of course, with a moment's reflection we all realize that there is more to the surface of the tightrope than the left/right dimension we can directly perceive. Although difficult to see with the naked eye from a great distance, the surface of the tightrope has a second dimension: the clockwise/counterclockwise dimension that is "wrapped" around it. With the aid of a modest telescope, this circular dimension becomes visible and we see that the worm can move not only in the long, unfurled left/right direction but also in the short, "curled-up" clockwise/counterclockwise direction. That is, at every point on the tightrope, the worm has two independent directions in which it can move (that's what we mean when we say the tightrope's surface is two-dimensional*), so it can safely stay out of Philippe's way either by slithering ahead of him, as we initially envisioned, or by crawling around the tiny circular dimension and letting Philippe pass above.

The tightrope illustrates that dimensions—the independent directions in which anything can move—come in two qualitatively distinct varieties. They can be big and easy to see, like the left/right dimension of the tightrope's surface, or they can be tiny and more difficult to see, like the clockwise/counterclockwise dimension that circles around the tightrope's surface. In this example, it was not a major challenge to see the small circular girth of the tightrope's surface. All we needed was a reasonable magnifying instrument. But as you can imagine, the smaller a

*Were you to count left, right, clockwise, and counterclockwise all separately, you'd conclude that the worm can move in four directions. But when we speak of "independent" directions, we always group those that lie along the same geometrical axis—like left and right, and also clockwise and counterclockwise.

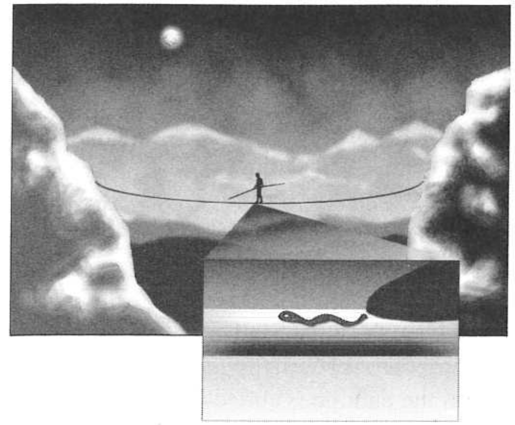


Figure 12.5 From a distance, a tightrope wire looks one-dimensional, although with a strong enough telescope, its second, curled-up dimension becomes visible.

curled-up dimension, the more difficult it is to detect. At a distance of a few miles, it's one thing to reveal the circular dimension of a tightrope's surface; it would be quite another to reveal the circular dimension of something as thin as dental floss or a narrow nerve fiber.

Klein's contribution was to suggest that what's true for an object *within* the universe might be true for the fabric of the universe itself. Namely, just as the tightrope's surface has both large and small dimensions, so does the fabric of space. Maybe the three dimensions we all know about—left/right, back/forth, and up/down—are like the horizontal extent of the tightrope, dimensions of the big, easy-to-see variety. But just as the surface of the tightrope has an additional, small, curled-up, circular dimension, maybe the fabric of space also has a small, curled-up, circular dimension, one so small that no one has powerful enough magnifying equipment to reveal its existence. Because of its tiny size, Klein argued, the dimension would be hidden.

How small is small? Well, by incorporating certain features of quantum mechanics into Kaluza's original proposal, Klein's mathematical analysis revealed that the radius of an extra circular spatial dimension would likely be roughly the Planck length,[16] certainly way too small for experimental accessibility (current state-of-the-art equipment cannot resolve anything smaller than about a thousandth the size of an atomic nucleus, falling short of the Planck length by more than a factor of a million billion). Yet, to an imaginary, Planck-sized worm, this tiny, curled-up circular dimension would provide a new direction in which it could roam just as freely as an ordinary worm negotiates the circular dimension of the tightrope in Figure 12.5. Of course, just as an ordinary worm finds that there isn't much room to explore in the clockwise direction before it finds itself back at its starting point, a Planck-sized worm slithering along a curled-up dimension of space would also repeatedly circle back to its starting point. But aside from the length of the travel it permitted, a curled-up dimension would provide a direction in which the tiny worm could move just as easily as it does in the three familiar unfurled dimensions.

To get an intuitive sense of what this looks like, notice that what we've been referring to as the tightrope's curled-up dimension—the clockwise/counterclockwise direction—*exists at each point along its extended dimension.* The earthworm can slither around the circular girth of the tightrope at any point along its outstretched length, and so the tightrope's surface can be described as having one long dimension, with a tiny, circular direction tacked on at each point, as in Figure 12.6. This is a useful image to have in mind because it also applies to Klein's proposal for hiding Kaluza's extra dimension of space.

To see this, let's again examine the fabric of space by sequentially showing its structure on ever smaller distance scales, as in Figure 12.7. At the first few levels of magnification, nothing new is revealed: the fabric of space still appears three-dimensional (which, as usual, we schematically represent on the printed page by a two-dimensional grid). However, when we get down to the Planck scale, the highest level of magnification in the figure, Klein suggested that a new, curled-up dimension becomes visible.

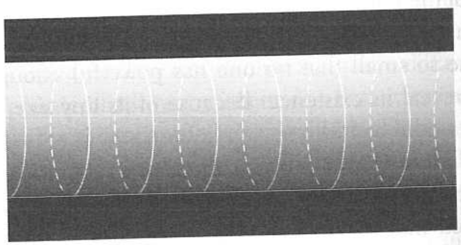


Figure 12.6 The surface of a tightrope has one long dimension with a circular dimension tacked on at each point.

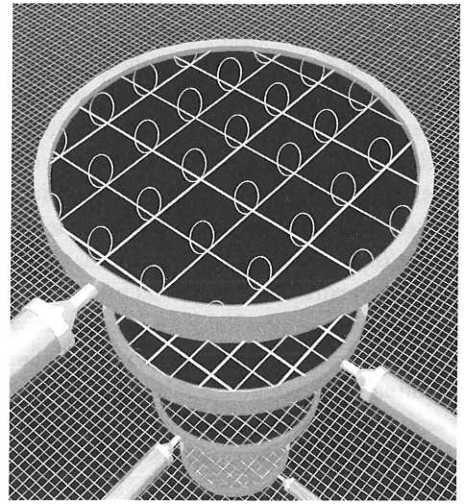


Figure 12.7 The Kaluza-Klein proposal is that on very small scales, space has an extra circular dimension tacked on to each familiar point.

Just as the circular dimension of the tightrope exists at each point along its big, extended dimension, the circular dimension in this proposal exists at each point in the familiar three extended dimensions of daily life. In Figure 12.7, we illustrate this by drawing the additional circular dimension at various points along the extended dimensions (since drawing the circle at every point would obscure the image) and you can immediately see the similarity with the tightrope in Figure 12.6. In Klein's proposal, therefore, space should be envisioned as having three unfurled dimensions (of which we show only two in the figure) with an additional circular dimension tacked on to each point. Notice that the extra dimension is not a bump or a loop within the usual three spatial dimensions, as the graphic limitations of the figure might lead you to think. Instead, the extra dimension is a new direction, completely distinct from the three we know about, which exists at every point in our ordinary three-dimensional space, but is so small that it escapes detection even with our most sophisticated instruments.

With this modification to Kaluza's original idea, Klein provided an answer to how the universe might have more than the three space dimensions of common experience that could remain hidden, a framework that has since become known as *Kaluza-Klein theory*. And since an extra dimension of space was all Kaluza needed to merge general relativity and electromagnetism, Kaluza-Klein theory would seem to be just what Einstein was looking for. Indeed, Einstein and many others became quite excited about unification through a new, hidden space dimension, and a vigorous effort was launched to see whether this approach would work in complete detail. But it was not long before Kaluza-Klein theory encountered its own problems. Perhaps most glaring of all, attempts to incorporate the electron into the extra-dimensional picture proved unworkable.[17] Einstein continued to dabble in the Kaluza-Klein framework until at least the early 1940s, but the initial promise of the approach failed to materialize, and interest gradually died out.

Within a few decades, though, Kaluza-Klein theory would make a spectacular comeback.

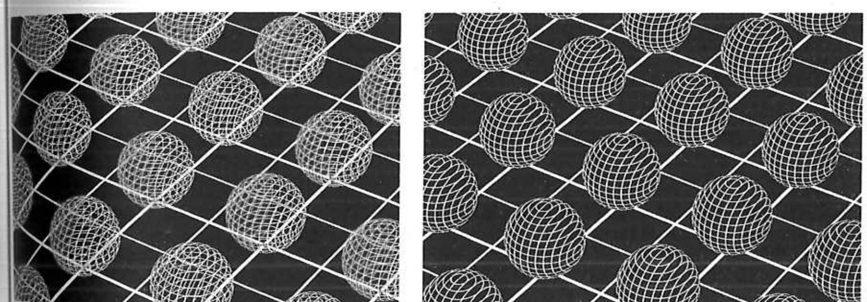
String Theory and Hidden Dimensions

In addition to the difficulties Kaluza-Klein theory encountered in trying to describe the microworld, there was another reason scientists were hesitant about the approach. Many found it both arbitrary and extravagant to postulate a hidden spatial dimension. It is not as though Kaluza was led to the idea of a new spatial dimension by a rigid chain of deductive reasoning. Instead, he pulled the idea out of a hat, and upon analyzing its implications discovered an unexpected link between general relativity and electromagnetism. Thus, although it was a great discovery in its own right, it lacked a sense of inevitability. If you asked Kaluza and Klein *why* the universe had five spacetime dimensions rather than four, or six, or seven, or 7,000 for that matter, they wouldn't have had an answer much more convincing than "Why not?"

More than three decades later, the situation changed radically. String theory is the first approach to merge general relativity and quantum mechanics; moreover, it has the potential to unify our understanding of all forces and all matter. But the quantum mechanical equations of string theory don't work in four spacetime dimensions, nor in five, six, seven, or 7,000. Instead, for reasons discussed in the next section, the equations of string theory work only in ten spacetime dimensions—nine of space, plus time. String theory *demands* more dimensions.

This is a fundamentally different kind of result, one never before encountered in the history of physics. Prior to strings, no theory said anything at all about the number of spatial dimensions in the universe. Every theory from Newton to Maxwell to Einstein assumed that the universe had three space dimensions, much as we all assume the sun will rise tomorrow. Kaluza and Klein proffered a challenge by suggesting that there were four space dimensions, but this amounted to yet another assumption—a different assumption, but an assumption nonetheless. Now, for the first time, string theory provided equations that *predicted* the number of space dimensions. A calculation—not an assumption, not a hypothesis, not an inspired guess—determines the number of space dimensions according to string theory, and the surprising thing is that the calculated number is not three, but nine. String theory leads us, *inevitably*, to a universe with six extra space dimensions and hence provides a compelling, ready-made context for invoking the ideas of Kaluza and Klein.

The original proposal of Kaluza and Klein assumed only one hidden dimension, but it's easily generalized to two, three, or even the six extra dimensions required by string theory. For example, in Figure 12.8a we replace the additional circular dimension of Figure 12.7, a one-dimensional shape, with the surface of a sphere, a two-dimensional shape (recall from the discussion in Chapter 8 that the surface of a sphere is two-dimensional because you need two pieces of information—like latitude



(a) (b)

Figure 12.8 A close-up of a universe with the three usual dimensions, represented by the grid, and **(a)** two curled-up dimensions, in the form of hollow spheres, and **(b)** three curled-up dimensions in the form of solid balls.

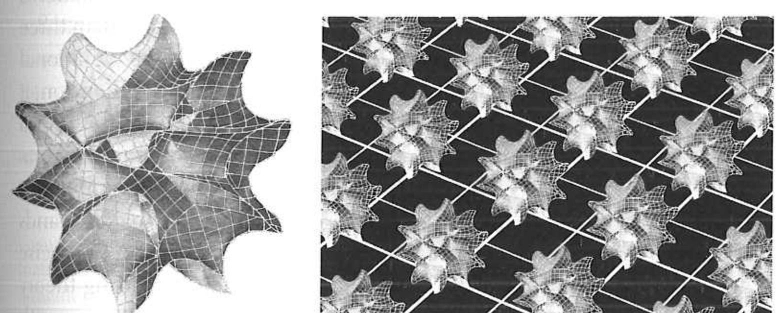
and longitude on the earth's surface—to specify a location). As with the circle, you should envision the sphere tacked on to every point of the usual dimensions, even though in Figure 12.8a, to keep the image clear, we draw only those that lie on the intersections of grid lines. In a universe of this sort, you would need a total of five pieces of information to locate a position in space: three pieces to locate your position in the big dimensions (street, cross street, floor number) and two pieces to locate your position on the sphere (latitude, longitude) tacked on at that point. Certainly, if the sphere's radius were tiny—billions of times smaller than an atom—the last two pieces of information wouldn't matter much for comparatively large beings like ourselves. Nevertheless, the extra dimension would be an integral part of the ultramicroscopic makeup of the spatial fabric. An ultramicroscopic worm would need all five pieces of information and, if we include time, it would need six pieces of information in order to show up at the right dinner party at the right time.

Let's go one dimension further. In Figure 12.8a, we considered only the surface of the spheres. Imagine now that, as in Figure 12.8b, the fabric of space also includes the interior of the spheres—our little Planck-sized worm can burrow into the sphere, as ordinary worms do with apples, and freely move throughout its interior. To specify the worm's location would now require *six* pieces of information: three to locate its position in the usual extended spatial dimensions, and three more to locate its position in the ball tacked on to that point (latitude, longitude, depth of penetration). Together with time, this is therefore an example of a universe with *seven* spacetime dimensions.

Now comes a leap. Although it is impossible to draw, imagine that at every point in the three extended dimensions of everyday life, the universe has not one extra dimension as in Figure 12.7, not two extra dimensions as in Figure 12.8a, not three extra dimensions as in Figure 12.8b, but six extra space dimensions. I certainly can't visualize this and I've never met anyone who can. But its meaning is clear. To specify the spatial location of a Planck-sized worm in such a universe requires *nine* pieces of information: three to locate its position in the usual extended dimensions and six more to locate its position in the curled-up dimensions tacked on to that point. When time is also taken into account, this is a ten-spacetime-dimensional universe, as required by the equations of string theory. If the extra six dimensions are curled up small enough, they would easily have escaped detection.

The Shape of Hidden Dimensions

The equations of string theory actually determine more than just the number of spatial dimensions. They also determine the kinds of shapes the extra dimensions can assume.[18] In the figures above, we focused on the simplest of shapes—circles, hollow spheres, solid balls—but the equations of string theory pick out a significantly more complicated class of six-dimensional shapes known as Calabi-Yau shapes or Calabi-Yau spaces. These shapes are named after two mathematicians, Eugenio Calabi and Shing-Tung Yau, who discovered them mathematically long before their relevance to string theory was realized; a rough illustration of one example is given in Figure 12.9a. Bear in mind that in this figure a two-dimensional graphic illustrates a six-dimensional object, and this results in a variety of significant distortions. Even so, the picture gives a rough sense of what these shapes look like. If the particular Calabi-Yau shape in Figure 12.9a constituted the extra six dimensions in string theory, on ultramicroscopic scales space would have the form illustrated in Figure 12.9b. As the Calabi-Yau shape would be tacked on to every point in the usual three dimensions, you and I and everyone else would right now be surrounded by and filled with these little shapes. Literally, as you walk



(a) (b)

Figure 12.9: (**a**) One example of a Calabi-Yau shape. (**b**) A highly magnified portion of space with additional dimensions in the form of a tiny Calabi-Yau shape.

from one place to another, your body would move through all nine dimensions, rapidly and repeatedly circumnavigating the entire shape, on average making it seem as if you weren't moving through the extra six dimensions at all.

If these ideas are right, the ultramicroscopic fabric of the cosmos is embroidered with the richest of textures.

String Physics and Extra Dimensions

The beauty of general relativity is that the physics of gravity is controlled by the geometry of space. With the extra spatial dimensions proposed by string theory, you'd naturally guess that the power of geometry to determine physics would substantially increase. And it does. Let's first see this by taking up a question that I've so far skirted. Why does string theory require ten spacetime dimensions? This is a tough question to answer nonmathematically, but let me explain enough to illustrate how it comes down to an interplay of geometry and physics.

Imagine a string that's constrained to vibrate only on the two-dimensional surface of a flat tabletop. The string will be able to execute a variety of vibrational patterns, but only those involving motion in the left/right and back/forth directions of the table's surface. If the string is then released to vibrate in the third dimension, motion in the up/down dimension that leaves the table's surface, additional vibrational patterns become accessible. Now, although it is hard to picture in more than three dimensions, this conclusion—more dimensions means more vibrational patterns—is general. If a string can vibrate in a fourth spatial dimension, it can execute more vibrational patterns than it could in only three; if a string can vibrate in a fifth spatial dimension, it can execute more vibrational patterns than it could in only four; and so on. This is an important realization, because there is an equation in string theory that demands that the number of independent vibrational patterns meet a very precise constraint. If the constraint is violated, the mathematics of string theory falls apart and its equations are rendered meaningless. In a universe with three space dimensions, the number of vibrational patterns is too small and the constraint is not met; with four space dimensions, the number of vibrational patterns is still too small; with five, six, seven, or eight dimensions it is still too small; but with nine space dimensions, the constraint on the number of vibrational patterns is satisfied perfectly. And that's how string theory determines the number of space dimensions.*[19]

While this illustrates well the interplay of geometry and physics, their association within string theory goes further and, in fact, provides a way to address a critical problem encountered earlier. Recall that, in trying to make detailed contact between string vibrational patterns and the known particle species, physicists ran into trouble. They found that there were far too many massless string vibrational patterns and, moreover, the detailed properties of the vibrational patterns did not match those of the known matter and force particles. But what I didn't mention earlier, because we hadn't yet discussed the idea of extra dimensions, is that although those calculations took account of the *number* of extra dimensions (explaining, in part, why so many string vibrational patterns were found), they did not take account of the small size and complex *shape* of the extra dimensions—they assumed that all space dimensions were flat and fully unfurled—and that makes a substantial difference.

Strings are so small that even when the extra six dimensions are crumpled up into a Calabi-Yau shape, the strings still vibrate into those directions. For two reasons, that's extremely important. First, it ensures that the strings always vibrate in all nine space dimensions, and hence the constraint on the number of vibrational patterns continues to be satisfied, even when the extra dimensions are tightly curled up. Second, just as the vibrational patterns of air streams blown through a tuba are affected by the twists and turns of the instrument, the vibrational patterns of strings are influenced by the twists and turns in the geometry of the extra six dimensions. If you were to change the shape of a tuba by making a passageway narrower or by making a chamber longer, the air's vibrational patterns and hence the sound of the instrument would change. Similarly, if the shape and size of the extra dimensions were modified, the precise properties of

*Let me prepare you for one relevant development we will encounter in the next chapter. String theorists have known for decades that the equations they generally use to mathematically analyze string theory are approximate (the exact equations have proven difficult to identify and understand). However, most thought that the approximate equations were sufficiently accurate to determine the required number of extra dimensions. More recently (and to the shock of most physicists in the field), some string theorists showed that the approximate equations *missed* one dimension; it is now accepted that the theory needs *seven* extra dimensions. As we will see, this does not compromise the material discussed in this chapter, but shows that it fits within a larger, in fact more unified, framework.[20]

each possible vibrational pattern of a string would also be significantly affected. And since a string's vibrational pattern determines its mass and charge, this means that the extra dimensions play a pivotal role in determining particle properties.

This is a key realization. *The precise size and shape of the extra dimensions has a profound impact on string vibrational patterns and hence on particle properties.* As the basic structure of the universe—from the formation of galaxies and stars to the existence of life as we know it—depends sensitively on the particle properties, the code of the cosmos may well be written in the geometry of a Calabi-Yau shape.

We saw one example of a Calabi-Yau shape in Figure 12.9, but there are at least hundreds of thousands of other possibilities. The question, then, is which Calabi-Yau shape, if any, constitutes the extra-dimensional part of the spacetime fabric. This is one of the most important questions string theory faces since only with a definite choice of Calabi-Yau shape are the detailed features of string vibrational patterns determined. To date, the question remains unanswered. The reason is that the current understanding of string theory's equations provides no insight into how to pick one shape from the many; from the point of view of the known equations, each Calabi-Yau shape is as valid as any other. The equations don't even determine the size of the extra dimensions. Since we don't see the extra dimensions, they must be small, but precisely how small remains an open question.

Is this a fatal flaw of string theory? Possibly. But I don't think so. As we will discuss more fully in the next chapter, the exact equations of string theory have eluded theorists for many years and so much work has used *approximate* equations. These have afforded insight into a great many features of string theory, but for certain questions—including the exact size and shape of the extra dimensions—the approximate equations fall short. As we continue to sharpen our mathematical analysis and improve these approximate equations, determining the form of the extra dimensions is a prime—and in my opinion attainable—objective. So far, this goal remains beyond reach.

Nevertheless, we can still ask whether *any* choice of Calabi-Yau shape yields string vibrational patterns that closely approximate the known particles. And here the answer is quite gratifying.

Although we are far from having investigated every possibility, examples of Calabi-Yau shapes have been found that give rise to string vibrational patterns in rough agreement with Tables 12.1 and 12.2. For instance, in the mid-1980s Philip Candelas, Gary Horowitz, Andrew Strominger, and Edward Witten (the team of physicists who realized the relevance of Calabi-Yau shapes for string theory) discovered that each hole—the term is used in a precisely defined mathematical sense—contained within a Calabi-Yau shape gives rise to a *family* of lowest-energy string vibrational patterns. A Calabi-Yau shape with three holes would therefore provide an explanation for the repetitive structure of three families of elementary particles in Table 12.1. Indeed, a number of such three-holed Calabi-Yau shapes have been found. Moreover, among these preferred Calabi-Yau shapes are ones that also yield just the right number of messenger particles as well as just the right electric charges and nuclear force properties to match the particles in Tables 12.1 and 12.2.

This is an extremely encouraging result; by no means was it ensured. In merging general relativity and quantum mechanics, string theory might have achieved one goal only to find it impossible to come anywhere near the equally important goal of explaining the properties of the known matter and force particles. Researchers take heart in the theory's having blazed past that disappointing possibility. Going further and calculating the precise masses of the particles is significantly more challenging. As we discussed, the particles in Tables 12.1 and 12.2 have masses that deviate from the lowest-energy string vibrations—zero times the Planck mass—by less than one part in a million billion. Calculating such infinitesimal deviations requires a level of precision way beyond what we can muster with our current understanding of string theory's equations.

As a matter of fact, I suspect, as do many other string theorists, that the tiny masses in Tables 12.1 and 12.2 arise in string theory much as they do in the standard model. Recall from Chapter 9 that in the standard model, a Higgs field takes on a nonzero value throughout all space, and the mass of a particle depends on how much drag force it experiences as it wades through the Higgs ocean. A similar scenario likely plays out in string theory. If a huge collection of strings all vibrate in just the right coordinated way throughout all of space, they can provide a uniform background that for all intents and purposes would be indistinguishable from a Higgs ocean. String vibrations that initially yielded zero mass would then acquire tiny nonzero masses through the drag force they experience as they move and vibrate through the string theory version of the Higgs ocean.

Notice, though, that in the standard model, the drag force experienced by a given particle—and hence the mass it acquires—is determined by experimental measurement and specified as an input to the theory. In the string theory version, the drag force—and hence the masses of the vibrational patterns—would be traced back to interactions between strings (since the Higgs ocean would be made of strings) and should be *calculable*. String theory, at least in principle, allows all particle properties to be determined by the theory itself.

No one has accomplished this, but as emphasized, string theory is still very much a work in progress. In time, researchers hope to realize fully the vast potential of this approach to unification. The motivation is strong because the potential payoff is big. With hard work and substantial luck, string theory may one day explain the fundamental particle properties and, in turn, explain why the universe is the way it is.

The Fabric of the Cosmos According to String Theory

Even though much about string theory still lies beyond the bounds of our comprehension, it has already exposed dramatic new vistas. Most strikingly, in mending the rift between general relativity and quantum mechanics, string theory has revealed that the fabric of the cosmos may have many more dimensions than we perceive directly—dimensions that may be the key to resolving some of the universe's deepest mysteries. Moreover, the theory intimates that the familiar notions of space and time do not extend into the sub-Planckian realm, which suggests that space and time as we currently understand them may be mere approximations to more fundamental concepts that still await our discovery.

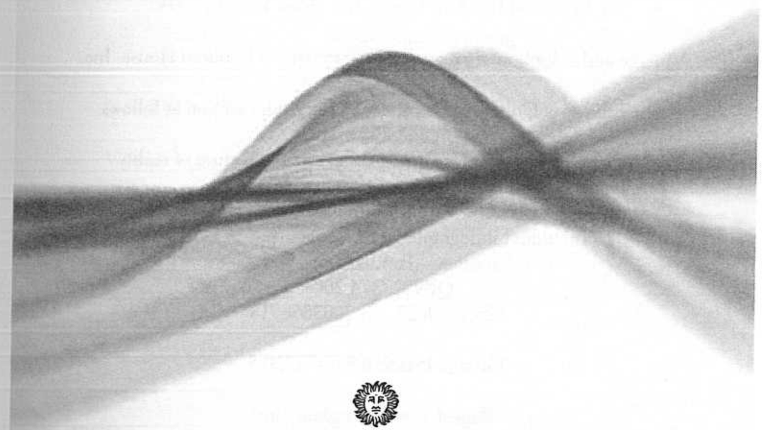
In the universe's initial moments, these features of the spacetime fabric that, today, can be accessed only mathematically, would have been manifest. Early on, when the three familiar spatial dimensions were also small, there would likely have been little or no distinction between what we now call the big and the curled-up dimensions of string theory. Their current size disparity would be due to cosmological evolution which, in a way that we don't yet understand, would have had to pick three of the spatial dimensions as special, and subject only them to the 14 billion years of expansion discussed in earlier chapters. Looking back in time even further, the entire observable universe would have shrunk into the sub-Planckian domain, so that what we've been referring to as the fuzzy patch (in Figure 10.6), we can now identify as the realm where familiar space and time have yet to emerge from the more fundamental entities—whatever they may be—that current research is struggling to comprehend.

Further progress in understanding the primordial universe, and hence in assessing the origin of space, time, and time's arrow, requires a significant honing of the theoretical tools we use to understand string theory—a goal that, not too long ago, seemed noble yet distant. As we'll now see, with the development of M-theory, progress has exceeded many of even the optimists' most optimistic predictions.

THE FABRIC OF THE COSMOS

SPACE, TIME, AND THE TEXTURE OF REALITY

BRIAN GREENE



VINTAGE BOOKS
A DIVISION OF RANDOM HOUSE, INC. • NEW YORK