

ALSO BY BRIAN GREENE

Icarus at the Edge of Time

The Fabric of the Cosmos

The Elegant Universe

THE HIDDEN REALITY

*Parallel Universes and the
Deep Laws of the Cosmos*

Brian Greene



ALFRED A. KNOPF NEW YORK 2011

CHAPTER 9

Black Holes and Holograms

The Holographic Multiverse

Plato likened our view of the world to that of an ancient forebear watching shadows meander across a dimly lit cave wall. He imagined our perceptions to be but a faint inkling of a far richer reality that flickers beyond reach. Two millennia later, it seems that Plato's cave may be more than a metaphor. To turn his suggestion on its head, reality—not its mere shadow—may take place on a distant boundary surface, while everything we witness in the three common spatial dimensions is a projection of that faraway unfolding. Reality, that is, may be akin to a hologram. Or, really, a holographic movie.

Arguably the strangest parallel world entrant, the *holographic principle* envisions that all we experience may be fully and equivalently described as the comings and goings that take place at a thin and remote locus. It says that if we could understand the laws that govern physics on that distant surface, and the way phenomena there link to experience here, we would grasp all there is to know about reality. A version of Plato's shadow world—a parallel but thoroughly unfamiliar encapsulation of everyday phenomena—would *be* reality.

The journey to this peculiar possibility combines developments deep and far flung—insights from general relativity; from research on black holes; from thermodynamics; quantum mechanics; and, most recently, string theory. The thread linking these diverse areas is the nature of information in a quantum universe.

Information

Beyond John Wheeler's knack for finding and mentoring the world's most gifted young scientists (besides Hugh Everett, Wheeler's students included Richard Feynman, Kip Thorne, and, as we will shortly see, Jacob Bekenstein), he had an uncanny ability to identify issues whose exploration could change our fundamental paradigm of nature's workings. During a lunch we had at Princeton in 1998, I asked him what he thought the dominant theme in physics would be in the decades going forward. As he had already done frequently that day, he put his head down, as if his aging frame had grown weary of supporting such a massive intellect. But now the length of his silence left me wondering, briefly, whether he didn't want to answer or whether, perhaps, he had forgotten the question. He then slowly looked up and said a single word: "Information."

I wasn't surprised. For some time, Wheeler had been advocating a view of physical law quite unlike what a fledgling physicist learns in the standard academic curriculum. Traditionally, physics focuses on *things*—planets, rocks, atoms, particles, fields—and investigates the forces that affect their behavior and govern their interactions. Wheeler was suggesting that *things*—matter and radiation—should be viewed as secondary, as carriers of a more abstract and fundamental entity: information. It's not that Wheeler was claiming that matter and radiation were somehow illusory; rather, he argued that they should be viewed as the material manifestations of something more basic. He believed that information—where a particle is, whether it is spinning one way or another, whether its charge is positive or negative, and so on—forms an irreducible kernel at the heart of reality. That such information is instantiated in real particles, occupying real positions, having definite spins and charges, is something like an architect's drawings being realized as a skyscraper. The fundamental information is in the blueprints. The skyscraper is but a physical realization of the information contained in the architect's design.

From this perspective, the universe can be thought of as an information processor. It takes information regarding how things are now and produces information delineating how things will be at the next now, and the now after that. Our senses become aware of such processing by detecting how the physical environment changes over time. But the physical environment itself is emergent; it arises from the fundamental ingredient, information, and evolves according to the fundamental rules, the laws of physics.

I don't know whether such an information-theoretic stance will reach the dominance in physics that Wheeler envisioned. But recently, driven largely by the work of physicists Gerard 't Hooft and Leonard Susskind, a major shift in thinking has resulted from puzzling questions regarding information in one particularly exotic context: black holes.

Black Holes

Within a year of general relativity's publication, the German astronomer Karl Schwarzschild found the first exact solution to Einstein's equations, a result that determined the shape of space and time in the vicinity of a massive spherical object such as a star or a planet. Remarkably, not only had Schwarzschild found his solution while calculating artillery trajectories on the Russian front during World War I, but also he had beaten the master at his own game: to that point, Einstein had found only approximate solutions to the equations of general relativity. Impressed, Einstein publicized Schwarzschild's achievement, presenting the work before the Prussian Academy, but even so he failed to appreciate a point that would become Schwarzschild's most tantalizing legacy.

Schwarzschild's solution shows that familiar bodies like the sun and the earth produce a modest curvature, a gentle depression in the otherwise flat spacetime trampoline. This matched well the approximate results Einstein had managed to work out earlier, but by dispensing with approximations, Schwarzschild could go further. His exact solution revealed something startling: if enough mass were crammed into a small enough ball, a gravitational abyss would form. The spacetime curvature would become so extreme that anything venturing too close would be trapped. And because "anything" includes light, such regions would fade to black, a characteristic that inspired the early term "dark stars." The extreme warping would also bring time to a grinding halt at the star's edge; hence another early label, "frozen stars." Half a century later, Wheeler, who was nearly as adept at marketing as he was at physics, popularized such stars both within and beyond the scientific community with a new and more memorable name: black holes. It stuck.

When Einstein read Schwarzschild's paper, he agreed with the mathematics as applied to ordinary stars or planets. But as to what we now call black holes? Einstein scoffed. In those early days it was a challenge, even for Einstein, to fully understand the intricate mathematics of general relativity. While the modern understanding of black holes was still decades

away, the intense folding of space and time already apparent in the equations was, in Einstein's view, too radical to be real. Much as he would resist cosmic expansion a few years later, Einstein refused to believe that such extreme configurations of matter were anything more than mathematical manipulations—based on his own equations—run amok.¹

When you see the numbers that are involved, it's easy to come to a similar conclusion. For a star as massive as the sun to be a black hole, it would need to be squeezed into a ball about three kilometers across; a body as massive as the earth would become a black hole only if squeezed to a centimeter across. The idea that there might be such extreme arrangements of matter seems nothing short of ludicrous. Yet, in the decades since, astronomers have gathered overwhelming observational evidence that black holes are both real and plentiful. There is wide agreement that a great many galaxies are powered by an enormous black hole at their center; our very own Milky Way galaxy is believed to revolve around a black hole whose mass is about three million times that of the sun. There's even a chance, as discussed in Chapter 4, that the Large Hadron Collider may produce tiny black holes in the laboratory by packing the mass (and energy) of violently colliding protons into such a minuscule volume that Schwarzschild's result again applies, though on microscopic scales. Extraordinary emblems of math's ability to illuminate the dark corners of the cosmos, black holes have become the cynosures of modern physics.

Besides serving as a boon for observational astronomy, black holes have also been a fertile source of inspiration for theoretical research by providing a mathematical playground in which physicists can push ideas to their limits, conducting pen-and-paper explorations of one of nature's most extreme environments. As a weighty case in point, in the early 1970s Wheeler realized that when the venerable Second Law of Thermodynamics—a guiding light for over a century in understanding the interplay between energy, work, and heat—was considered in the vicinity of a black hole, it seemed to flounder. The fresh thinking of Wheeler's young graduate student Jacob Bekenstein came to the rescue, and in doing so planted the seeds of the holographic proposal.

The Second Law

The aphorism "less is more" takes many forms. "Let's have the executive summary." "Just the facts." "TMI." "You had me at hello." These idioms

are so common because every moment of every day we're bombarded with information. Thankfully, in most cases our senses pare down the details to those that really matter. If I'm out on the savanna and encounter a lion, I don't care about the motion of every photon reflecting off his body. Way TMI. I just want particular overall features of those photons, the very ones our eyes have evolved to sense and our brains to rapidly decode. Is the lion coming toward me? Is he crouched and stalking? Provide me with a moment-to-moment catalog of every reflected photon and, sure, I'll be in possession of all the details. What I won't have is any understanding. Less would indeed be very much more.

Similar considerations play a central role in theoretical physics. Sometimes we want to know every microscopic detail of a system we're studying. At the locations along the Large Hadron Collider's seventeen-mile-long tunnel where particles are steered into head-on collisions, physicists have placed mammoth detectors capable of tracking, with extreme precision, the motion of the particle fragments produced. Essential for gaining insight into the fundamental laws of particle physics, the data are so detailed that a year's worth would fill a stack of DVDs about fifty times as tall as the Empire State Building. But, as in that impromptu meeting with a lion, there are other situations in physics where that level of detail would obscure, not clarify. A nineteenth-century branch of physics called *thermodynamics* or, in its more modern incarnation, *statistical mechanics*, focuses on such systems. The steam engine, the technological innovation that initially drove thermodynamics—as well as the Industrial Revolution—provides a good illustration.

The core of a steam engine is a vat of water vapor that expands when heated, driving the engine's piston forward, and contracts when cooled, returning the piston to its initial position, ready to drive forward once again. In the late nineteenth and early twentieth centuries, physicists worked out the molecular underpinnings of matter, which among other things provided a microscopic picture of the steam's action. As steam is heated, its H₂O molecules pick up increasing speed and career into the underside of the piston. The hotter they are, the faster they go and the bigger the push. A simple insight, but one essential to thermodynamics, is that to understand the steam's force we don't need the details of which particular molecules happen to have this or that velocity or which happen to hit the piston precisely here or there. Provide me with a list of billions and billions of molecular trajectories, and I'll look at you just as blankly as I would if you listed the photons bouncing off the lion. To figure out the piston's push, I need only the average number of molecules that will hit it

in a given time interval, and the average speed they'll have when they do. These are much coarser data, but it's exactly such pared-down information that's useful.

In crafting mathematical methods for systematically sacrificing detail in favor of such higher-level aggregate understanding, physicists honed a wide range of techniques and developed a number of powerful concepts. One such concept, encountered briefly in earlier chapters, is *entropy*. Initially introduced in the mid-nineteenth century to quantify energy dissipation in combustion engines, the modern view, emerging from Ludwig Boltzmann's work in the 1870s, is that entropy provides a characterization of how finely arranged—or not—the constituents of a given system need to be for it to have the overall appearance that it does.

To get a feel for this, imagine that Felix is frantic because he believes the apartment he shares with Oscar has been broken into. "They've ransacked us!" he tells Oscar. Oscar brushes him off—surely Felix is having one of his moments. To make his point, Oscar throws open the door to his bedroom, revealing clothing, empty pizza boxes, and crushed beer cans strewn everywhere. "It looks just like it always does," Oscar barks. Felix isn't swayed. "Of course it looks the same—ransack a pigsty and you get a pigsty. But look at my room." And he throws open his own door. "Ransacked," mocks Oscar; "it's neater than a straight whiskey." "Neat, yes. But the intruders have left their mark. My vitamin bottles? Not lined up in order of size. My collected works of Shakespeare? Out of alphabetical order. And my sock drawer? Look at this—some black pairs are in the blue bin! Ransacked, I tell you. Obviously ransacked."

Putting Felix's hysteria aside, the scenario makes plain a simple but essential point. When something is highly disordered, like Oscar's room, a great many possible rearrangements of its constituents leave its overall appearance intact. Grab the twenty-six crumpled shirts that were scattered across the bed, floor, and dresser, and toss them this way and that, fling the forty-two crushed beer cans randomly here and there, and the room will look the same. But when something is highly ordered, like Felix's room, even small rearrangements are easily detected.

This distinction underlies Boltzmann's mathematical definition of entropy. Take any system and count the number of ways its constituents can be rearranged without affecting its gross, overall, macroscopic appearance. That number is the system's entropy.* If there's a large number of such rearrangements, then entropy is high: the system is highly disor-

*This loose definition will suffice for now; in a moment, I'll be more precise.

dered. If the number of such rearrangements is small, entropy is low: the system is highly ordered (or, equivalently, has low disorder).

For more conventional examples, consider a vat of steam and a cube of ice. Focus only on their overall macroscopic properties, those you can measure or observe without accessing the detailed state of either's molecular constituents. When you wave your hand through the steam, you rearrange the positions of billions upon billions of H_2O molecules, and yet the vat's uniform haze looks undisturbed. But randomly change the positions and speeds of that many molecules in a piece of ice, and you'll immediately see the impact—the ice's crystalline structure will be disrupted. Fissures and fractures will appear. The steam, with H_2O molecules randomly flitting through the container, is highly disordered; the ice, with H_2O molecules arranged in a regular, crystalline pattern, is highly ordered. The entropy of the steam is high (many rearrangements will leave it looking the same); the entropy of the ice is low (few rearrangements will leave it looking the same).

By assessing the sensitivity of a system's macroscopic appearance to its microscopic details, entropy is a natural concept in a mathematical formalism that focuses on aggregate physical properties. The Second Law of Thermodynamics developed this line of insight quantitatively. The law states that, over time, the total entropy of a system will increase.² Understanding why requires only the most elementary grasp of chance and statistics. By definition, a higher-entropy configuration can be realized through many more microscopic arrangements than a lower-entropy configuration. As a system evolves, it's overwhelmingly likely to pass through higher-entropy states since, simply put, there are more of them. *Many* more. When bread is baking, you smell it throughout the house because there are trillions more arrangements of the molecules streaming from the bread that are spread out, yielding a uniform aroma, than there are arrangements in which the molecules are all tightly packed in a corner of the kitchen. The random motions of the hot molecules will, with near certainty, drive them toward one of the numerous spread-out arrangements, and not toward one of the few clustered configurations. The collection of molecules evolves, that is, from lower to higher entropy, and that's the Second Law in action.

The idea is general. Glass shattering, a candle burning, ink spilling, perfume pervading: these are different processes, but the statistical considerations are the same. In each, order degrades to disorder and does so because there are so many ways to be disordered. The beauty of this kind of analysis—the insight provided one of the most potent “Aha!” moments in my physics education—is that, without getting lost in the microscopic

details, we have a guiding principle to explain why a great many phenomena unfold the way they do.

Notice, too, that, being statistical, the Second Law does not say that entropy *can't* decrease, only that it is extremely unlikely to do so. The milk molecules you just poured into your coffee might, as a result of their random motions, coalesce into a floating figurine of Santa Claus. But don't hold your breath. A floating milk Santa has very low entropy. If you move around a few billion of his molecules, you'll notice the result—Santa will lose his head or an arm, or he'll disperse into abstract white tendrils. By comparison, a configuration in which the milk molecules are uniformly spread around has enormously more entropy: a vast number of rearrangements continue to look like ordinary coffee with milk. With a huge likelihood, then, the milk poured into your dark coffee will turn it a uniform tan, with nary a Santa in sight. Similar considerations hold for the vast majority of high-to-low-entropy evolutions, making the Second Law appear inviolable.

The Second Law and Black Holes

Now to Wheeler's point about black holes. Back in the early 1970s, Wheeler noticed that when black holes amble onto the scene, the Second Law appears compromised. A nearby black hole seems to provide a ready-made and reliable means for reducing overall entropy. Throw whatever system you're studying—smashed glass, burned candles, spilled ink—into the hole. Since nothing escapes from a black hole, the system's disorder would appear permanently gone. Crude the approach may be, but it seems easy to lower total entropy if you have a black hole to work with. The Second Law, many thought, had met its match.

Wheeler's student Bekenstein was not convinced. Perhaps, Bekenstein suggested, entropy is not lost to the black hole but merely transferred to it. After all, no one claimed that, in gorging themselves on dust and stars, black holes provide a mechanism for violating the First Law of Thermodynamics, the conservation of energy. Instead, Einstein's equations show that when a black hole gorges, it gets bigger and heftier. The energy in a region can be redistributed, with some falling into the hole and some remaining outside, but the total is preserved. Maybe, Bekenstein suggested, the same idea applies to entropy. Some entropy stays outside a given black hole and some entropy falls in, but none gets lost.

This sounds reasonable, but experts shot Bekenstein down. Schwarzschild's solution, and much work that followed, seemed to establish that

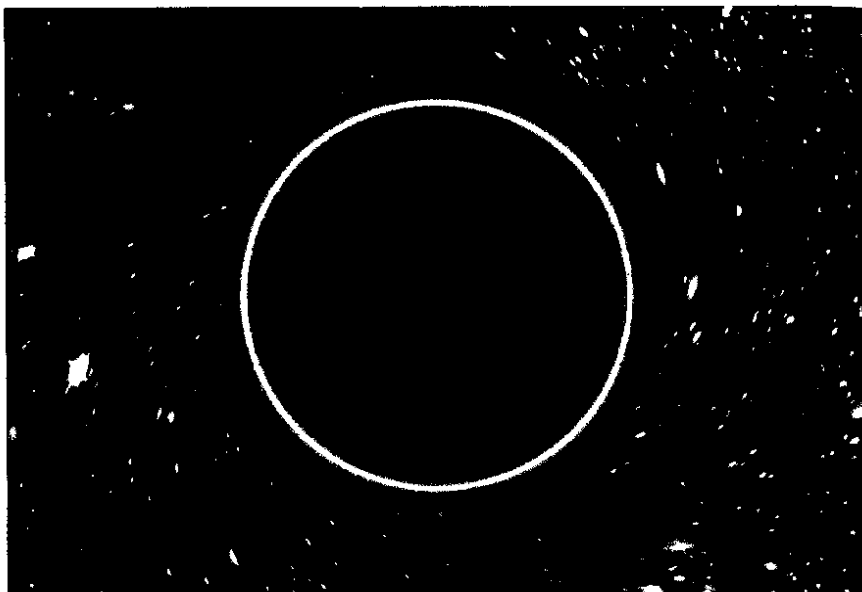


Figure 9.1 A black hole comprises a region of spacetime surrounded by a surface of no return, the event horizon.

black holes are the epitome of order. Infalling matter and radiation, however messy and disordered, are crushed to infinitesimal size at a black hole's center: a black hole is the ultimate in orderly trash compaction. True, no one knows exactly what happens during such powerful compression, because the extremes of curvature and density disrupt Einstein's equations; but there just doesn't seem to be any capacity for a black hole's center to harbor disorder. And outside the center, a black hole is nothing but an empty region of spacetime extending to the boundary of no return—the event horizon—as in Figure 9.1. With no atoms or molecules wafting this way and that, and thus no constituents to rearrange, a black hole would seem to be entropy-free.

In the 1970s, this view was reinforced by the so-called *no hair theorems*, which established mathematically that black holes, much like the bald performers of Blue Man Group, have a dearth of distinguishing characteristics. According to the theorems, any two black holes that have the same mass, charge, and angular momentum (rate of rotation) are *identical*. Lacking any other intrinsic traits—as the Blue Men lack bangs, mullets, or dreads—black holes seemed to lack the underlying differences that would harbor entropy.

By itself, this was a fairly convincing argument, but there was a yet

more damning consideration that seemed to definitively undercut Bekenstein's idea. According to basic thermodynamics, there's a close association between entropy and temperature. Temperature is a measure of the average motion of an object's constituents: hot objects have fast-moving constituents, cold objects have slow-moving constituents. Entropy is a measure of the possible rearrangements of these constituents that, from a macroscopic viewpoint, would go unnoticed. Both entropy and temperature thus depend on aggregate features of an object's constituents; they go hand in hand. When worked out mathematically, it became clear that if Bekenstein was right and black holes carried entropy, they should also have a temperature.³ That idea set off alarm bells. Any object with a nonzero temperature radiates. Hot coal radiates visible light; we humans, typically, radiate in the infrared. If a black hole has a nonzero temperature, the very laws of thermodynamics that Bekenstein was seeking to preserve state that it too should radiate. But that conflicts blatantly with the established understanding that nothing can escape a black hole's gravitational grip. Most everyone concluded that Bekenstein was wrong. Black holes do not have a temperature. Black holes do not harbor entropy. Black holes are entropy sinkholes. In their presence, the Second Law of Thermodynamics fails.

Despite the evidence mounting against him, Bekenstein had one tantalizing result on his side. In 1971, Stephen Hawking realized that black holes obey a curious law. If you have a collection of black holes with various masses and sizes, some engaged in stately orbital waltzes, others pulling in nearby matter and radiation, and still others crashing into each other, *the total surface area of the black holes increases over time*. By "surface area," Hawking meant the area of each black hole's event horizon. Now, there are many results in physics that ensure quantities don't change over time (conservation of energy, conservation of charge, conservation of momentum, and so on), but there are very few that require quantities to increase. It was natural, then, to consider a possible relation between Hawking's result and the Second Law. If we envision that, somehow, the surface area of a black hole is a measure of the entropy it contains, then the increase in total surface area could be read as an increase in total entropy.

It was an enticing analogy, but no one bought it. The similarity between Hawking's area theorem and the Second Law was, in almost everyone's view, nothing more than a coincidence. Until, that is, a few years later, when Hawking completed one of the most influential calculations in modern theoretical physics.

Hawking Radiation

Because quantum mechanics plays no role in Einstein's general relativity, Schwarzschild's black hole solution is based purely in classical physics. But proper treatment of matter and radiation—of particles like photons, neutrinos, and electrons that can carry mass, energy, and entropy from one location to another—requires quantum physics. To fully assess the nature of black holes and understand how they interact with matter and radiation, we must update Schwarzschild's work to include quantum considerations. This isn't easy. Notwithstanding advances in string theory (as well as in other approaches we haven't discussed, such as loop quantum gravity, twistors, and topos theory), we are still at an early stage in our attempt to meld quantum physics and general relativity. Back in the 1970s, there was still less theoretical basis for understanding how quantum mechanics would affect gravity.

Even so, a number of early researchers developed a partial union of quantum mechanics and general relativity by considering quantum fields (the quantum part) evolving in a fixed but curved spacetime environment (the general relativity part). As I pointed out in Chapter 4, a full union would, at the very least, consider not only the quantum jitters of fields within spacetime but the jitters of spacetime itself. To facilitate progress, the early work steadfastly avoided this complication. Hawking embraced the partial union and studied how quantum fields would behave in a very particular spacetime arena: that created by the presence of a black hole. What he found knocked physicists clear off their seats.

A well-known feature of quantum fields in ordinary, empty, uncurved spacetime is that their jitters allow pairs of particles, for instance an electron and its antiparticle the positron, to momentarily erupt out of the nothingness, live briefly, and then smash into each other, with mutual annihilation the result. This process, *quantum pair production*, has been intensively studied both theoretically and experimentally, and is thoroughly understood.

A novel characteristic of quantum pair production is that while one member of the pair has positive energy, the law of energy conservation dictates that the other must have an equal amount of *negative* energy—a concept that would be meaningless in a classical universe.* But the uncertainty principle provides a window of weirdness whereby negative-

*In Chapter 3, we discussed how the energy embodied by a gravitational field can be negative; this energy, however, is potential energy. The energy we're discussing here, kinetic energy, comes from the electron's mass and its motion. In classical physics this has to be positive.

energy particles are allowed as long as they don't overstay their welcome. If a particle exists only fleetingly, quantum uncertainty establishes that no experiment will have adequate time, even in principle, to determine the sign of its energy. This is the very reason why the particle pair is condemned by quantum laws to swift annihilation. So, over and over again, quantum jitters result in particle pairs being created and annihilated, created and annihilated, as the unavoidable rumbling of quantum uncertainty plays itself out in otherwise empty space.

Hawking reconsidered such ubiquitous quantum jitters not in the setting of empty space but near the event horizon of a black hole. He found that sometimes events look much as they ordinarily do. Pairs of particles are randomly created; they quickly find each other; they are destroyed. But every so often something new happens. If the particles are formed sufficiently close to the black hole's edge, one can get sucked in while the other careens into space. In the absence of a black hole this never happens, because if the particles failed to annihilate each other then the one with negative energy would outlive the protective haze of quantum uncertainty. Hawking realized that the black hole's radical twisting of space and time can cause particles that have negative energy, as determined by anyone outside the hole, to appear to have *positive* energy to any unfortunate observer inside the hole. In this way, a black hole provides the negative energy particles a safe haven, and so eliminates the need for a quantum cloak. The erupting particles can forgo mutual annihilation and blaze their own separate trails.⁴

The positive-energy particles shoot outward from just above the black hole's event horizon, so to someone watching from afar they look like radiation, a form since named *Hawking radiation*. The negative-energy particles are not directly seen, because they fall into the black hole, but they nevertheless have a detectable impact. Much as a black hole's mass increases when it absorbs anything that carries positive energy, so its mass decreases when it absorbs anything that carries negative energy. In tandem, these two processes make the black hole resemble a piece of burning coal: the black hole emits a steady outward stream of radiation as its mass gets ever smaller.⁵ When quantum considerations are included, black holes are thus not completely black. This was Hawking's bolt from the blue.

Which is not to say that your average black hole is red hot, either. As particles stream from just outside the black hole, they fight an uphill battle to escape the strong gravitational pull. In doing so, they expend energy and, because of this, cool down substantially. Hawking calculated that an observer far from the black hole would find that the temperature for the resulting "tired" radiation was inversely proportional to the black hole's

mass. A huge black hole, like the one at the center of our galaxy, has a temperature that's less than a trillionth of a degree above absolute zero. A black hole with the mass of the sun would have a temperature less than a millionth of a degree, minuscule even compared with the 2.7-degree cosmic background radiation left to us by the big bang. For a black hole's temperature to be high enough to barbecue the family dinner, its mass would need to be about a ten-thousandth of the earth's, extraordinarily small by astrophysical standards.

But the magnitude of a black hole's temperature is secondary. Although the radiation coming from distant astrophysical black holes won't light up the night sky, the fact that they *do* have a temperature, that they *do* emit radiation, suggests that the experts had too quickly rejected Bekenstein's suggestion that black holes *do* have entropy. Hawking then nailed the case. His theoretical calculations determining a given black hole's temperature and the radiation it emits gave him all the data he needed to determine the amount of entropy the black hole should contain, according to the standard laws of thermodynamics. And the answer he found is proportional to the surface area of the black hole, just as Bekenstein had proposed.

So by the end of 1974, the Second Law was law once again. The insights of Bekenstein and Hawking established that in any situation, total entropy increases, as long as you account for not only the entropy of ordinary matter and radiation but also that contained within black holes, as measured by their total surface area. Rather than being entropy sinks that subvert the Second Law, black holes play an active part in upholding the law's pronouncement of a universe with ever-increasing disorder.

The conclusion provided a welcome relief. To many physicists, the Second Law, emerging from seemingly unassailable statistical considerations, came as close to sacred as just about anything in science. Its restoration meant that, once again, all was right with the world. But, in time, a vital little detail in the entropy accounting made it clear that the Second Law's balance sheet was not the deepest issue in play. *That* honor went to identifying *where entropy is stored*, a matter whose importance becomes clear when we recognize the deep link between entropy and the central theme of this chapter: information.

Entropy and Hidden Information

So far, I've described entropy, loosely, as a measure of disorder and, more quantitatively, as the number of rearrangements of a system's microscopic

constituents that leave its overall macroscopic features unchanged. I've left implicit, but will now make explicit, that you can think of entropy as measuring the *gap in information* between the data you have (those overall macroscopic features) and the data you don't (the system's particular microscopic arrangement). Entropy measures the additional information hidden within the microscopic details of the system, which, should you have access to it, would distinguish the configuration at a micro level from all the macro look-alikes.

To illustrate, imagine that Oscar has straightened up his room, except that the thousand silver dollars he won in last week's poker game remain scattered across the floor. Even after he gathers them in a neat cluster, Oscar sees only a haphazard assortment of dollar coins, some heads and others tails. Were you to randomly change some heads to tails and other tails to heads, he'd never notice—evidence that the thousand-dropped-silver-dollar system has high entropy. Indeed, this example is so explicit that we can do the entropy counting. If there were only two coins, there'd be four possible configurations: (heads, heads), (heads, tails), (tails, heads), and (tails, tails)—two possibilities for the first dollar, times two for the second. With three coins, there'd be eight possible arrangements: (heads, heads, heads), (heads, heads, tails), (heads, tails, heads), (heads, tails, tails), (tails, heads, heads), (tails, heads, tails), (tails, tails, heads), (tails, tails, tails), arising from two possibilities for the first, times two for the second, times two for the third. With a thousand coins, the number of possibilities follows exactly the same pattern—a factor of 2 for each coin—yielding a total of 2^{1000} , which is 1071508607186267320948425 0490600018105614048117055336074437503883703510511249361224 9319837881569585812759467291755314682518714528569231404359 8457757469857480393456777482423098542107460506237114187795 4182153046474983581941267398767559165543946077062914571196 477686542167660429831652624386837205668069376. The vast majority of these heads-tails arrangements would have no distinguishing features, so they would not stand out in any way. Some *would*, for instance, if all 1,000 coins were heads or all were tails, or if 999 were heads, or 999 tails. But the number of such unusual configurations is so extraordinarily small, compared with the huge total number of possibilities, that removing them from the count would hardly make a difference.*

From our earlier discussion, you'd deduce that the number 2^{1000} is the entropy of the coins. And, for some purposes, that conclusion would be

*Besides flipping the coins, you could also swap around their locations, but for the purpose of illustrating the main ideas, we can safely ignore this complication.

fine. But to draw the strongest link between entropy and information, I need to sharpen up the description I gave earlier. The entropy of a system is *related to* the number of indistinguishable rearrangements of its constituents, but properly speaking is not equal to the number itself. The relationship is expressed by a mathematical operation called a *logarithm*; don't be put off if this brings back bad memories of high school math class. In our coin example, it simply means that you pick out the exponent in the number of rearrangements—that is, the entropy is defined as 1,000 rather than 2^{1000} .

Using logarithms has the advantage of allowing us to work with more manageable numbers, but there's a more important motivation. Imagine I ask you how much information you'd need to supply in order to describe one particular heads-tails arrangement of the 1,000 coins. The simplest response is that you'd need to provide the list—heads, heads, tails, heads, tails, tails . . .—that specifies the disposition of each of the 1,000 coins. Sure, I respond, that would tell me the details of the configuration, but that wasn't my question. I asked *how much information* is contained in that list.

So, you start to ponder. What actually *is* information, and what does it do? Your response is simple and direct. Information answers questions. Years of research by mathematicians, physicists, and computer scientists have made this precise. Their investigations have established that the most useful measure of information content is the *number of distinct yes-no questions the information can answer*. The coins' information answers 1,000 such questions: Is the first dollar heads? Yes. Is the second dollar heads? Yes. Is the third dollar heads? No. Is the fourth dollar heads? No. And so on. A datum that can answer a single yes-no question is called a *bit*—a familiar computer-age term that is short for *binary digit*, meaning a 0 or a 1, which you can think of as a numerical representation of *yes* or *no*. The heads-tails arrangement of the 1,000 coins thus contains 1,000 bits' worth of information. Equivalently, if you take Oscar's macroscopic perspective and focus only on the coins' overall haphazard appearance while eschewing the "microscopic" details of the heads-tails arrangement, the coins' "hidden" information content is 1,000 bits.

Notice that the value of the entropy and the amount of hidden information are equal. That's no accident. The number of possible heads-tails rearrangements *is* the number of possible answers to the 1,000 questions—(yes, yes, no, no, yes, . . .) or (yes, no, yes, yes, no, . . .) or (no, yes, no, no, no, . . .), and so on—namely, 2^{1000} . With entropy defined as the logarithm of the number of such rearrangements—1,000 in this case—

entropy *is* the number of yes-no questions any one such sequence answers.

I've focused on the 1,000 coins so as to offer a specific example, but the link between entropy and information is general. The microscopic details of any system contain information that's hidden when we take account of only macroscopic, overall features. For instance, you know the temperature, pressure, and volume of a vat of steam, but did an H_2O molecule just hit the upper right-hand corner of the box? Did another just hit the mid-point of the lower left edge? As with the dropped dollars, *a system's entropy is the number of yes-no questions that its microscopic details have the capacity to answer, and so the entropy is a measure of the system's hidden information content.*⁶

Entropy, Hidden Information, and Black Holes

How does this notion of entropy, and its relation to hidden information, apply to black holes? When Hawking worked out the detailed quantum mechanical argument linking a black hole's entropy to its surface area, he

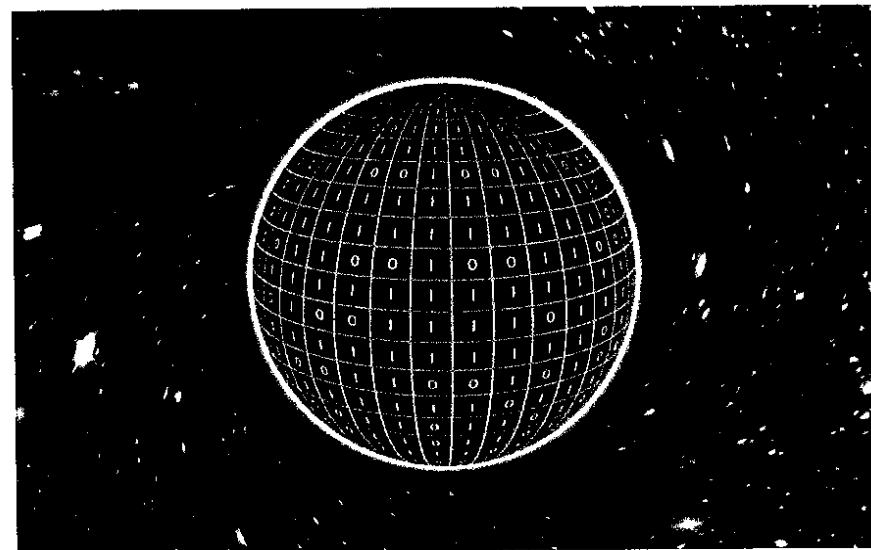


Figure 9.2 Stephen Hawking showed mathematically that the entropy of a black hole equals the number of Planck-sized cells that it takes to cover its event horizon. It's as if each cell carries one bit, one basic unit of information.

not only brought quantitative precision to Bekenstein's original suggestion, he also provided an algorithm for calculating it. Take the event horizon of a black hole, Hawking instructed, and divide it into a gridlike pattern in which the sides of each cell are one Planck length (10^{-33} centimeters) long. Hawking proved mathematically that the black hole's entropy is the number of such cells needed to cover its event horizon—the black hole's surface area, that is, as measured in square Planck units (10^{-66} square centimeters per cell). In the language of hidden information, it's as if each such cell secretly carries a single bit, a 0 or a 1, that provides the answer to a single yes-no question delineating some aspect of the black hole's microscopic makeup.⁷ This is schematically illustrated in Figure 9.2.

Einstein's general relativity, as well as the black hole no-hair theorems, ignores quantum mechanics and so completely misses this information. Choose values for its mass, its charge, and its angular momentum, and you've uniquely specified a black hole, says general relativity. But the most straightforward reading of Bekenstein and Hawking tells us you haven't. Their work established that there must be many different black holes with the same macroscopic features that, nevertheless, differ microscopically. And much as is the case in more commonplace settings—coins on the floor, steam in a vat—the black hole's entropy reflects information hidden within the finer details.

Exotic as black holes may be, these developments suggested that, when it comes to entropy, black holes behave much like everything else. But the results also raised puzzles. Although Bekenstein and Hawking tell us how much information is hidden within a black hole, they don't tell us what that information is. They don't tell us the specific yes-no questions the information answers, nor do they even specify the microscopic constituents that the information is meant to describe. The mathematical analyses pinned down the *quantity* of information a given black hole contains, without providing insight into the information itself.⁸

These were—and remain—perplexing issues. But there's yet another puzzle, one that seems even more basic: Why would the amount of information be dictated by the area of the black hole's surface? I mean, if you asked me how much information was stored in the Library of Congress, I'd want to know about the available space *inside* the Library of Congress. I'd want to know the capacity, within the library's cavernous interior, for shelving books, filing microfiche, and stacking maps, photographs, and documents. The same goes for the information in my head, which seems tied to the volume of my brain, the available space for neural interconnections. And it goes for the information in a vat of steam,

which is stored in the properties of the particles that fill the container. But, surprisingly, Bekenstein and Hawking established that for a black hole, the information storage capacity is determined not by the volume of its interior but by the area of its surface.

Prior to these results, physicists had reasoned that since the Planck length (10^{-33} centimeters) was apparently the shortest length for which the notion of "distance" continues to have meaning, the smallest meaningful volume would be a tiny cube whose edges were each one Planck length long (a volume of 10^{-99} cubic centimeters). A reasonable conjecture, widely believed, was that irrespective of future technological breakthroughs, the smallest possible volume could store no more than the smallest unit of information—one bit. And so the expectation was that a region of space would max out its information storage capacity when the number of bits it contained equaled the number of Planck cubes that could fit inside it. That Hawking's result involved the Planck length was therefore not surprising. The surprise was that the black hole's storehouse of hidden information was determined by the number of Planck-sized squares covering its surface and not by the number of Planck-sized cubes filling its volume.

This was the first hint of holography—information storage capacity determined by the area of a bounding surface and not by the volume interior to that surface. Through twists and turns across three subsequent decades, this hint would evolve into a dramatic new way of thinking about the laws of physics.

Locating a Black Hole's Hidden Information

The Planckian chessboard with 0s and 1s scattered across the event horizon, Figure 9.2, is a symbolic illustration of Hawking's result for the amount of information harbored by a black hole. But how literally can we take the imagery? When the math says that a black hole's store of information is measured by its surface area, does that merely reflect a numerical accounting, or does it mean that the black hole's surface is where the information is actually stored?

It's a deep issue and has been pursued for decades by some of the most renowned physicists.* The answer depends sensitively on whether you view the black hole from the outside or from the inside—and from the

*If you're interested in the full story, I highly recommend Leonard Susskind's excellent book *The Black Hole Wars*.

outside, there's good reason to believe that information is indeed stored at the horizon.

To anyone familiar with the finer details of how general relativity depicts black holes, this is an astoundingly odd claim. General relativity makes clear that were you to fall through a black hole's event horizon, you would encounter nothing—no material surface, no signposts, no flashing lights—that would in any way mark your crossing the boundary of no return. It's a conclusion that derives from one of Einstein's simplest but most pivotal insights. Einstein realized that when you (or any object) assume free-fall motion, you become weightless; jump from a high diving board, and a scale strapped to your feet falls with you and so its reading drops to zero. In effect, you cancel gravity by giving in to it fully. From this, Einstein leaped to an immediate consequence. Based on what you experience in your immediate environment, there's no way for you to distinguish between freely falling toward a massive object and freely floating in the depths of empty space: in both situations you are perfectly weightless. Sure, if you look beyond your immediate environment and see, say, the earth's surface rapidly getting closer, that's a pretty good clue that it's time to pull your parachute cord. But if you are confined to a small, windowless capsule, the experiences of free fall and free float are indistinguishable.⁹

In the early years of the twentieth century, Einstein seized on this simple but profound interconnection between motion and gravity; after a decade of development, he leveraged it into his general theory of relativity. Our application here is more modest. Suppose you are in that capsule and are freely falling not toward the earth but toward a black hole. The very same reasoning ensures that there's no way for your experience to be any different from floating in empty space. And that means that nothing special or unusual will happen as you freely fall through the black hole's horizon. When you eventually hit the black hole's center, you'll no longer be in free fall, and that experience will certainly distinguish itself. And spectacularly so. But until then, you could just as well be aimlessly floating in the dark depths of outer space.

This realization renders the black hole's entropy all the more puzzling. If as you pass through the horizon of a black hole you find nothing there, nothing at all to distinguish it from empty space, how can it store information?

An answer that has gained traction over the last decade resonates with the duality theme encountered in early chapters. Recall that duality refers to a situation in which there are complementary perspectives that seem completely different, and yet are intimately connected through a shared physical anchor. The Albert-Marilyn image of Figure 5.2 provides a good

visual metaphor; mathematical examples come from the mirror shapes of string theory's extra dimensions (Chapter 4) and the naively distinct yet dual string theories (Chapter 5). In recent years, researchers, led by Susskind, have realized that black holes present another context in which complementary yet widely divergent perspectives yield fundamental insight.

One essential perspective is yours, as you freely fall toward a black hole. Another is that of a distant observer, watching your journey through a powerful telescope. The remarkable thing is that as you pass uneventfully through a black hole's horizon, the distant observer perceives a very different sequence of events. The discrepancy has to do with the black hole's Hawking radiation.* When the distant observer measures the Hawking radiation's temperature, she finds it to be tiny; let's say it's 10^{-13} K, indicating that the black hole is roughly the size of the one at the center of our galaxy. But the distant observer knows that the radiation is cold only because the photons, traveling to her from just outside the horizon, have expended their energy valiantly fighting against the black hole's gravitational pull; in the description I gave earlier, the photons are tired. She deduces that as you get ever closer to the black hole's horizon, you'll encounter ever-fresher photons, ones that have only just begun their journey and so are ever more energetic and ever hotter. Indeed, as she watches you approach to within a hair's breadth of the horizon, she sees your body bombarded by increasingly intense Hawking radiation, until finally all that's left is your charred remains.

Happily, however, what you experience is much more pleasant. You don't see or feel or otherwise obtain any evidence of this hot radiation. Again, because your free-fall motion cancels the effects of gravity,¹⁰ your experience is indistinguishable from that of floating in empty space. And one thing we know for sure is that when you float in empty space, you don't suddenly burst into flames. So the conclusion is that from your perspective, you pass seamlessly through the horizon and (less happily) hurtle on toward the black hole's singularity, while from the distant observer's perspective, you are immolated by a scorching corona that surrounds the horizon.

Which perspective is right? The claim advanced by Susskind and others is that both are. Granted, this is hard to square with ordinary logic—the logic by which you are either alive or not alive. But this is no ordinary

*The reader familiar with black holes will note that even without the quantum considerations that lead to Hawking radiation, the two perspectives would differ with regards to the rate of time's passage. Hawking radiation makes the perspectives yet more distinct.

situation. Most saliently, the wildly different perspectives can never confront each other. You can't climb out of the black hole and prove to the distant observer that you are alive. And, as it turns out, the distant observer can't jump into the black hole and confront you with evidence that you're not. When I said that the distant observer "sees" you immolated by the black hole's Hawking radiation, that was a simplification. The distant observer, by closely examining the tired radiation that reaches her, can piece together the story of your fiery demise. But for the information to reach her takes time. And the math shows that by the time she can conclude you've burned, she won't have enough time left to then hop into the black hole and catch up with you before you're destroyed by the singularity. Perspectives can differ, but physics has a built-in fail-safe against paradoxes.

What about information? From your perspective, all your information, stored in your body and brain and in the laptop you're holding, passes with you through the black hole's horizon. From the perspective of the distant observer, all the information you carry is absorbed by the layer of radiation incessantly bubbling just above the horizon. The bits contained in your body, brain, and laptop would be preserved, but would become thoroughly scrambled as they joined, jostled, and intermingled with the sizzling hot horizon. Which means that to the distant observer, the event horizon is a real place, populated by real things that give physical expression to the information symbolically depicted in the chessboard, Figure 9.2.

The conclusion is that the distant observer—us—infers that a black hole's entropy is determined by the area of its horizon because the horizon is where the entropy is stored. Said that way, it seems utterly sensible. But don't lose sight of how unexpected it is that the storage capacity isn't set by the black hole's volume. And, as we will now see, this result doesn't merely highlight a peculiar feature of black holes. Black holes don't just tell us about how black holes store information. Black holes inform us about information storage in any context. This paves a direct path to the holographic perspective.

Beyond Black Holes

Consider any object or collection of objects—the collections of the Library of Congress, all of Google's computers, the CIA's archives—situated in some region of space. For ease, imagine that we highlight the

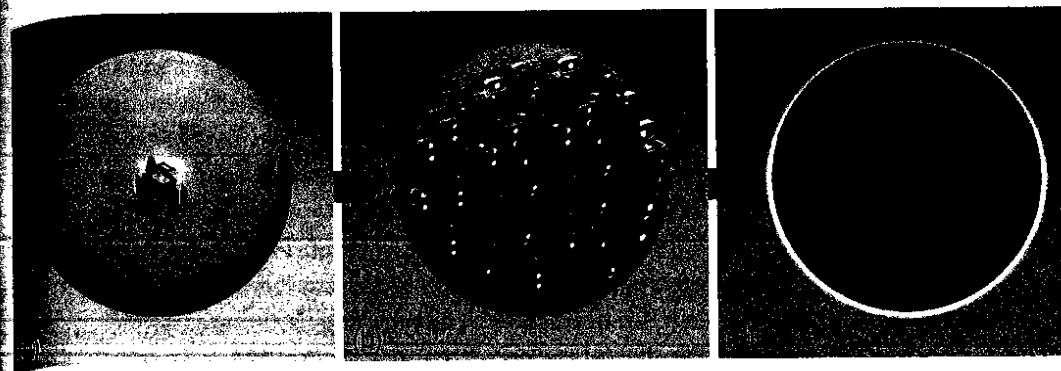


Figure 9.3 (a) A variety of objects that store information, situated within a well-marked region of space. (b) We augment the region's capacity for storing information. (c) When the amount of matter crosses a threshold (whose value can be calculated from general relativity),¹¹ the region becomes a black hole.

region by surrounding it with an imaginary sphere, as in Figure 9.3a. Assume further that the total mass of the objects, compared with the volume they fill, is of such an ordinary run-of-the-mill magnitude that it's nowhere near what it takes to create a black hole. That's the setup. Now for the pivotal question: What is the maximum amount of information that can be stored within the region of space?

Those unlikely bedfellows, the Second Law and black holes, provide the answer. Imagine adding matter to the region, with the aim of augmenting its information storage capacity. You might insert high-capacity memory chips or voluminous hard drives into the bank of Google's computers; you might provide books or jam-packed Kindles to augment the Library of Congress collection. Since even raw matter carries information—Are the steam's molecules here or there? Are they moving at this speed or that?—you also cram every nook and cranny of the region with as much matter as you can get your hands on. Until you reach a critical juncture. At some point, the region will be so thoroughly stuffed that were you to add even a single grain of sand, the interior would go dark as the region turned into a black hole. When that happens, game over. A black hole's size is determined by its mass, so if you try to increase the information storage capacity by adding yet more matter, the black hole will respond by growing larger. And since we want to focus on the information that can inhabit a given *fixed* volume of space, this result falls afoul of the basic setup. You can't increase the black hole's information capacity without forcing the black hole to enlarge.¹²

Two observations take us across the finish line. The Second Law ensures that entropy increases throughout the entire process, and so the information hidden within the hard drives, Kindles, old-fashioned paper books, and everything else you packed into the region is less than that hidden in the black hole. From the results of Bekenstein and Hawking, we know that the black hole's hidden information content is given by the area of its event horizon. Moreover, because you were careful not to overspill the original region of space, the black hole's event horizon coincides with the region's boundary, so the black hole's entropy equals the area of this surrounding surface. We thus learn an important lesson. *The amount of information contained within a region of space, stored in any objects of any design, is always less than the area of the surface that surrounds the region (measured in square Planck units).*

This is the conclusion we've been chasing. Notice that although black holes are central to the reasoning, the analysis applies to *any* region of space, whether or not a black hole is actually present. If you max out a region's storage capacity, you'll create a black hole, but as long as you stay under the limit, no black hole will form.

I hasten to add that in any practical sense, the information storage limit is of no concern. Compared with today's rudimentary storage devices, the potential storage capacity on the surface of a spatial region is humongous. A stack of five off-the-shelf terabyte hard drives fits comfortably within a sphere of radius 50 centimeters, whose surface is covered by about 10^{70} Planck cells. The surface's storage capacity is thus about 10^{70} bits, which is about a billion, trillion, trillion, trillion, trillion terabytes, and so enormously exceeds anything you can buy. No one in Silicon Valley cares much about these theoretical constraints.

Yet, as a guide to how the universe works, the storage limitations are telling. Think of any region of space, such as the room in which I'm writing or the one in which you're reading. Take a Wheelerian perspective and imagine that whatever happens in the region amounts to information processing—information regarding how things are right now is transformed by the laws of physics into information regarding how they will be in a second or a minute or an hour. Since the physical processes we witness, as well as those by which we're governed, seemingly take place within the region, it's natural to expect that the information those processes carry is also found within the region. But the results just derived suggest an alternative view. For black holes, we found that the link between information and surface area goes beyond mere numerical accounting; there's a concrete sense in which information is stored on their surfaces. Susskind and 't Hooft stressed that the lesson should be

general: since the information required to describe physical phenomena within *any* given region of space can be fully encoded by data on a surface that surrounds the region, then there's reason to think that the surface is where the fundamental physical processes actually happen. Our familiar three-dimensional reality, these bold thinkers suggested, would then be likened to a holographic projection of those distant two-dimensional physical processes.

If this line of reasoning is correct, then there are physical processes taking place on some distant surface that, much like a puppeteer pulls strings, are fully linked to the processes taking place in my fingers, arms, and brain as I type these words at my desk. Our experiences here, and that distant reality there, would form the most interlocked of parallel worlds. Phenomena in the two—I'll call them *Holographic Parallel Universes*—would be so fully joined that their respective evolutions would be as connected as me and my shadow.

The Fine Print

That familiar reality may be mirrored, or perhaps even produced, by phenomena taking place on a faraway, lower-dimensional surface ranks among the most unexpected developments in all of theoretical physics. But how confident should we be that the holographic principle is right? We are navigating a realm deep in theoretical territory, and relying almost exclusively on developments that have not been experimentally tested, so there is surely grounds for skepticism. There are many places where the argument could be forced off course. Do black holes really have nonzero entropy and nonzero temperature, and, if so, do the values conform to theoretical predictions? Is the information capacity of a region of space really determined by the amount of information that can be stored on a surface that surrounds it? And on such a surface, is one bit per Planck area really the limit? We think the answer to each of these questions is yes because of the coherent, consistent, and carefully constructed theoretical edifice into which the conclusions perfectly fit. But since none of these ideas has been subject to the experimenter's scalpel, it is certainly possible (though in my view highly unlikely) that future advances will convince us that one or more of these essential intermediate steps are wrong. That could lay to waste the holographic idea.

Another important point is that throughout the discussion, we've spoken of a region of space, of a surface that surrounds it, and of the information content of each. But since our focus has been on entropy and the

Second Law—both of which concern themselves primarily with the *quantity* of information in a given context—we’ve not elaborated on the details of *how* that information is physically realized or stored. When we talk about information residing on a sphere surrounding a region of space, what does that really mean? How does the information manifest itself? What form does it take? To what extent can we develop an explicit dictionary that translates from phenomena taking place on the boundary to those taking place in the interior?

Physicists have yet to articulate a general framework for addressing these questions. Given that gravity and quantum mechanics are both central to the reasoning, you might expect that string theory would provide a potent context for theoretical explorations. But when ’t Hooft first formulated the holographic concept, he doubted that string theory would be able to advance the subject, noting, “Nature is much more crazy at the Planck scale than even string theorists could have imagined.”¹³ Less than a decade later, string theory proved ’t Hooft wrong by proving him right. In a landmark paper, a young theorist showed that string theory provides an explicit realization of the holographic principle.

String Theory and Holography

When I was called to the stage at the University of California, Santa Barbara, to give my talk at the annual international string theory conference in 1998, I did something I’d never done before and suspect will never do again. I faced the audience, threw my right hand to my left shoulder and my left to my right shoulder, and then with both hands in succession grabbed the seat of my pants, bunny-hopped, and made a quarter turn, followed, thankfully, by audience laughter, which covered the three remaining steps necessary to reach the podium, where I began my talk. The crowd got the joke. At the banquet the night before, the conference participants had performed a song-and-dance celebrating—as only physicists can—a spectacular result of the Argentinian string theorist Juan Maldacena. With lyrics like “Black holes used to be a great mystery; / Now we use D-branes to compute D-entropy,” the crowd had reveled in a string theory version of the 1990s momentary dance craze, the Macarena—a touch more animated than Al Gore’s version at the Democratic National Convention, a touch less mellifluous than Los del Rio’s original one-hit wonder, but second to none in passion. I was one of the few at the conference whose talk was not focused on Maldacena’s breakthrough, so when I

took the stage the next morning I felt it only appropriate to preface my remarks with a personal gesture of appreciation.

Now, more than a decade later, many would agree that no work in string theory since is of comparable magnitude and influence. Of the numerous ramifications of Maldacena’s result, one is directly relevant to the line we’ve been following. In a particular hypothetical setting, Maldacena’s result *realized explicitly the holographic principle, and in doing so provided the first mathematical example of Holographic Parallel Universes*. Maldacena achieved this by considering string theory in a universe whose shape differs from ours but for the purpose at hand proves easier to analyze. In a precise mathematical sense, the shape has a boundary, an impenetrable surface that completely surrounds its interior. By zeroing in on this surface, Maldacena argued convincingly that everything taking place within the specified universe is a reflection of laws and processes acting themselves out on the boundary.

Although Maldacena’s method may not seem directly applicable to a universe with the shape of ours, his results are decisive because they established a mathematical proving ground in which ideas regarding holographic universes could be made explicit and investigated quantitatively. The results of such studies won over a great many physicists who had previously eyed the holographic principle with much misgiving, and thus set off an avalanche of research that has yielded thousands of articles and considerably deeper understanding. Most exciting of all, there’s now evidence that a link between these theoretical insights and physics in our universe *can* be forged. In the next few years, that link may very well allow the holographic ideas to be experimentally tested.

The rest of this and the next section will be devoted to explaining how Maldacena achieved this breakthrough; the material is the most difficult we will cover. I’ll begin with a short summary, a CliffsNotes version that doubles as a guilt-free pass to jump to the last section should, at any point, the material overwhelm your appetite for detail.

Maldacena’s inspired move was to invoke a new version of the duality arguments discussed in Chapter 5. Recall the branes—the “slice of bread” universes—introduced there. Maldacena considered, from two complementary perspectives, the properties of a tightly stacked collection of three-dimensional branes, as in Figure 9.4. One perspective, an “intrinsic” perspective, focused on strings that move, vibrate, and wiggle along the branes themselves. The other perspective, an “extrinsic” perspective, focused on how the branes influence their immediate environment gravitationally, much as the sun and the earth influence theirs. Maldacena