PHYSICS

BURNING RINGS OF FIRE

"Firewalls" of particles may border black holes, confounding both general relativity and quantum mechanics

By Joseph Polchinski

FALLING INTO A BLACK HOLE WAS NEVER GOING TO be fun. As soon as physicists realized that black holes exist, we knew that getting too close to one spelled certain death. But we used to think that an astronaut falling past the point of no return—the so-called event horizon—would not feel anything special. According to Einstein's general theory of relativity, no signposts would mark the spot where the chance of escape dropped to zero. Anyone journeying past the horizon would just seem to fall down, down, down into a pit of blackness.

Recently, however, my colleagues and I have recast that picture in light of some new information about the effects of quantum mechanics on black holes. It now seems that our astronaut would have an experience very different from Albert Einstein's prediction. Rather than

Illustration by Kenn Brown, Mondolithic Studios falling seamlessly into the interior, the astronaut would encounter a “firewall” of high-energy particles at the horizon that would be instantly lethal. The wall might even mark the end of space.

Three years ago four of us, all then at the University of California, Santa Barbara—my colleague Donald Marolf, then graduate students Ahmed Almheiri and James Sully, and I (now known by the acronym AMPS)—arrived at this conclusion after using ideas from string theory to take a closer look at the physics of black holes, particularly at an interesting argument put forward in the 1970s by Stephen Hawking. Hawking had identified a deep conflict between the predictions of quantum theory and relativity in these extreme environments. According to his reasoning, either quantum mechanics or Einstein’s depiction of spacetime is flawed. The battle over which view is correct has swung back and forth ever since.

As with Hawking’s original claim, our recent firewall proposal has raised a storm of disbelief, and no satisfactory alternative has yet emerged. If quantum mechanics is to be trusted, firewalls are the consequence. Yet their existence raises theoretical puzzles as well. It seems that physicists must give up one of our widely cherished beliefs, but we cannot agree on which one. We hope, however, that out of this confusion will come a more complete understanding of quantum mechanics and relativity—and, ideally, a way to finally resolve the apparent contradictions between these two reigning theories of physics.

THE SINGULARITY

GENERAL RELATIVITY, which gave birth to the very concept of black holes, derives its picture of these mysterious entities and their event horizons from an understanding of gravity’s effect on space and time. According to the theory, if enough mass comes together, gravity’s pull will cause it to start collapsing. Nothing can stop this process until all the mass is compressed into a single point where spacetime is infinitely dense and infinitely curved, called the singularity—in other words, a black hole.

Any space travelers who pass the black hole’s event horizon boundary will be unable to escape the gravitational pull and will soon be drawn into the singularity. Even light, once it is past the horizon, cannot escape. The singularity is a very dramatic place, but the horizon itself is supposed to be unremarkable, according to what is called the equivalence principle of general relativity; individuals falling freely into a black hole will see the same physical laws as anywhere else as they cross the horizon. Theorists are fond of saying that the entire solar system could be falling into a giant black hole right now, and we would not experience anything out of the ordinary.

BLACK HOLE RADIATION

THE CHALLENGE HAWKING POSED to the traditional picture of black holes began in 1974, when he considered a strange prediction of quantum mechanics. According to this theory, pairs of particles and their antimatter counterparts constantly pop into existence and then disappear almost at once. If such fluctuations happen just outside the horizon of the black hole, Hawking showed, the pair could separate. One would fall into the singularity, and the other would escape from the black hole and carry away some of its mass. Eventually the black hole’s entire mass could be depleted through this process, termed Hawking evaporation.

Joseph Polchinski is a professor of physics at the University of California, Santa Barbara, and a permanent member of the Kavli Institute for Theoretical Physics there. He works on many areas of theoretical physics, but he is guided by two big questions: How does duality work, and what is quantum gravity? In his spare time, he rides his bike in the hills of Santa Barbara.

For black holes found in nature, evaporation is unimportant: these black holes add mass at a much more rapid rate from gas and dust falling in than they lose to radiation. But for theoretical purposes, we can investigate what would happen if a black hole were completely isolated and we had enough time to watch the full process of evaporation. By pursuing such a thought experiment, Hawking revealed two apparent contradictions between general relativity and quantum mechanics.

The entropy problem. In pondering the isolated black hole, Hawking noted that the light spectrum of the eponymous radiation streaming away from it would look the same as that of a radiating hot body, meaning that the black hole has a temperature. In general, temperature arises from the motion of atoms inside objects. The thermal nature of Hawking radiation, then, suggested that the black hole should have a microscopic structure made of some kind of discrete building blocks or bits. Physicist Jacob D. Bekenstein, now at the Hebrew University of Jerusalem, had also reached this conclusion two years earlier by engaging in thought experiments involving throwing things into black holes. The work of Bekenstein and Hawking gives a formula for the number of bits, a measure known as the black hole entropy. Entropy is a gauge of disorder, which becomes greater as the number of states that an object can have grows. The larger the number of bits in a black hole, the more possible arrangements they can have and the greater the entropy.

In contrast, general relativity describes a black hole as having a smooth geometry and indicates that every black hole of given mass, spin and charge should be exactly the same: in the words of the late physicist John Wheeler of Princeton University, “Black holes have no hair.” So here is a contradiction: relativity says no

IN BRIEF

Stephen Hawking’s discovery that particles leak out of black holes revealed a fissure in scientists’ understanding of physics. These escaped particles seem to imply that information is destroyed inside black holes—something quantum mechanics forbids.

An attempt to resolve this quandary using string theory looked promising, but recent calculations show that black holes are even more perplexing than was thought.

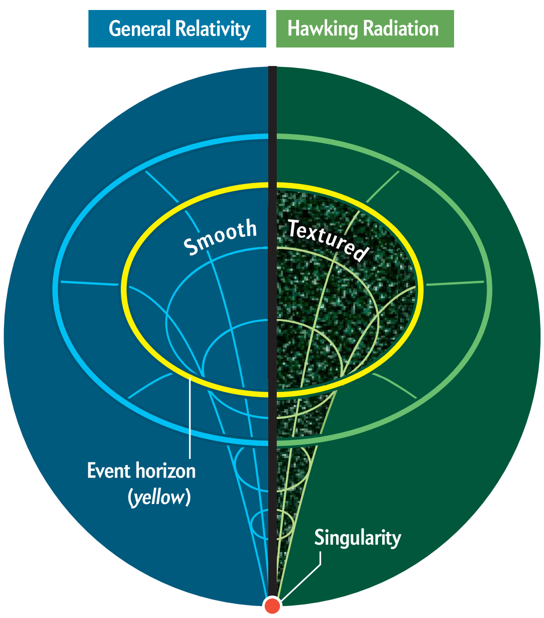
Barriers of high-energy particles called firewalls surround black holes, according to calculations by the author and his colleagues. Such firewalls may represent the end of space itself. Resolving the paradoxes of firewalls could offer a path toward unifying quantum mechanics and general relativity.

Resolving Black Hole Conundrums

In 1974 Stephen Hawking showed that a small amount of radiation leaks out of black holes. According to quantum mechanics, pairs of particles and their antimatter counterparts constantly spring into existence and then disappear moments later all over the universe. Hawking noted that when a pair shows up near the horizon of a black hole, one particle could fall in while the other escapes. This phenomenon, called Hawking radiation, raises some puzzles about the laws of physics inside black holes.

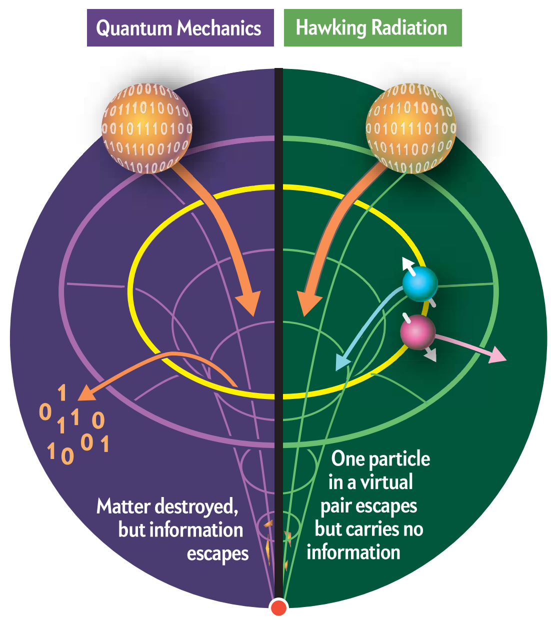
The Entropy Problem

The radiation spectrum of Hawking emission suggests that black holes have temperatures. Traditionally, heat arises from the motion of atoms within an object. The temperature of black holes implies that they have substructure—some type of internal building blocks that can rearrange themselves. The possibility of different arrangements gives black holes a measure of disorder, or "entropy," according to the quantum-mechanical picture of Hawking radiation. Entropy is forbidden to black holes by general relativity, however, because the theory requires them to be completely smooth, without substructure.



The Information Paradox

According to the standard picture of quantum mechanics, information can never be destroyed. Even when you burn a letter, for example, the original information encoded in the atoms of the letter is preserved in the ashes. Hawking radiation, however, implies that black holes destroy the information of the matter that falls into them because the particles that escape do not depend at all on the properties of the atoms that initially fell into the hole. Hawking suggested that quantum mechanics might have to be modified to allow for information loss.



Earlier Conjectures (*not shown*) ...

In an effort to resolve these puzzles, physicists looked for new ways to combine general relativity and quantum mechanics into a coherent theory that could describe black holes. One breakthrough was string theory, which posits that particles are actually tiny loops of vibrating string. This theory appeared to solve elements of the information paradox and the entropy problem.

... Led to Firewalls

Yet the string theory solutions eventually led to a surprising conclusion: black holes might be surrounded by firewalls—walls of high-energy particles that would obliterate any object that encountered them. Firewalls seem to imply a drastic breakdown of the laws of physics at the boundary of black holes and could lead to extreme conclusions, such as the possibility that firewalls mark the end of space and time altogether.

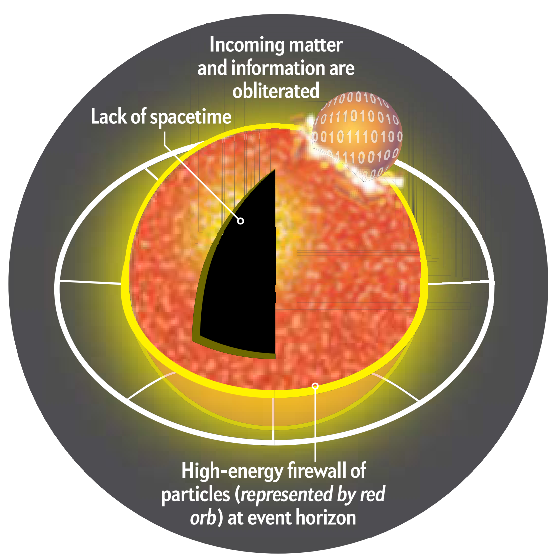


Illustration by Jen Christiansen

hair, whereas quantum mechanics says black holes have a large amount of entropy, meaning some microscopic structure, or hair.

The information paradox. Hawking evaporation also gives rise to a challenge to quantum theory. According to Hawking's calculation, the particles that escape from a black hole do not depend at all on the properties of the material that went into the hole—usually a massive star that collapsed. For example, we could send a note with a message into the black hole, and there would then be no way to reconstruct the message from the final particles that would emerge. Once the note passed through the horizon, it could not influence anything that came out later, because no information can escape from the interior. In quantum mechanics, every system is described by a formula called the wave function, which encodes the chances that the system will be in any particular state.

In Hawking's thought experiment, the loss of information means that we have no way to predict the wave function of Hawking radiation based on the properties of the mass that went into the black hole. Information loss is forbidden by quantum mechanics, so Hawking concluded that the laws of quantum physics had to be modified to allow for such loss in black holes.

You might be saying to yourself, "Of course, black holes destroy information—they destroy everything that enters them." But compare what happens if we simply burned the note. The message would certainly be scrambled, and it would be impractical to reconstruct it from the smoke. But the process of burning is described by ordinary quantum mechanics, applied to the atoms in the note, and the quantum description of the smoke would be a definite wave function that would depend on the original message. In theory, then, the message could be reconstructed through the wave function. In the case of black holes, however, there would be no definite wave function for the resulting radiation.

Based on this analogy, many theorists concluded that Hawking was wrong, that he had mistaken the scrambling of information for actual information loss. Further, some argued, if information can be lost, then it will not just happen in the exotic situation of black hole evaporation but everywhere and all the time—in quantum physics, anything that can happen will happen. If Hawking were right, we would see the signs in everyday physics, probably including severe violations of the law of conservation of energy.

Hawking's argument, though, stands up to simple objections. Unlike burning paper, black holes have horizons beyond which information cannot escape. Thus, we seem to have a sharp paradox: either modify quantum mechanics to allow information loss or modify relativity to allow information to escape from the black hole interior.

A third possibility also exists—that the black hole does not evaporate completely but ends up as a microscopic remnant containing all the information of the original star that created it. This "solution" has its own difficulties, however. For example, such a small object containing so much information would violate the Bekenstein-Hawking idea of entropy.

BLACK HOLES AND BRANES

STRING THEORY is one attempt to rectify some of the problems that arise when relativity and quantum mechanics collide, as in the case of black holes. This theory replaces the pointlike particles of previous theories with tiny loops or strands of string; these strings manage to eliminate some of the mathematical difficulties that arise when quantum mechanics and relativity are combined. Replacing points with strings does not, however, immediately change the black hole story.

A break came in 1995, when I was looking at another kind of thought experiment, studying strings in small spaces. Building on work that I and several others had done a few years earlier, I showed that string theory, as it was then understood, was not complete. Rather it required the existence of objects with more dimensions than the three of space and one of time we are familiar with. In black holes these higher-dimensional objects, called D-branes, would be tiny—wrapped up in hidden dimensions too small for us to detect. The next year Andrew Strominger and Cumrun Vafa, both now at Harvard University, showed that strings and D-branes together provide the precise number of bits to account for black hole entropy, at least for certain very symmetrical black holes. The entropy puzzle was partly solved.

The next question was, What about information loss? Then, in 1997, Juan Maldacena, now at the Institute for Advanced Study in Princeton, N.J., came up with a way around the information loss problem—a solution sometimes called the Maldacena duality. A duality is a surprising equivalence between two things that seem very different. Maldacena's duality shows that the mathematics of a theory combining quantum mechanics and gravity—a quantum theory of gravity—based on string theory are equivalent to the mathematics of an ordinary quantum theory under a special set of circumstances. In particular, the quantum physics of a black hole is equivalent to that of an ordinary gas of hot nuclear particles. It also means that spacetime is fundamentally different from what we perceive, more like a three-dimensional hologram projected from a more fundamental two-dimensional surface of a sphere.

Using Maldacena's duality, physicists also get a way to describe the quantum mechanics of black holes in the bargain. If Maldacena's assumptions are true, then ordinary quantum laws would apply to gravity as well, and information cannot be lost. By a less direct argument, evaporating black holes cannot leave behind any remnants, so it must be that the information gets out with the Hawking radiation.

Maldacena's duality is arguably the closest we have come to unifying general relativity and quantum mechanics, and Maldacena discovered it by chasing down the black hole puzzles of entropy and information loss. It is not yet proved to be true, but it is supported by much evidence—enough that in 2004 Hawking announced that he had changed his mind about the need for black holes to lose information and publicly paid off a bet with physicist John Preskill at the International Conference on General Relativity and Gravitation in Dublin.

Physicists widely believed that no single observer would see any violation of relativity or any other laws near a black hole that lived by Maldacena's rules, although his duality falls short in not giving a clear explanation for how information gets from the inside of a black hole to the outside.

About 20 years ago Leonard Susskind of Stanford University and Gerard 't Hooft of Utrecht University in the Netherlands proposed a solution to the original information problem that involves a kind of relativity principle called black hole complementarity. In essence, the argument holds that an observer who jumps into a black hole sees the information inside, whereas one who stays outside sees it come out. There is no contradiction because these two observers cannot communicate.

THE FIREWALL

MALDACENA'S DUALITY and black hole complementarity seemed to dispel all the paradoxes, but many of the details had yet to be filled in. Three years ago my own AMPS collaboration tried to make a model of how the combined picture would work, building on ideas of physicists Samir D. Mathur of Ohio State University and Steven Giddings of U.C. Santa Barbara (and extending, unbeknownst to us, an earlier argument of Samuel Braunstein of the University of York in England). After failing repeatedly to make a successful model, we realized that the problem ran deeper than our mathematical shortcomings and that a contradiction remained.

This contradiction pops up when considering the phenomenon of quantum entanglement—the most unintuitive part of quantum theory and the one furthest from our experience. If particles were like dice, entangled particles would be two dice that always added to seven: if you roll the dice, and the first comes up as two, then the second will always come up as five, and so on. Similarly, when scientists measure the properties of one entangled particle, the measurement also determines the characteristics of its partner. It is a further consequence of quantum theory that a particle can be fully entangled only with one other: if particle B is entangled with particle A, then it cannot also be entangled with particle C. Entanglement is monogamous.

In the case of the black hole, think about a Hawking photon; call it "B," emitted after the black hole is at least halfway evaporated. The Hawking process implies that B is part of a pair; call its partner that falls into the black hole "A." A and B are entangled. Furthermore, the information that originally fell into the black hole has been encoded into all the Hawking radiation particles. Now, if information is not lost, and the outgoing Hawking photon B ends up in a definite quantum state, then B must be entangled with some combination, "C," of the other Hawking particles that already escaped (otherwise, the output would not preserve the information). But then we have a contradiction: polygamy!

The price of saving quantum mechanics, keeping the entanglement between B and C and not having anything else out of the ordinary on the outside of the black hole, is the loss of entanglement between A and B. The Hawking photons A and B began just inside and outside the horizon when they arose as an ephemeral particle-antiparticle pair. In quantum theory, the cost of breaking this entanglement, like the cost of breaking a chemical bond, is energy. Breaking the entanglement for all the Hawking pairs implies that the horizon is a wall of high-energy particles, which we termed a firewall. An infalling astronaut, rather than moving freely through the horizon, encounters something dramatic.

Finding such a large departure from general relativity—a wall of energy in a place where nothing unusual should be happening—was disturbing, but the argument was simple, and we could not find a flaw. In a sense, we had just run Hawking's original argument backward, assuming that information is not lost and seeing where that assumption would lead. We concluded that, rather than the subtle effects of complementarity, there was a drastic breakdown of general relativity. As we began to describe the argument to others, the common reaction was first skepticism and then the same puzzlement that we experienced.

Either these strange firewalls actually exist, or it seems we must again consider letting go of some of the deeply held doctrines of quantum theory. Information may not be destroyed, but perhaps some rewriting of quantum mechanics is in store. Unfortunately, observing real black holes will not decide the issue—any radiation from a firewall would be weakened by the gravitational pull of the black hole, making the firewall very hard to see.

THE END OF SPACE

FURTHERMORE, if the firewall exists, what is it? One idea is that the firewall is simply the end of space. Perhaps the conditions for spacetime to form do not exist inside the black hole. As Marolf once remarked, maybe the interior cannot form, because "the black hole's quantum memory is full." If spacetime cannot occur inside, then space ends at the horizon, and an infalling astronaut who hits it dissolves into quantum bits residing on this boundary.

To avoid such bizarre scenarios, physicists have attempted to circumvent the firewall conclusion. One idea is that because Hawking radiation particle B must be entangled with both A and C, then A must be part of C: the photon behind the horizon is somehow the same bit that is encoded in the earlier Hawking radiation, even though they are in very different places. This notion is something like the original idea of black hole complementarity, but to make a concrete model of this scenario, it seems, one ends up modifying quantum mechanics again. The most radical idea, from Maldacena and Susskind, is that every pair of entangled particles is connected by a microscopic spacetime wormhole, so that large regions of spacetime, such as the black hole interior, can be built up from large amounts of entanglement.

Hawking had proposed that general relativity works for black holes but that quantum mechanics breaks down. Maldacena concluded that quantum mechanics is unmodified but that spacetime is holographic. Perhaps the truth is somewhere in the middle.

Many other ideas have been proposed, most of which give up one long-standing principle or another, and there is no consensus as to the right direction to resolve the problems. A common question is, What do firewalls imply for real-life black holes, such as the one in the center of our Milky Way galaxy? It is too early to say.

For now investigators are excited that we have discovered a new contradiction between two of the central theories of physics. Our inability to say definitively whether or not the firewall is real exposes a limitation in our current formulations of quantum gravity, and theoretical physicists are rethinking their basic assumptions about the workings of the universe. Out of this may come a deeper understanding of the nature of space and time and of the principles underlying all the laws of physics. Ultimately, by unraveling the quandaries at the heart of black hole firewalls, we may finally get the break we need to unify quantum mechanics and general relativity into a single working theory. SA

MORE TO EXPLORE

The Black Hole War: My Battle with Stephen Hawking to Make the World Safe for Quantum Mechanics. Leonard Susskind. Little, Brown, 2008.

Black Holes: Complementarity or Firewalls? Ahmed Almheiri, Donald Marolf, Joseph Polchinski and James Sully in *Journal of High Energy Physics,* Vol. 2013, No. 2, Article No. 62; February 2013. Preprint available at **http://arxiv.org/abs/arXiv:1207.3123**

FROM OUR ARCHIVES

Black Holes and the Information Paradox. Leonard Susskind; April 1997.

Information in the Holographic Universe. Jacob D. Bekenstein; August 2003.

The Illusion of Gravity. Juan Maldacena; November 2005.

scientificamerican.com/magazine/sa