The Hole in Einstein's Theory

David Wakeham

Albert Einstein is best known for his theory of *special relativity*: a Dalí-esque realm where clocks melt, long poles fit into short barns, and mass equals energy. But his best was yet to come. In 1915, Einstein formulated *general relativity*, a theory which replaces Newtonian gravity the same way special relativity supersedes Newtonian mechanics. With its elegant connections between physics and geometry, it is, by many accounts, the most beautiful scientific theory ever devised. In John Wheeler's pithy summary,

Spacetime tells matter how to move.
Matter tells spacetime how to curve.

The sun curves spacetime like a rubber sheet (Fig. 1), and planets roll along their circular orbits like marbles. A theory this beautiful, Einstein believed, could not have holes in it.

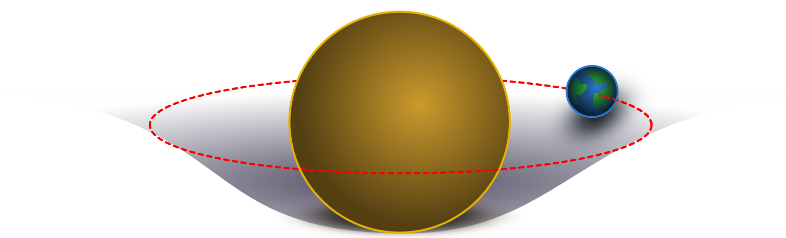

Figure 1. The sun bends spacetime and planets simply roll along this curved geometry.

A few weeks after Einstein presented his results, unexpected progress occurred on the Russian frontlines. Karl Schwarzschild, a German astrophysicist and artillery officer, solved Einstein's equations exactly to determine how a sphere of matter dimples the spacetime around it. The answer was sensible, giving the orbits we see in the sky. But for a very dense sphere, two surprises emerged: first, there is a *light-trapping region* around the sphere; second, if the sphere is squished to a point, the nearby spacetime is curved so dramatically that general relativity itself breaks down (Fig. 2). This is called a *singularity*, mathematical parlance for "disaster", and a sign the theory was less complete than Einstein thought. Schwarzschild died of pemphigus a year later, but bequeathed to the world the enigma of black holes.

Einstein admired the simplicity of Schwarzschild's solution, but was not prepared to accept there were gaps in general relativity. He argued that the singularity was an artefact, a bug due to Schwarzschild's assumption of perfect spherical symmetry. It would disappear, he concluded, "in the real world", where symmetries are harder to come by and matter would strenuously object to being squished. The maestro had spoken, and for most physicists, that settled the matter.

In 1964, almost 50 years after Schwarzschild's discovery and 10 years after Einstein's death, a professor of mathematics crossed a busy London street. He had been pondering the robustness

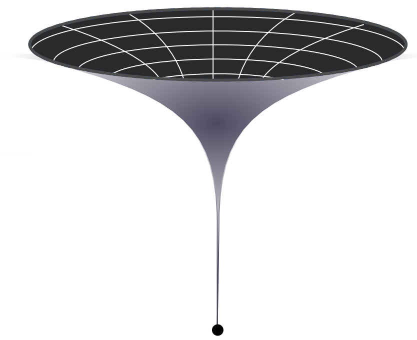

Figure 2. A Schwarzschild black hole and its singularity.

of singularities to asymmetry and matter, without making headway. Suddenly, an impression flashed into his mind—a skin of light, pulsing outwards and *getting smaller*—but he lost it again in the hustle and bustle of midday traffic. The course of 20th century physics might be very different if he had not by luck recalled it the following day. The professor was Roger Penrose, and his mental image would inspire the most important technical development since Schwarzschild: the *Penrose singularity theorem*, a mathematical result showing that singularities are not mere bugs arising from symmetry, but inevitable features of a curvy universe. It is perhaps a mercy Einstein did not live to see the most elegant theory ever devised, the apotheosis of his genius, robustly predict its own demise.

Penrose shared the 2020 Nobel Prize in Physics with Andrea Ghez and Reinhard Genzel, astronomers who discovered black holes in our very own Milky Way. In a way, Penrose laid the groundwork for these observations; if Einstein's prejudice still held sway, no one would bother searching for black holes! To understand where the holes in Einstein's theory come from, we will unpack Penrose's proof. The starting point is the idea that *gravity is attractive*, focusing light like a magnifying glass. If light rays are already converging, gravity cannot pry them apart and they will collide (Fig. 3). Penrose's flash of inspiration, as he crossed the street, was the notion of a *closed trapped surface*: a complete surface like a sphere, with an inside and an outside, such that light heading *outwards* from the sphere converges. We call this a "light trap". It may sound impossible for outgoing light rays to get closer, but this is exactly what happens in the light-trapping region of a black hole.

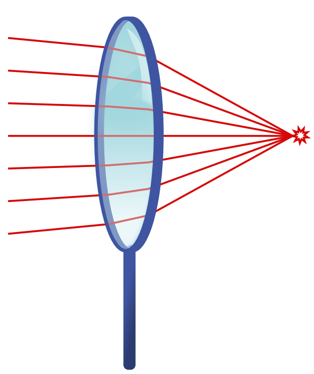

Figure 3. If rays are already converging, gravity focuses them like a magnifying glass.

Remember that gravity is a magnifying glass. This means that rays going in either direction from the light trap are focused and soon collide. That's why we call it a light trap! We can picture this scenario in two dimensions (Fig. 4), with space on the horizontal axis and time in the vertical direction. The trap is represented by two points—the equivalent of a sphere in one dimension—with an inside (dotted black line) and outside (solid black line). The dark rays head outwards and lighter red head inwards, meeting at the stars.

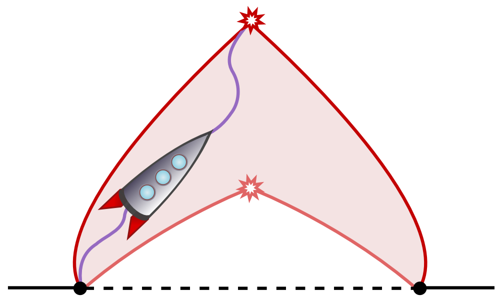

Figure 4. The future of the light trap is all the places a rocket can visit.

Penrose showed carefully that the *future* of the light trap—the points in spacetime that a rocket starting on the trap can visit—is the pink blob in Fig. 4. It is finite, and cannot be extended, so time effectively ends at the top of the blob. To see why, imagine that a light ray is allowed to pass through a collision, as in Fig. 5(a). There will be a point on the extended ray which can be reached by a "zig-zag" light ray path which swaps between rays (Fig. 5(b)). The kink in this zig-zag path can be smoothed out, like pulling on a crimped thread, to form a slower path that a rocket ship could use to arrive at the same point (Fig. 5(c)). And because rocket ships can always be a little slower, or a little faster, this path can be varied to reach a small ball of points in the vicinity (Fig. 5(d)). This zig-zag argument shows that a rocket can visit any point on an extended light ray.

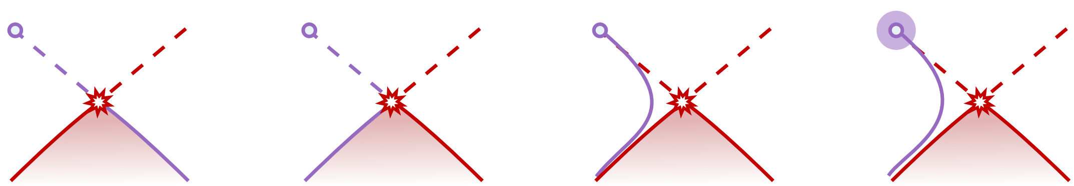

Figure 5. A point on an extended ray can also be reached by a zig-zag path which switches rays at the collision. This zig-zag can be smoothed into a path a rocket could take, then varied slightly to visit a ball of nearby points.

Penrose realized that this was impossible. Since nothing can travel faster than light, and rockets are strictly slower, the *edge* of the future consists of points reached only by light ray. Let's expand on this. For a point on the edge, any small ball we draw around it will partly fall outside the future (Fig. 6). Otherwise, it's not on the edge, but in the middle! If a rocket could reach such a point, we could vary the rocket's speed a little to reach a whole ball of nearby points, completely contained in the future; by definition, this cannot be true for a point on the edge, so only light rays, whose speed cannot be varied, are allowed to get there. The zig-zag argument establishes that any point on an extended light ray can be reached by rocket. It follows that no extended ray is on the edge of the future. Perhaps the rays can be extended, but they will be inside the future, and not on its boundary. Rockets can't go everywhere, however, so the future must have an edge. Since our original light rays can't be extended to provide it, and there are no other light in town, Penrose concluded that the *future itself ends* in a singularity at the top of the blob. So passes the glory of spacetime.

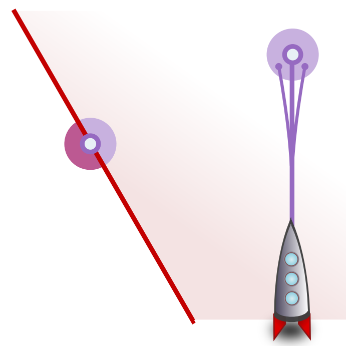

FIGURE 6. On the left, a point on the edge of the future. Any ball around it lies partly outside the accessible region. On the right, a point in the interior of the future, surrounded by a ball of rocket-accessible points.

In contrast to Schwarzschild's perfect sphere, a light trapping surface can be puckered, deformed, and filled with matter; the singularity will form regardless. Evidently, the bugs are not in in Schwarzchild's mathematics, but Einstein's beautifully incomplete theory. Whether singularities are truly bugs in Nature, or are somehow patched up by new physics, remains to be seen. But as Penrose's fellow laureates discovered, black holes—Schwarzschild's hole in Einstein's perfect theory —can be found in our galactic backyard.