Black Holes, Thermodynamics, and Information Loss

Robert M. Wald

Black Holes

Black Holes: A *black hole* is a region of spacetime where gravity is so strong that nothing—not even light—that enters that region can ever escape from it.

Michell (1784); Laplace (1798):

Escape velocity:

$$\frac{1}{2} m v_e^2 = G \frac{mM}{R}$$

so $v_e > c$ if

$$R < R_S \equiv \frac{2GM}{c^2} \approx 3 \frac{M}{M_\odot} \mathrm{km}$$

Michell and Laplace predicted that stars with $R < R_S$ would appear to be black.

Black Holes in General Relativity

In special and general relativity, nothing can travel faster than light, so if light is "pulled back", then so is everything else. A body with $R < \frac{2GM}{c^2}$ cannot exist in equilibrium; it must undergo complete gravitational collapse to a singularity. There is considerable (but mainly indirect!) evidence in favor of the "cosmic censorship conjecture": The end product of this collapse is always a black hole, with the singularity hidden within the black hole.

Formation of Black Holes

There are 3 basic processes by which black holes may have formed in our universe:

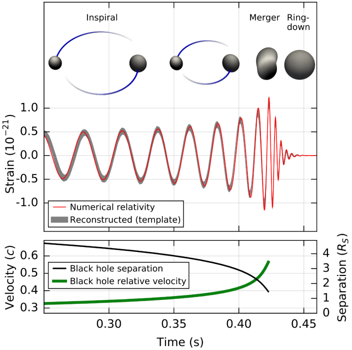
- Complete gravitational collapse of stars that have exhausted their thermonuclear fuel and are too massive to be supported by electron degeneracy pressure (white dwarfs, $M < 1.4 M_\odot$) or neutron degeneracy pressure/nuclear forces (neutron stars, $M \lesssim 2\text{–}3 M_\odot$). Mass range: $2M_\odot \lesssim M \lesssim 100 M_\odot$
- Collapse of the central region of a galaxy or dense star cluster. The existence of extremely high redshift quasars shows that such black holes form extremely early, and it is not obvious how this happens. Mass range: $(10^5 M_\odot \lesssim M \lesssim 10^{10} M_\odot)$
- Primodial black holes that, hypothetically, could have formed by the collapse of over-dense regions in the very early universe. Mass range: anything

Observational Evidence for Black Holes

Considerable observation evidence for black holes has accumulated over the past 50 years. There are ~ 20 known binary X-ray systems that contain a compact component with mass $> 3M_\odot$. Almost all nearby galaxies show clear evidence of a massive black hole at their center. Our own galaxy is one of the best examples: There is a mass concentration of $\sim 4 \times 10^6 M_\odot$ at the center of our galaxy that is within 1000 Schwarzschild radii.

However, the most dramatic observational evidence has come in recent years.

LIGO A signal from a binary black hole merger



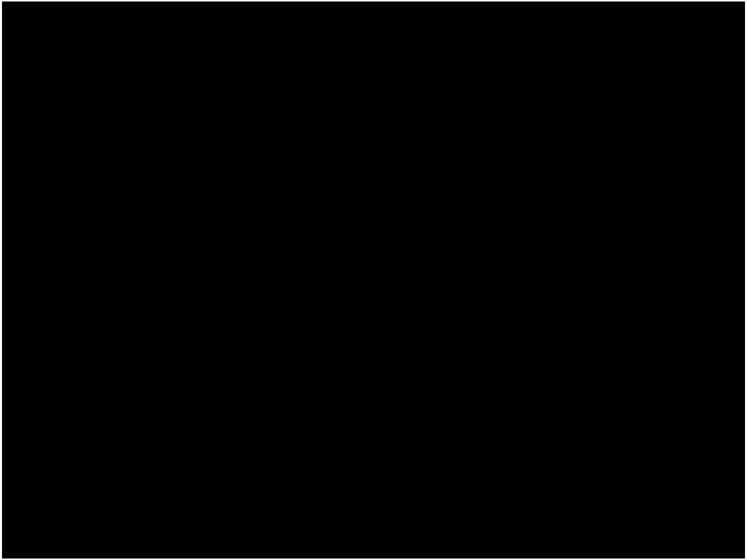
LIGO-G1600258

The GWTC catalog now has 90 events, almost all of which are binary black hole mergers.

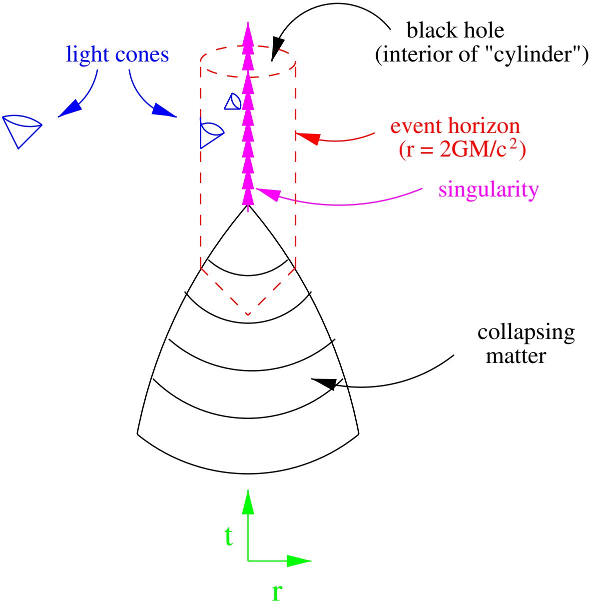
EHT Image of Central Region of M87

This confirms the presence of a $\sim 6 \times 10^9 M_\odot$ black hole at the center of M87.

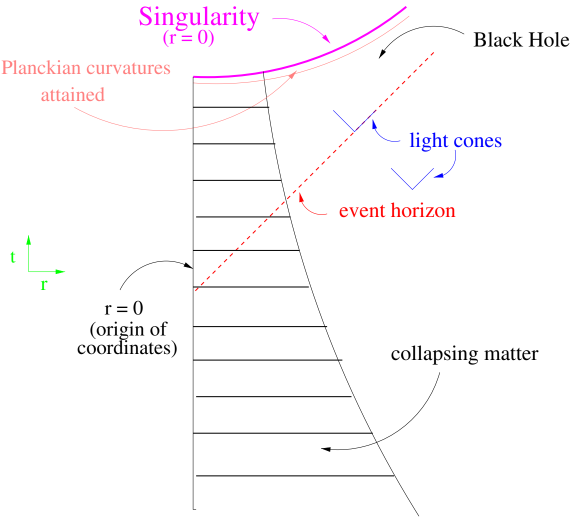
What a Black Hole Looks Like



Spacetime Diagram of Gravitational Collapse



Spacetime Diagram of Gravitational Collapse with Angular Directions Suppressed and Light Cones "Straightened Out"



Black Holes and Thermodynamics

Stationary black hole $\leftrightarrow$ Body in thermal equilibrium

Consider an ordinary system composed of a large number of particles, such as the gas in a box. If one waits long enough after one has filled a box with gas, the gas will "settle down" to final state of thermal equilibrium, characterized by a small number of "state parameters", such as the total energy, E, the total volume, V, and the total number of particles, N. Similarly, if one forms a black hole by gravitational collapse, it is expected that the black hole will quickly "settle down" to a stationary final state. This final state is uniquely characterized by by its total mass, M, total angular momentum, J, and total electric charge, Q.

0th Law

Thermodynamics: The temperature, T, is constant over a body in thermal equilibrium.

Black holes: The surface gravity, κ, is constant over the horizon of a stationary black hole. (κ is the limit as one approaches the horizon of the acceleration needed to remain stationary times the "redshift factor".)

1st and 2nd Laws

1st Law

Thermodynamics:

$$\delta E = T\delta S - P\delta V$$

Black holes:

$$\delta M = \frac{1}{8\pi}\kappa\delta A + \Omega_H \delta J + \Phi_H \delta Q$$

2nd Law

Thermodynamics:

$$\delta S \geq 0$$

Black holes:

$$\delta A \geq 0$$

Analogous Quantities

$M \leftrightarrow E \leftarrow$ But M really is E!

$\frac{1}{2\pi}\kappa \leftrightarrow T$

$\frac{1}{4}A \leftrightarrow S$

Particle Creation by Black Holes

Black holes are perfect black bodies! As Stephen Hawking discovered in 1974, as a result of particle creation effects in quantum field theory, a distant observer will see an exactly thermal flux of all species of particles appearing to emanate from the black hole. The temperature of this radiation is

$$kT = \frac{\hbar\kappa}{2\pi}\,.$$

For a Schwarzshild black hole $(J = Q = 0)$ we have $\kappa = c^3/4GM$, so

$$T \sim 10^{-7}\frac{M_\odot}{M}\,{}^\circ\mathrm{K}\,.$$

Black Hole Evaporation

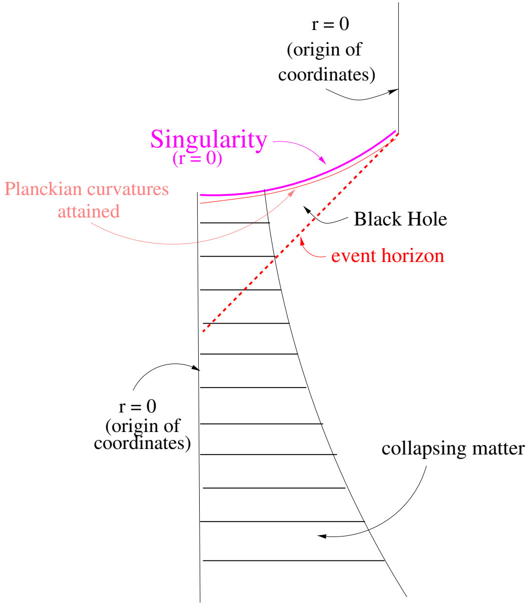
The mass loss of a black hole due to the particle creation process is

$$\frac{dM}{dt} \sim AT^4 \propto M^2 \frac{1}{M^4} = \frac{1}{M^2}\,.$$

Thus, an isolated black hole should "evaporate" completely in a time

$$\tau \sim 10^{73} (\frac{M}{M_\odot})^3 \mathrm{sec}\,.$$

Spacetime Diagram of Evaporating Black Hole



Analogous Quantities

$M \leftrightarrow E \leftarrow$ But M really is E!

$\frac{1}{2\pi}\kappa \leftrightarrow T \leftarrow$ But $\kappa/2\pi$ really is the (Hawking) temperature of a black hole!

$\frac{1}{4}A \leftrightarrow S$

Problems with the 2nd Law

Ordinary 2nd law: $\delta S \geq 0$

Classical black hole area theorem: $\delta A \geq 0$

However, when a black hole is present, it really is physically meaningful to consider only the matter outside the black hole. But then, can decrease S by dropping matter into the black hole. So, can get $\delta S < 0$.

Although classically A never decreases, it *does* decrease during the quantum particle creation process. So, can get $\delta A < 0$.

The Generalized 2nd Law

The problems with both second laws are resolved if, as first suggested by Bekenstein (prior to Hawking's particle creation calculation!) we have only one 2nd law:

$$\delta S' \geq 0$$

where

$$S' \equiv S + \frac{1}{4}\frac{c^3}{G\hbar}A$$

where S = entropy of matter outside black holes and A = black hole area.

A careful analysis of gedanken experiments strongly suggests that the generalized 2nd law is valid!

Analogous Quantities

$M \leftrightarrow E \leftarrow$ But M really is E!

$\frac{1}{2\pi}\kappa \leftrightarrow T \leftarrow$ But $\kappa/2\pi$ really is the (Hawking) temperature of a black hole!

$\frac{1}{4}A \leftrightarrow S \leftarrow$ Apparent validity of the generalized 2nd law strongly suggests that $A/4$ really is the physical entropy of a black hole!

Does the Entropy of a Black Hole Count States?

If so, where do these states "reside"?

Possible answers:

- Inside the black hole. But it is hard to see how the number of states inside the black hole influence thermodynamic properties outside of the black hole. It is also hard to see how the number of states inside the black hole are "lost" during black hole evaporation. It is also hard to see how the simple formula $S = A/4$ could emerge from counting states in the interior of the black hole.
- On the event horizon. The event horizon is a globally defined quantity. How does a local region of spacetime on the event horizon "know" that it is on the event horizon so that its degrees of freedom should count towards black hole entropy?

- In the "thermal atmosphere" surrounding the black hole. In this view, the entropy of a black hole would reside entirely in the thermal distribution of Hawking radiation as seen by stationary observers (but not inertial observers) outside the black hole. Since $T \to \infty$ as one approaches the horizon, the total entropy diverges, but if one puts in a Planck scale cutoff, one will get $S \propto A$. However, it is hard to see why $S = A/4$ for all black holes in general relativity, independently, e.g., of the number of species of matter. Also, in other theories of gravity, the temperature distribution of the thermal atmosphere is determined by the surface gravity and redshift in the same way as in general relativity, but the formula for black hole entropy is quite different, which does not seem reasonable if black hole entropy corresponds to counting states in the thermal atmosphere.

Quantum Entanglement

If a quantum system consists of two subsystems, described by Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, then the joint system is described by the Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$. In addition to simple product states $|\Psi_1\rangle \otimes |\Psi_2\rangle$, the Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$ contains linear combinations of such product states that cannot be re-expressed as a simple product. If the state of the joint system is not a simple product, the subsystems are said to be *entangled* and the state of each subsystem is said to be *mixed*. Interactions between subsystems generically result in entanglement. Entanglement is a ubiquitous feature of quantum field theory. At small spacelike separations, a quantum field is always strongly entangled with itself, as illustrated by the following formula for a massless KG field in Minkowski spacetime:

$$\langle 0|\phi(x)\phi(y)|0\rangle = \frac{1}{4\pi^2}\frac{1}{\sigma(x,y)}$$

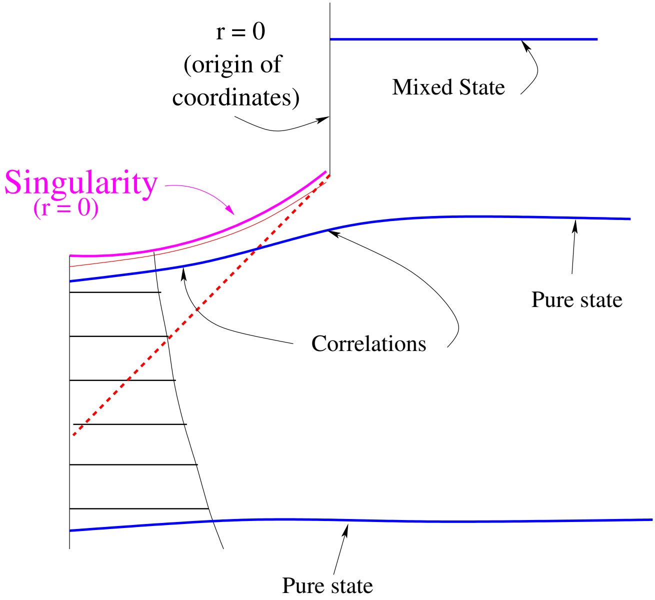
(If no entanglement, $\langle 0|\phi(x)\phi(y)|0\rangle = \langle 0|\phi(x)|0\rangle\langle 0|\phi(y)|0\rangle = 0$.)

Information Loss

In a spacetime in which a black hole forms, there will be entanglement between the state of quantum field observables inside and outside of the back hole. This entanglement is intimately related to the Hawking radiation emitted by the black hole. In addition to the strong quantum field entanglement arising on small scales near the horizon associated with Hawking radiation, there may also be considerable additional entanglement because the matter that forms (or later falls into) the black hole may be highly entangled with matter that remains outside of the black hole.

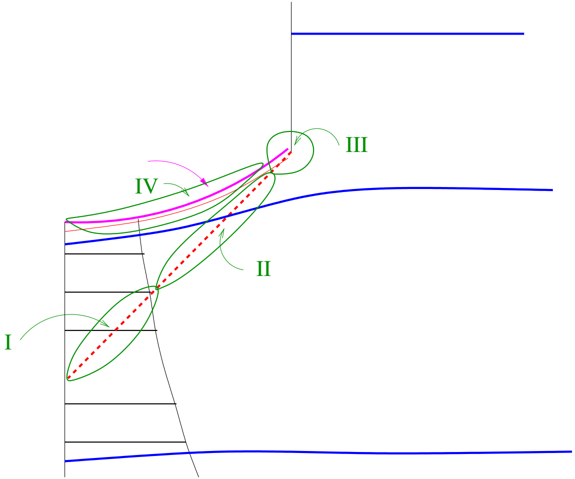
In a semiclassical treatment, if the black hole evaporates completely, the final state will be mixed, i.e., one will have dynamical evolution from a pure state to a mixed state. In this sense, there will be irreversible "information loss" into black holes.

Spacetime Diagram Illustratrating Information Loss



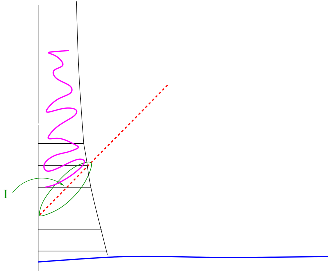
What's Wrong With This Picture?

If the semiclassical picture is wrong, there are basically 4 places where it could be wrong in such a way as to modify the conclusion of information loss:



Possibility I: No Black Hole Ever Forms (Fuzzballs, Gravistars)

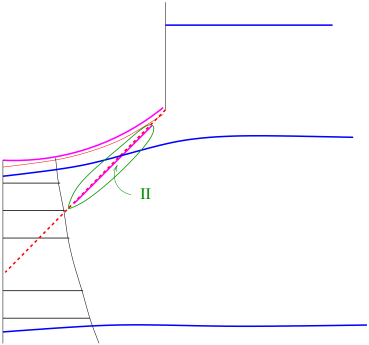
This is quite a radical alternative. Both (semi-)classical general relativity and quantum field theory would have to break down in an arbitrarily low curvature/low energy regime.



Note that if the fuzzball or other structure doesn't form at just the right moment, it will be "too late" to do anything without a major violation of causality/locality in a low curvature regime as well.

Possibility II: Major Departures from Semiclassical Theory Occur During Evaporation (Firewalls, Tunneling via Wormholes)

This is also a radical alternative, since the destruction of entanglement between the inside and outside of the black hole during evaporation requires a breakdown of quantum field theory in an arbitrarily low curvature regime or a major violation of causality.



Firewalls

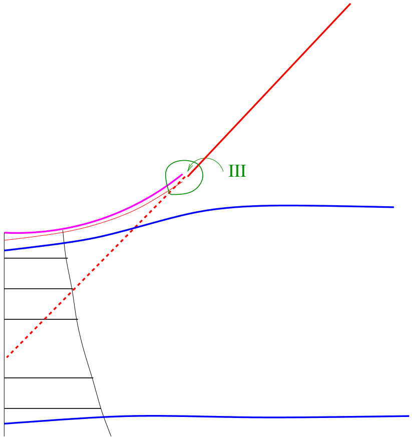
"Firewalls" would need to come into existence at (or very near) the horizon in order to destroy entanglement. There is no theory of firewalls, but they would not only require a major breakdown of local laws of physics near the horizon but also require major violations of causality/locality in order to bring the entanglement from deep inside the black hole to outside the horizon.

Tunneling via Wormholes

The proposal here is that the information deep inside the black hole tunnels to infinity, thereby purifying the state of Hawking radiation. This would involve a drastic violation of causality. This proposal receives support from the "island formula" for the entropy of the Hawking radiation obtained from "replica wormhole" calculations. However, there is no calculation of how the state of the Hawking radiation is modified and no (plausible) description of a mechanism for how this works for black holes but doesn't result in causality violation in other contexts. (Quote from a recent review by Almheiri, Hartman, Maldacena, Shaghoulian, and Tajdini: "The idea is that very complex computations in the radiation can create wormholes that reach into that interior and pull out the information stored there.")

Possibility III: Remnants

This is not a radical alternative, since the breakdown of the semi-classical picture occurs only near the Planck scale.

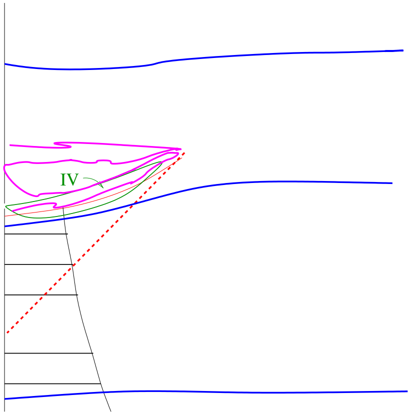


Remnants

However, there are severe problems with invoking remnants to maintain a pure state. If the remnants cannot interact with the external world, it is not clear what "good" they do (since the "information," although still present, is inaccessible). If they can interact with the external world, then there are serious thermodynamic problems with them, since they must contain arbitrarily many states at tiny (Planck scale) energy and thus should be thermodynamically favored over all other forms of matter.

Possibility IV: A Final Burst

This alternative requires an arbitrarily large amount of "information" to be released from an object of Planck mass and size.



Final Burst Carrying All the Information?

This is not necessarily as crazy as it might initially sound: Hotta, Schutzhold, and Unruh have considered the model of an accelerating mirror in $1+1$ spacetime dimensions that emits Hawking-like radiation. The "partner particles" to the Hawking radiation are indistinguishable from vacuum fluctuations, and thus the information is "carried off" by vacuum fluctuations that are correlated with the emitted particles—at no energy cost!

However, I analyzed this several years ago and showed that in higher than $1+1$ dimensions it is not possible to perform a similar entanglement with vacuum fluctuations emanating from a small spatial region without a very large auxilliary particle and energy cost. Thus, this does not appear to be a viable option.

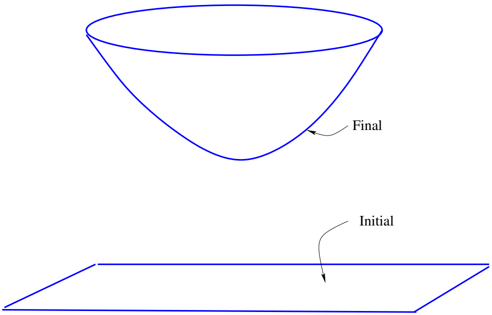
Arguments Against Information Loss: Violation of Unitarity

In scattering theory, the word "unitarity" has 2 completely different meanings: (1) Conservation of probability; (2) Evolution from pure states to pure states.

Failure of (1) would represent a serious breakdown of quantum theory (and, indeed, of elementary logic). However, that is not what is being proposed by the semiclassical picture.

Failure of (2) would be expected to occur in any situation where the final "time" is not a Cauchy surface, and it is entirely innocuous.

Violation of Unitarity



For example, we get "pure $\rightarrow$ mixed" for the evolution of a massless Klein-Gordon field in Minkowski spacetime if the final "time" is chosen to be a hyperboloid. This is a *prediction* of quantum theory, not a *violation* of quantum theory.
The "pure $\rightarrow$ mixed" evolution predicted by the semiclassical analysis of black hole evaporation is of an entirely similar character.

Arguments Against Information Loss: Failure of Energy and Momentum Conservation

Banks, Peskin, and Susskind argued that evolution laws taking "pure $\rightarrow$ mixed" would lead to violations of energy and momentum conservation. However, they considered only a "Markovian" type of evolution law (namely, the Lindblad equation). This would not be an appropriate model for black hole evaporation, as the black hole clearly should retain a "memory" of what energy it previously emitted.

Unruh has provided a very nice example of a quantum mechanical system that interacts with a "hidden spin system" in such a way that "pure $\rightarrow$ mixed" for the quantum system but exact energy conservation holds. Thus, there is no problem of principle with maintaining exact energy and momentum conservation in quantum mechanics with an evolution wherein "pure $\rightarrow$ mixed."

Arguments Against Information Loss: AdS/CFT

The one sentence version of AdS/CFT argument against the semiclassical picture is simply that if gravity in asymptotically AdS spacetimes is dual to a conformal field theory, then since the conformal field theory does not admit "pure $\rightarrow$ mixed" evolution, such evolution must also not be possible in quantum gravity.

AdS/CFT is a conjecture. The difficulty I have with the use of the AdS/CFT correspondence in arguments against information loss is not that this conjecture has not been *proven*, but rather that it has not been *formulated* with the degree of precision needed to use it reliably in such an argument. For example, if some of the late time degrees of freedom of the CFT correspond to early time degrees of freedom in the bulk that propagate into the black hole, there would be no contradiction with effective "information loss" in the bulk.

Conclusions

The study of black holes has led to the discovery of a remarkable and deep connection between gravitation, quantum theory, and thermodynamics. It is my hope and expectation that further investigations of black holes will lead to additional fundamental insights.