Black Holes and Thermodynamics

Robert M. Wald

I. Classical Black Holes

II. The First Law of Black Hole Mechanics

III. Quantum Black Holes, the Generalized 2nd Law and the 'Information Paradox'

Black Holes and Thermodynamics

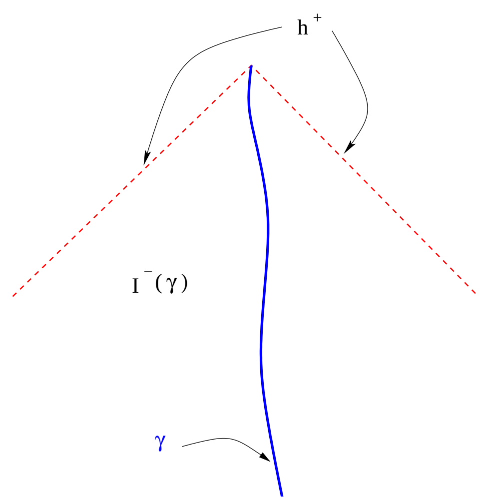
I: Classical Black Holes

Robert M. Wald

General references: R.M. Wald *General Relativity* University of Chicago Press (Chicago, 1984); R.M. Wald Living Rev. Rel. **4**, 6 (2001).

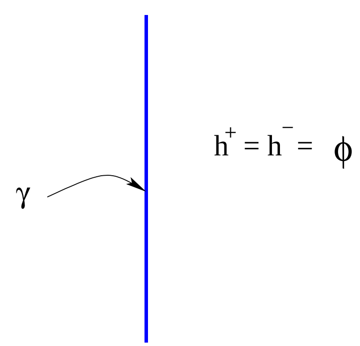
Horizons

An observer in a spacetime (M, g_{ab}) is represented by an inextendible timelike curve γ. Let $I^-(\gamma)$ denote the chronological past of γ. The future horizon, h^+, of γ is defined to be the boundary, $\dot{I}^-(\gamma)$ of $I^-(\gamma)$.

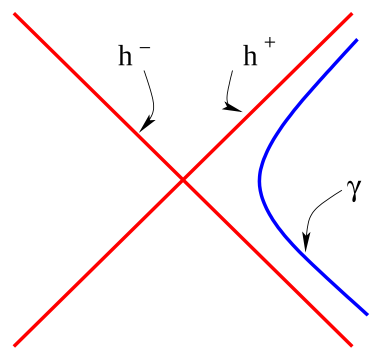


Theorem: Each point $p \in h^+$ lies on a null geodesic segment contained entirely within h^+ that is future inextendible. Furthermore, the convergence of these null geodesics that generate h^+ cannot become infinite at a point on h^+.

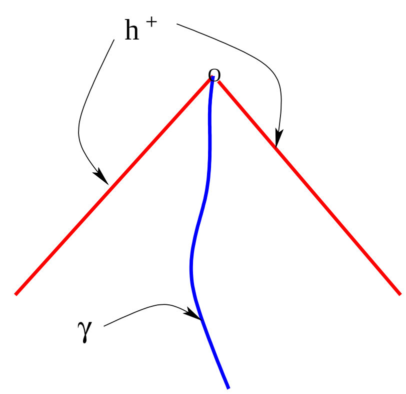
Can similarly define a past horizon, h^-. Can also define h^+ and h^- for families of observers.



Inertial Observer in Minkowski Spacetime



Uniformly Accelerating Observer in Minkowski Spacetime



Observer in Minkowski Spacetime with a point removed

Black Holes and Event Horizons

Consider an asymptotically flat spacetime (M, g_{ab}). (The notion of asymptotic flatness can be defined precisely using the notion of conformal null infinity.) Consider the family of observers Γ who escape to arbitrarily large distances at late times. If the past of these observers $I^-(\Gamma)$ fails to be the entire spacetime, then a black hole $B \equiv M - I^-(\Gamma)$ is said to be present. The horizon, h^+, of these observers is called the future event horizon of the black hole.

This definition allows "naked singularities" to be present.

Cosmic Censorship

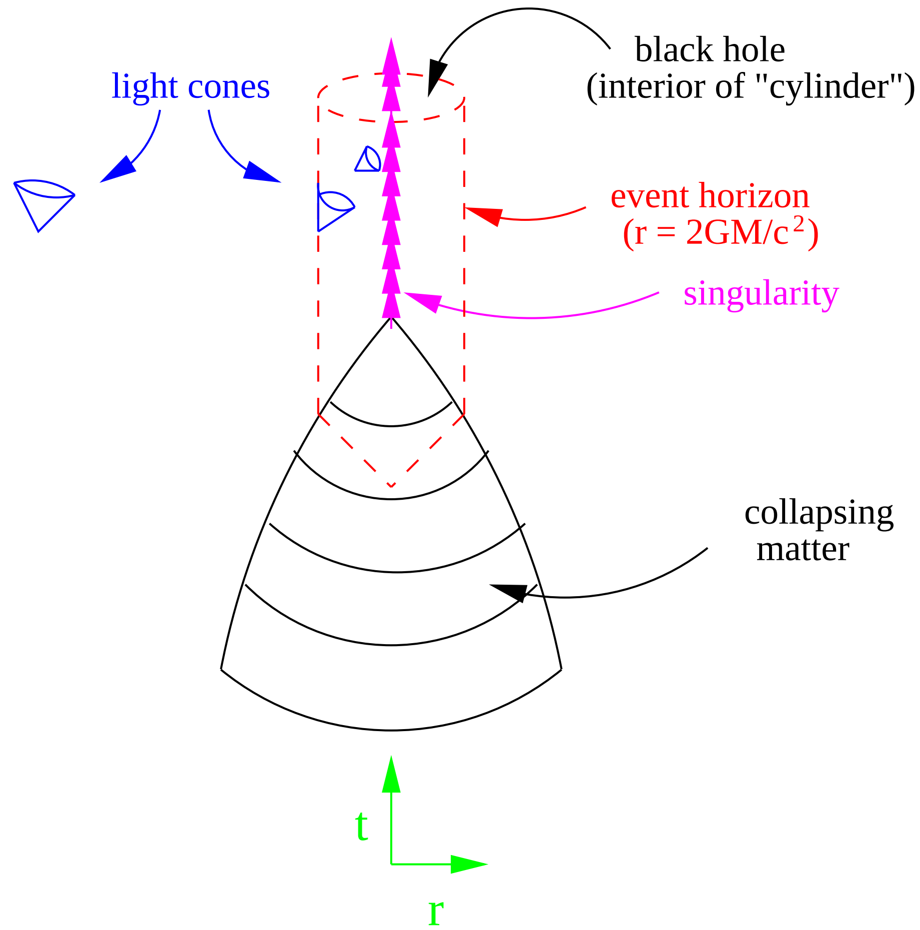
A Cauchy surface, $\mathcal{C}$, in a (time orientable) spacetime (M, g_{ab}) is a set with the property that every inextendible timelike curve in M intersects $\mathcal{C}$ in precisely one point. (M, g_{ab}) is said to be globally hyperbolic if it possesses a Cauchy surface $\mathcal{C}$. This implies that M has topology $\mathbf{R} \times \mathcal{C}$.

An asymptotically flat spacetime (M, g_{ab}) possessing a black hole is said to be predictable if there exists a region of M containing the entire exterior region and the event horizon, h^+, that is globally hyperbolic. This expresses the idea that no "naked singularities" are present.

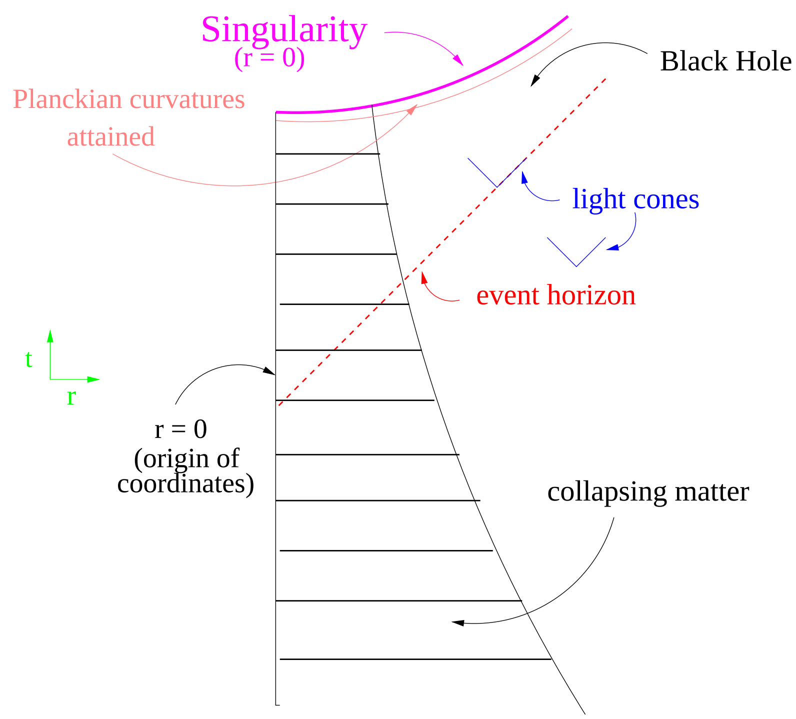
Cosmic Censor Hypothesis: The maximal Cauchy evolution—which is automatically globally hyperbolic—of an asymptotically flat initial data set (with suitable matter fields) generically yields an asymptotically flat spacetime with complete null infinity.

The validity of the cosmic censor hypothesis would assure that any observer who stays outside of black holes could not be causally influenced by singularities.

Spacetime Diagram of Gravitational Collapse



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



Null Geodesics and the Raychauduri Equation

For a congruence of null geodesics with affine parameter λ and null tangent k^a, define the expansion, θ, by

$$\theta = \nabla_a k^a$$

The area, A of an infinitesimal area element transported along the null geodesics varies as

$$\frac{d(\ln A)}{d\lambda} = \theta$$

For null geodesics that generate a null hypersurface (such as the event horizon of a black hole), the twist, ω_{ab}, vanishes. The Raychauduri equation—which is a direct consequence of the geodesic deviation equation—then yields

$$\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - \sigma_{ab}\sigma^{ab} - R_{ab}k^a k^b$$

where σ_{ab} is the shear of the congruence. Thus, provided that $R_{ab}k^a k^b \geq 0$ (i.e., the null energy condition holds), we have

$$\frac{d\theta}{d\lambda} \leq -\frac{1}{2}\theta^2$$

which implies

$$\frac{1}{\theta(\lambda)} \leq \frac{1}{\theta_0} + \frac{1}{2}\lambda$$

Consequently, if $\theta_0 < 0$, then $\theta(\lambda_1) = -\infty$ at some $\lambda_1 < 2/|\theta_0|$ (provided that the geodesic can be extended that far).

The Area Theorem

Any horizon h^+, is generated by future inextendible null geodesics; cannot have $\theta = -\infty$ at any point of h^+. Thus, if the horizon generators are complete, must have $\theta \geq 0$. However, for a predictable black hole, can show that $\theta \geq 0$ without having to assume that the generators of the event horizon are future complete—by a clever argument involving deforming the horizon outwards at a point where $\theta < 0$.

Let S_1 be a Cauchy surface for the globally hyperbolic region appearing in the definition of predictable black hole. Let S_2 be another Cauchy surface lying to the future of S_1. Since the generators of h^+ are future complete, all of the generators of h^+ at S_1 also are present at S_2. Since $\theta \geq 0$, it follows that the area carried by the generators of h^+ at S_2 is greater or equal to $A[S_1 \cap h^+]$. In addition, new horizon generators may be present at S_2. Thus, $A[S_2 \cap h^+] \geq A[S_1 \cap h^+]$, i.e., we have the following theorem:

Area Theorem: For a predictable black hole with $R_{ab}k^a k^b \geq 0$, the surface area A of the event horizon h^+ never decreases with time.

Killing Vector Fields

An isometry is a diffeomorphism ("coordinate transformation") that leaves the metric, g_{ab} invariant. A Killing vector field, ξ^a, is the infinitesimal generator of a one-parameter group of isometries. It satisfies

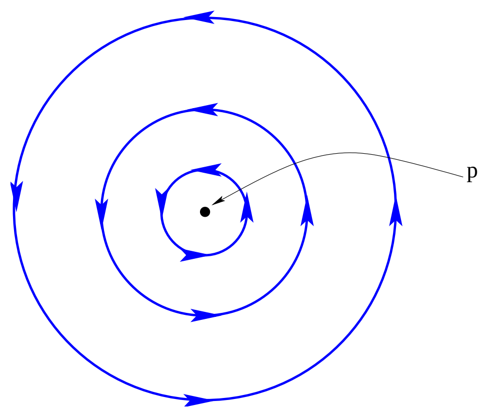
$$0 = \mathcal{L}_\xi g_{ab} = 2\nabla_{(a}\xi_{b)}$$

For a Killing field ξ^a, let $F_{ab} = \nabla_a \xi_b = \nabla_{[a}\xi_{b]}$. Then ξ^a is uniquely determined by its value and the value of F_{ab} at an aribitrarily chosen single point p.

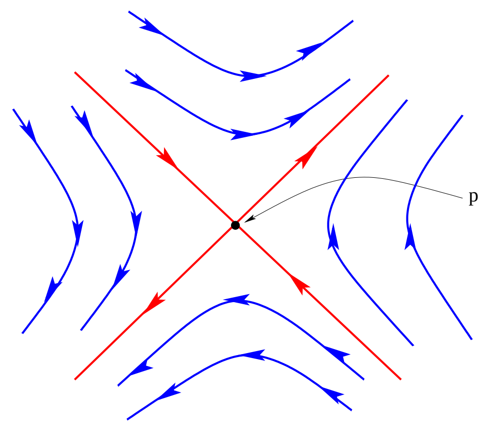
Bifurcate Killing Horizons

2-dimensions: Suppose a Killing field ξ^a vanishes at a point p. Then ξ^a is determined by F_{ab} at p. In 2-dimensions, $F_{ab} = \propto \epsilon_{ab}$, so ξ^a is unique up to scaling

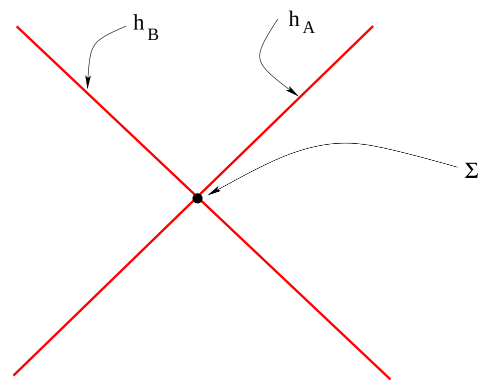
If g_{ab} is Riemannian, the orbits of the isometries generated by ξ^a must be closed and, near p, the orbit structure is like a rotation in flat space:



Similarly, if g_{ab} is Lorentzian, the isometries must carry the null geodesics through p into themselves and, near p, the orbit structure is like a Lorentz boost in 2-dimensional Minkowski spacetime:



4-dimensions: Similar results to the 2-dimensional case hold if ξ^a vanishes on a 2-dimensional surface Σ. In particular, if g_{ab} is Lorentzian and Σ is spacelike, then, near Σ, the orbit structure of ξ^a will look like a Lorentz boost in 4-dimensional Minkowski spacetime. The pair of intersecting (at Σ) null surfaces h_A and h_B generated by the null geodesics orthogonal to Σ is called a bifurcate Killing horizon.



It follows that ξ^a is normal to both h_A and h_B. More generally, any null surface h having the property that a Killing field is normal to it is called a Killing horizon.

Surface Gravity and the Zeroth Law

Let h be a Killing horizon associated with Killing field ξ^a. Let U denote an affine parameterization of the null geodesic generators of h and let k^a denote the corresponding tangent. Since ξ^a is normal to h, we have

$$\xi^a = f k^a$$

where $f = \partial U/\partial u$ where u denotes the Killing parameter along the null generators of h. Define the surface gravity, κ, of h by

$$\kappa = \xi^a \nabla_a \ln f = \partial \ln f/\partial u$$

Equivalently, we have $\xi^b \nabla_b \xi^a = \kappa \xi^a$ on h. It follows immediately that κ is constant along each generator of h.

Consequently, the relationship between affine parameter U and Killing parameter u on an arbitrary Killing horizon is given by

$$U = \exp(\kappa u)$$

Can also show that

$$\kappa = \lim_{h}(Va)$$

where $V \equiv [-\xi^a \xi_a]^{1/2}$ is the "redshift factor" and a is the proper acceleration of observers following orbits of ξ^a.

In general, κ can vary from generator to generator of h. However, we have the following three theorems:

Zeroth Law (1st version): Let h be a (connected) Killing horizon in a spacetime in which Einstein's equation holds with matter satisfying the dominant energy condition. Then κ is constant on h.

Zeroth Law (2nd version): Let h be a (connected) Killing horizon. Suppose that either (i) ξ^a is hypersurface orthogonal (static case) or (ii) there exists a second Killing field ψ^a which commutes with ξ^a and satisfies $\nabla_a(\psi^b \omega_b) = 0$ on h, where ω_a is the twist of ξ^a (stationary-axisymmetric case with "t-ϕ reflection symmetry"). Then κ is constant on h.

Zeroth Law (3rd version): Let h_A and h_B be the two null surfaces comprising a (connected) bifurcate Killing horizon. Then κ is constant on h_A and h_B.

Constancy of κ and Bifurcate Killing Horizons

As just stated, κ is constant over a bifurcate Killing horizon. Conversely, it can be shown that if κ is constant and non-zero over a Killing horizon h, then h can be extended locally (if necessary) so that it is one of the null surfaces (i.e., h_A or h_B) of a bifurcate Killing horizon.

In view of the first version of the 0th law, we see that apart from "degenerate horizons" (i.e., horizons with $\kappa = 0$), bifurcate horizons should be the only types of Killing horizons relevant to general relativity.

Event Horizons and Killing Horizons

Hawking Rigidity Theorem: Let (M, g_{ab}) be a stationary, asymptotically flat solution of Einstein's equation (with matter satisfying suitable hyperbolic equations) that contains a black hole. Then the event horizon, h^+, of the black hole is a Killing horizon.

The stationary Killing field, ξ^a, must be tangent to h^+. If ξ^a is normal to h^+ (so that h^+ is a Killing horizon of ξ^a), then it can be shown that ξ^a is hypersurface orhogonal, i.e., the spacetime is static. If ξ^a is not normal to h^+, then there must exist another Killing field, χ^a , that is normal to the horizon. It can then be further shown that there is a linear combination, ψ^a, of ξ^a and χ^a whose orbits are spacelike and closed, i.e., the spacetime is axisymmetric. Thus, a stationary black hole must be static or axisymmetric.

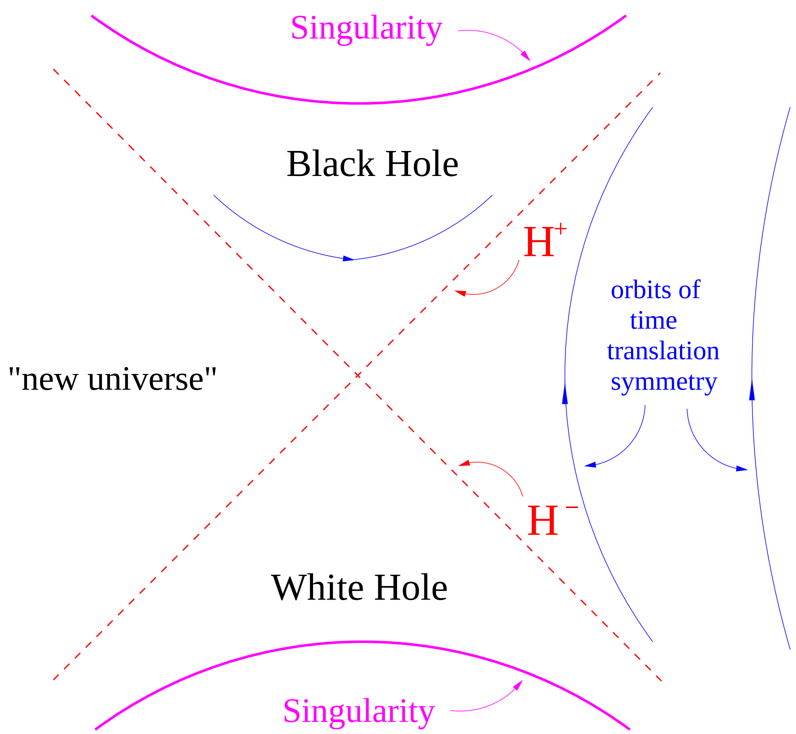
We can choose the normalization of χ^a so that

$$\chi^a = \xi^a + \Omega \psi^a$$

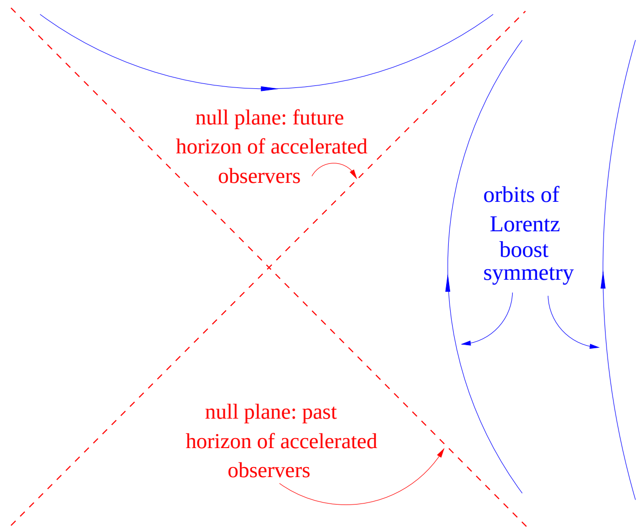
where Ω is a constant, called the angular velocity of the horizon.

Idealized ("Analytically Continued") Black Hole

"Equilibrium State"



A Close Analog: Lorentz Boosts in Minkowski Spacetime



Note: For a black hole with $M \sim 10^9 M_\odot$, the curvature at the horizon of the black hole is smaller than the curvature in this room! An observer falling into such a black hole would hardly be able to tell from local measurements that he/she is not in Minkowski spacetime.

Summary

- If cosmic censorship holds, then—starting with nonsingular initial conditions—gravitational collapse will result in a predictable black hole.
- The surface area of the event horizon of a black hole will be non-decreasing with time (2nd law).

It is natural to expect that, once formed, a black hole will quickly asymptotically approach a stationary ("equilibrium") final state. The event horizon of this stationary final state black hole:

- will be a Killing horizon
- will have constant surface gravity, κ (0th law)

- if $\kappa \neq 0$, will have bifurcate Killing horizon structure

Black Holes and Thermodynamics

II: The First Law of Black Hole Mechanics

Robert M. Wald

Based mainly on V. Iyer and RMW, Phys. Rev. **D50**, 846 (1994)

Variational Formulas

Lagrangian for vacuum general relativity:

$$L_{a_1 \dots a_D} = \frac{1}{16\pi} R \, \epsilon_{a_1 \dots a_D} \, .$$

First variation:

$$\delta L = E \cdot \delta g + d\theta \, ,$$

with

$$\theta_{a_1 \dots a_{d-1}} = \frac{1}{16\pi} g^{ac} g^{bd} (\nabla_d \delta g_{bc} - \nabla_c \delta g_{bd}) \epsilon_{c a_1 \dots a_{d-1}} \, .$$

Symplectic current ($(D-1)$-form):

$$\omega(g; \delta_1 g, \delta_2 g) = \delta_1 \theta(g; \delta_2 g) - \delta_2 \theta(g; \delta_1 g) \, .$$

Symplectic form:

$$\begin{aligned} W_\Sigma(g;\delta_1 g,\delta_2 g) &\equiv \int_\Sigma \omega(g;\delta_1 g,\delta_2 g) \\ &= -\frac{1}{32\pi}\int_\Sigma \left(\delta_1 h_{ab}\delta_2 p^{ab} - \delta_2 h_{ab}\delta_1 p^{ab}\right), \end{aligned}$$

with

$$p^{ab} \equiv h^{1/2}(K^{ab} - h^{ab}K)\,.$$

Noether current:

$$\begin{aligned} \mathcal{J}_X &\equiv \theta(g,\pounds_X g) - X\cdot L \\ &= X\cdot C + dQ_X\,. \end{aligned}$$

Fundamental variational identity:

$$\begin{aligned}\omega(g;\delta g,\pounds_X g) &= X\cdot[E(g)\cdot\delta g]+X\cdot\delta C\\ &\quad +d\left[\delta Q_X(g)-X\cdot\theta(g;\delta g)\right]\end{aligned}$$

ADM conserved quantities:

$$\delta H_X=\int_\infty[\delta Q_X(g)-X\cdot\theta(g;\delta g)]$$

For a stationary black hole, choose X to be the horizon Killing field

$$K^a=t^a+\sum\Omega_i\phi_i^a$$

Integration of the fundamental identity yields the first law of black hole mechanics:

$$0 = \delta M - \sum_i \Omega_i \delta J_i - \frac{\kappa}{8\pi}\delta A\,.$$

Black Holes and Thermodynamics

Stationary black hole $\leftrightarrow$ Body in thermal equilibrium

Just as bodies in thermal equilibrium are normally characterized by a small number of "state parameters" (such as E and V) a stationary black hole is uniquely characterized by M, J, Q.

0th Law

Black holes: The surface gravity, κ, is constant over the horizon of a stationary black hole.

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

1st Law

Black holes:

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

Thermodynamics:

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

2nd Law

Black holes:

$$\delta A \geq 0$$

Thermodynamics:

$$\delta S \geq 0$$

Analogous Quantities

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

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

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

Black Holes and Thermodynamics

III: Quantum Black Holes, the Generalized 2nd Law and the ‘Information Paradox’

Robert M. Wald

General reference: R.M. Wald *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics* University of Chicago Press (Chicago, 1994).

Particle Creation by Black Holes

Black holes are perfect black bodies! 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

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

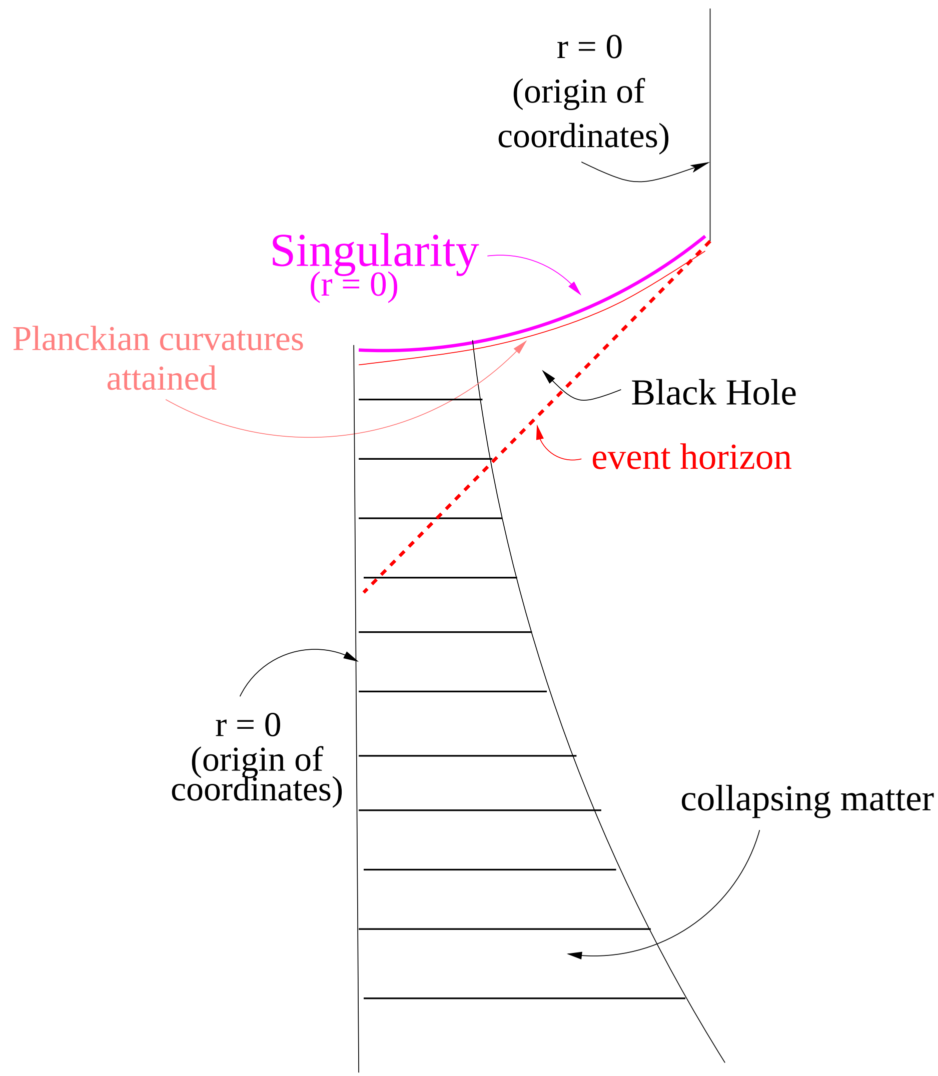
The mass loss of a black hole due to this process is

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

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

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

Spacetime Diagram of Evaporating Black Hole



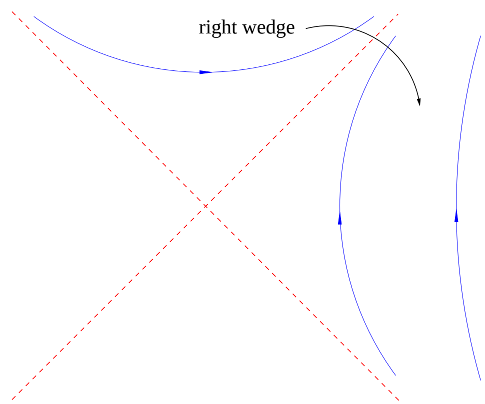
Analogous Quantities

$M \leftrightarrow E \leftarrow$ 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!

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

A Closely Related Phenomenon: The Unruh Effect



View the "right wedge" of Minkowski spacetime as a spacetime in its own right, with Lorentz boosts defining a notion of "time translation symmetry". Then, when restricted to the right wedge, the ordinary Minkowski vacuum state, $|0\rangle$, is a thermal state with respect to this notion of time translations (Bisognano-Wichmann theorem). A uniformly accelerating observer "feels himself to be in a thermal bath at temperature

$$kT = \frac{\hbar a}{2\pi c}$$

(i.e., in SI units, $T \sim 10^{-23} a$).

For a black hole, the temperature locally measured by a stationary observer is

$$kT = \frac{\hbar \kappa}{2\pi V c}$$

where $V = (-\xi^a \xi_a)^{1/2}$ is the redshift factor associated with the horizon Killing field. Thus, for an observer near the horizon, $kT \to \hbar a / 2\pi c$.

The Generalized Second 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 $\delta A < 0$.

However, as first suggested by Bekenstein, perhaps have

$$\delta S' \geq 0$$

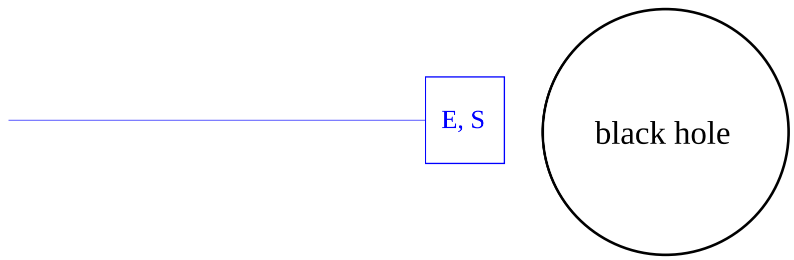
where

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

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

Can the Generalized 2nd Law be Violated?

Slowly lower a box with (locally measured) energy E and entropy S into a black hole.



Lose entropy S

Gain black hole entropy $\delta(\frac{1}{4}A) = \frac{\mathcal{E}}{T_{\rm b.h.}}$

But, classically, $\mathcal{E} = VE \to 0$ as the "dropping point" approaches the horizon, where V is the redshift factor. Thus, apparently can get $\delta S' = -S + \delta(\frac{1}{4}A) < 0$.

However: The temperature of the "acceleration radiation" felt by the box varies as

$$T_{\rm loc} = \frac{T_{\rm b.h.}}{V} = \frac{\kappa}{2\pi V}$$

and this gives rise to a "buoyancy force" which produces a quantum correction to $\mathcal{E}$ that is precisely sufficient to prevent a violation of the generalized 2nd law!

Analogous Quantities

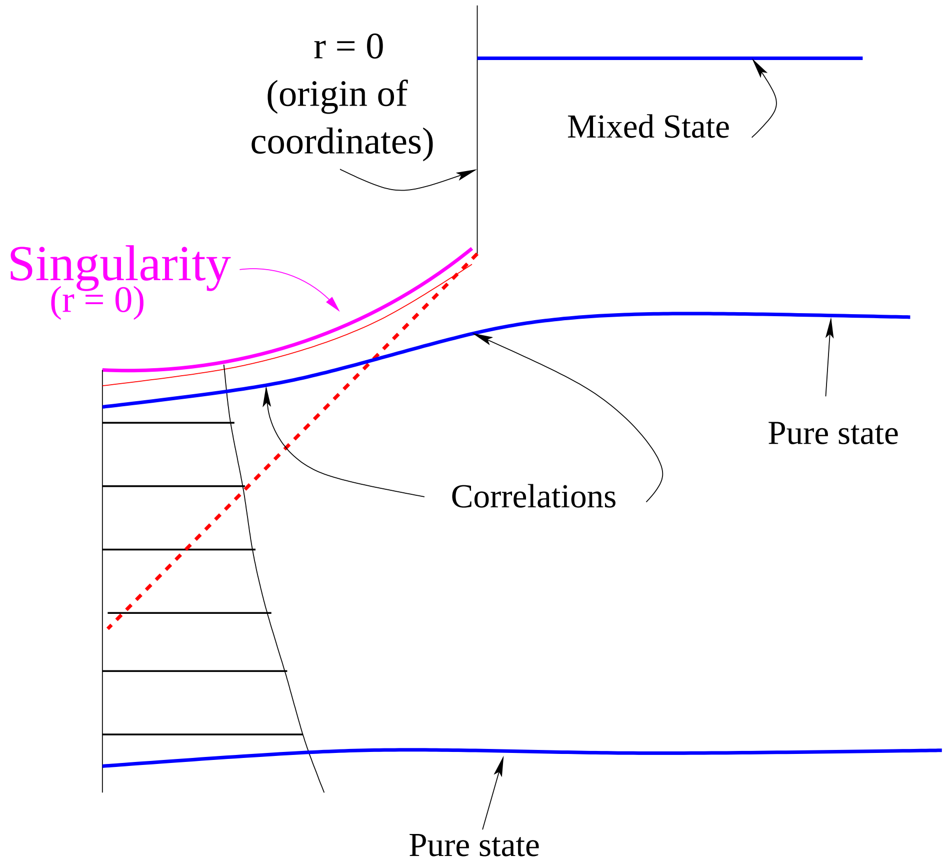
$M \leftrightarrow E \leftarrow$ But M really is E!

$\frac{1}{2\pi}\kappa \leftrightarrow \; T \leftarrow$ 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!

The Information "Paradox"

In a semiclassical description of the Hawking effect, correlations are build up over time between the state of the field outside and inside 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.

Logical possibilities (assuming black holes do form!):

- (i) Semiclassical picture is correct in its essential features and pure $\rightarrow$ mixed.
- (ii) Remnants remain behind.
- (iii) Correlations are restored either (a) in a final "burst" or (b) gradually during the evaporation process, so that the final state is pure.

Difficulties with Alternatives (ii) and (iii)

Alternatives (ii) and (iii a) would require the existence of a Planck scale object capable of storing (and, in the case of (iii a), releasing) arbitrarily large amounts of information. These alternatives have few, if any, advocates at the present time (and I am certainly not one of them).

Alternative (iii b) obviously requires a breakdown of quantum field theory in a regime where the spacetime curvature is far from Planckian scales, where one would expect the known laws of physics to be applicable. As AMPS has argued, this breakdown must involve either a severe violation of causality or the conversion of the horizon to a singularity (“firewall”). If the latter alternative held and the entanglement between the outside and inside of the black hole were eliminated, I would expect the Hawking radiation to shut off, making the entire analysis self-inconsistent.

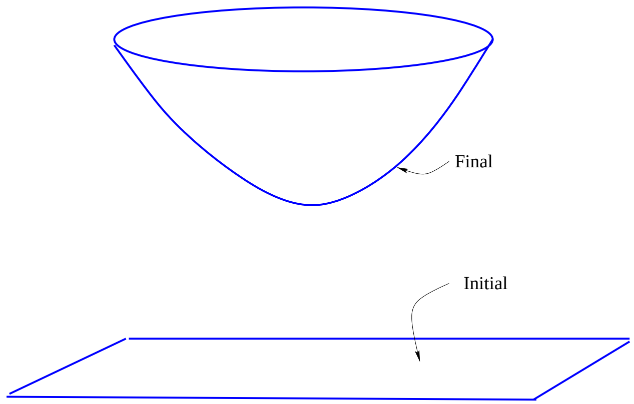
What about alternative (i)?

Against Alternative (i): 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 in alternative (i).

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.



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.

I find it ironic that some of the same people who declare "pure $\rightarrow$ mixed" to be a violation of quantum theory then endorse alternatives like (iii b), which *really are* violations of quantum theory in a regime where it should be valid. I have a deep and firm belief in the validity of the known laws of quantum theory (below the Planck scale), and I will continue to vigorously defend quantum theory against those who claim to be saving it but who are actually trying to destroy it.

Against Alternative (i): 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.

There appears to be a widespread belief that any quantum mechanical decoherence process requires energy exchange and therefore a failure of conservation of energy for the system under consideration. This is true if the "environment system" is taken to be a thermal bath of oscillators. However, it is not true in the case where the "environment system" is a spin bath. In any case, Unruh has recently provided an 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.

Bottom line: There is no problem with maintaining exact energy and momentum conservation in quantum mechanics with an evolution wherein "pure $\rightarrow$ mixed".

Against Alternative (i): AdS/CFT

"AdS/CFT" is a conjectured exact correspondence between states in quantum gravity (in asymptotically AdS spacetimes) and states of a conformal field theory (defined on the asymptotic AdS boundary). The evidence in favor of AdS/CFT consists mainly of examples of nontrivial and unexpected relationships and correspondences between various bulk and boundary theories in AdS, but there is, to date, no mathematically precise formulation of the conjecture. The AdS/CFT argument against alternative (i) is simply that the conformal field theory does not admit "pure $\rightarrow$ mixed" evolution, so such evolution must also not be possible in quantum gravity.

To my ears, this argument has a similar ring to arguing that the occurrence of various miracles suggests the existence of God and that, if God exists, His perfection should not allow ...; so therefore ... cannot exist. The problem is that if one is agnostic (rather than a true believer or an atheist), it is difficult to come to a definite conclusion about the validity of arguments concerning the true meaning miracles, the existence of God, or the exact nature of God's perfection.

I therefore hope that a clear argument against alternative (i) will emerge from the AdS/CFT ideas, in such a way that it make assertions about where, when, and how major departures from classical/semiclassical general relativity occur in the process of black hole formation and evaporation. Until then, I'm sticking with alternative (i)!

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.