PHYS 514: The Black Hole Information Paradox

Eric Hanson (# 260444061)

McGill University, Montreal, Quebec, Canada
(Dated: April 25, 2014)

CONTENTS

I. INTRODUCTION

The black hole information paradox concerns the intersection of quantum mechanics and general relativity, so its resolution could point the way towards a quantum theory of gravity. The paradox is a contradiction between our understanding of the limits of our physical theories and the consequences of them. With curvature on the scales of that in our solar system, current physical theories have been very successful experimentally; thus, there seems to exist a "solar system limit" in which the laws of physics reduce to those of quantum field theory. However, Hawking's argument [Haw75], as formalized into a theorem by Mathur [Mat11] (theorem 3 in this report), shows that a natural interpretation of this limit, along with the existence of a "traditional" black hole (such as a Schwarzschild black hole), leads to a violation of the unitarity of quantum mechanics or physically-objectionable "remnants". Two resolutions to this paradox are presented. The "firewalls" resolution formulated by [AMPS13] considers the observations of an external observer and an infalling observer, and concludes that high-energy quanta near the horizon of the black hole would prevent the paradox (and evade the theorem, as the black hole would no longer be "traditional"). A subsequent argument by Hayden and Harlow examines the computational complexity of the task presented to the infalling observer, and suggests that perhaps quantum field theory does not hold between "computationally-inaccessible" observables, allowing a resolution of the paradox without firewalls [HH13] (thus evading the theorem by interpreting the solar system limit in a different way).

II. HAWKING'S THEOREM

As presented by Mathur [Mat11], Hawking's argument [Haw75] can be formalized into a theorem. First however, the context needs to be established.

A. Preliminaries

One key observation is that in curved spacetime the notion of particle depends on the observer. It's worth a detailed look at how this comes about; in this, we follow Wald's discussion [Wal84, Ch. 14]:

First we consider a real Klein-Gordon scalar field φ in Minkowski space to see what we mean by a particle in flat space. In the Heisenberg picture, φ satisfies

$$\nabla_a \nabla^a \varphi - m^2 \varphi = 0 \tag{KG}$$

Additionally, the "particle number current"

$$j^a = -i(\bar{\varphi}\nabla^a \varphi - \varphi \nabla^a \bar{\varphi}) \tag{1}$$

is conserved via the Klein-Gordon equation.

To construct quantum states in this theory, we first define the Klein-Gordon "inner product", which is not necessarily positive definite:

$$(\alpha, \beta)_{\mathrm{KG}} = -\int_\Sigma j_a[\alpha, \beta] n^a \,\mathrm{d}V = i \sum_\Sigma (\bar{\alpha}\nabla_a \beta - \beta \nabla_a \bar{\alpha}) n^a \,\mathrm{d}V \tag{2}$$

where Σ is a Cauchy surface, α, β are solutions of KG, and n^a the normal vector. Note that this definition is independent of Σ, because j^a is conserved [Wal84]. To recover an inner product, we need to restrict to the subspace of positive frequency solutions of the KG equation: solutions φ such that for frequencies $\omega < 0$, the Fourier transform with respect to an inertial observer's time coordinate t

$$\tilde{\varphi}(\omega, \vec{x}) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{i\omega t} \varphi(t, \vec{x}) \,\mathrm{d}t \tag{3}$$

vanishes. Setting V as the vector space of smooth positive frequency KG solutions which vanish at spatial infinity with inner product given in equation 2, we obtain as the completion of V the Hilbert space $\mathcal{H}$, which is the *one particle Hilbert space*. The negative frequency solutions are in "natural linear correspondence" with $\bar{\mathcal{H}} = \mathcal{H}^*$ [Wal84].

The Hilbert space of all possible states of the field is

$$\mathcal{L}_S(\mathcal{H}) = \mathbb{C} \bigoplus \left[\bigoplus_{n=1}^{\infty} \left(\bigotimes_s^n \mathcal{H} \right) \right] \tag{4}$$

where $\mathbb{C}$ corresponds to the space for zero particles, $\mathcal{H}$ with one-particle, $\mathcal{H} \otimes \mathcal{H}$ with two particles, and so on, where the subscript s indicates that a symmetrized tensor product is to be taken. Then

$\Psi \in \mathcal{L}_s(\mathcal{H})$ is of the form

$$\Psi = (\alpha_0, \alpha_1, \alpha_, \ldots) \tag{5}$$

where $\alpha_0 \in \mathbb{C}$, $\alpha_n \in \bigotimes_s^n \mathcal{H}$. We can then define an annihilation operator $a(\bar{\sigma})$ for $\sigma \in \mathcal{H}$ as

$$a(\bar{\sigma})\Psi = (\bar{\sigma} \cdot \alpha_1, \sqrt{2}\bar{\sigma} \cdot \alpha_2, \sqrt{3}\bar{\sigma} \cdot \alpha_3, \ldots) \tag{6}$$

where $\bar{\sigma} \in \bar{\mathcal{H}}$ relates to σ by the complex conjugate map, so $\bar{\sigma} \cdot \alpha_n \in \bigotimes_s^{n-1} \mathcal{H}$ is found by "inserting $\bar{\sigma}$ into one of the 'slots' of the map α_n" [Wal84]. Then if $|0\rangle = (1, 0, 0, \ldots)$, $a(\bar{\sigma})\,|0\rangle = 0$, and this statement characterizes the vacuum up to phase. We can also define the creation operator $a^\dagger(\sigma)$ by the adjoint of the annihilation operator:

$$a^\dagger(\sigma)\Psi = (0, \alpha_0\sigma, \sqrt{2}a_1 \otimes_S \sigma, \sqrt{3}\alpha_2 \otimes_s \sigma, \ldots) \tag{7}$$

and then the quantum field operator $\hat{\varphi}(x)$:

$$\hat{\varphi}(x) = \sum_{i=1}^{\infty}[\sigma_i(x)a(\bar{\sigma}_i) + \bar{\sigma}_i(x)a^\dagger(\sigma_i)] \tag{8}$$

where $\{\sigma_i\}$ is an orthonormal basis of $\mathcal{H}$. Then $\hat{\varphi}$ solves the KG equation and the annihilation and creation operators are a basis of positive frequency solutions and their complex conjugates.

In curved spacetime, the states of the field are still vectors in a Hilbert space $\mathcal{L}$, the field φ is still described a an operator $\hat{\varphi}$ which satisfies the KG equation, and the current (equation 1) is still conserved. The difficulty comes in when attempting to restrict to a subspace of solutions of the KG equation so that the inner product (equation 2) is positive definite. In Minkowski space, that subspace was that of positive frequency solutions. Since in curved space there is no preferred time coordinate to take the Fourier transform with respect to, this procedure cannot continue; there is no natural definition of particle. If, however, the field's variation over spacetime is much smaller than the curvature of spacetime, an approximate notion of positive and negative frequency solutions exist, which allows a notion of particles which have wavelength less than the curvature radius of the universe R.

In particular, for particles with wavelength $\lambda \lesssim R$, one definition of particle will disagree with another by ~ 1 quanta of wavelength $\sim R$. This allows a notion of "empty" space.

Def. 1 A point on the horizon of a black-hole is called information-free if around this point, field modes with wavelengths $\ell_p \ll \lambda \lesssim M$ evolve by the semi-classical evolution of quantum fields on "empty" space, up to terms that vanish as $m_p/M \to 0$.

Def. 2 A black hole with an information-free horizon is called a "traditional black hole".

B. The theorem

Theorem 3 ***(Hawking's Theorem)*** *Assuming*

A1. The "niceness conditions":

(a) The quantum state is defined on a spacelike slice with small curvature. This is expressed via three conditions:

i. The intrinsic curvature of the spacelike slice on which the quantum state is defined is small ${}^3R \ll \frac{1}{\ell_p^2}$*, the Planck scale* *ii. The extrinsic curvature of the slice K is also small: $K \ll \frac{1}{\ell_p^2}$*

iii. The 4-curvature of the full space-time near the slice is small everywhere: ${}^{(4)}R \ll \frac{1}{\ell_p^2}$

(b) Quanta on the slice have wavelength $\lambda \gg \ell_p$, and small energy density: $U \ll \ell_p^{-4}$, and momentum density: $P \ll \ell_p^{-4}$. We also require all matter satisfy the dominant energy condition.

(c) States evolved from the initial state also satisfy property 1b as above, and the shift and lapse vectors change smoothly with position: $\frac{\mathrm{d}N^i}{\mathrm{d}s} \ll \frac{1}{\ell_p}$, $\frac{\mathrm{d}N}{\mathrm{d}s} \ll \frac{1}{\ell_p}$

yield local Hamiltonian evolution and

A2. A traditional black hole exists in the theory

Then the formation and evaporation of such a black hole will lead to either

C1. Mixed states: The initially pure state of the system evolves into a mixed final state.

or

C2. Remnants: given bounds $m_{remnant}, l_{remnant}$, there exists a set of objects with mass $m < m_{remnant}$ and size $l < l_{remnant}$ with arbitrarily high entanglement with systems far away from the object.

That the niceness conditions of A1 yield ordinary quantum evolution is motivated by the idea of a "solar system limit"; that is, with curvature at the scales of that found in our solar system, it seems that quantum gravity does not need to be taken into account in order to do physics. This is described in [HH13] as conditions under which the effective field theory is valid. These niceness conditions are often assumed implicitly.

Mixed states evolving from pure states violates the unitarity of quantum mechanics. Remnants do not directly violate quantum mechanics; however, for an object to have entanglement entropy $S_{\text{ent}} = n \ln 2$, the object needs to have at least n states, so the remnants must have unbounded degeneracy but bounded size and energy.

To avoid the consequences of mixed states or remnants then, either the "niceness" conditions are insufficient to allow ordinary local Hamiltonian evolution or traditional black holes do not exist. In section III, a resolution of each form is considered.

C. Proof sketch

This sketch will follow the proof presented in Mathur [Mat11]; Mathur not only presents a more detailed proof but also shows that corrections to the leading order state cannot avoid the consequences of C1 or C2.

We will assume the traditional black hole obeys the Schwarzschild metric

$$\mathrm{d}s^2 = -\left(1 - \frac{2M}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2M}{r}\right)^{-1}\mathrm{d}r^2 + r^2\,\mathrm{d}\Omega_2^2. \tag{9}$$

While this metric yields a time independent geometry, we will assume that the evaporation process is slow enough so that at any point of the evaporation, the black hole can be described by a metric like (9) (except at the end when the black hole becomes Planck-sized).

1. A nice slicing

To begin with, we need to find a set a spacelike slices so that the niceness conditions A1 hold and the details of the Hawking process are captured.

One slice will be constructed as shown in Figure 1:

1. For $r > 4M$, we set $t = t_1 = \text{constant}$
2. For $r < 2M$, we set $r = r_1 = \text{const}$ for $\frac{M}{2} < r_1 < \frac{3M}{2}$, to avoid the horizon and singularity
3. We join these with a smooth 'connector' $\mathcal{C}$ which can be chosen so that the niceness conditions A1 hold.

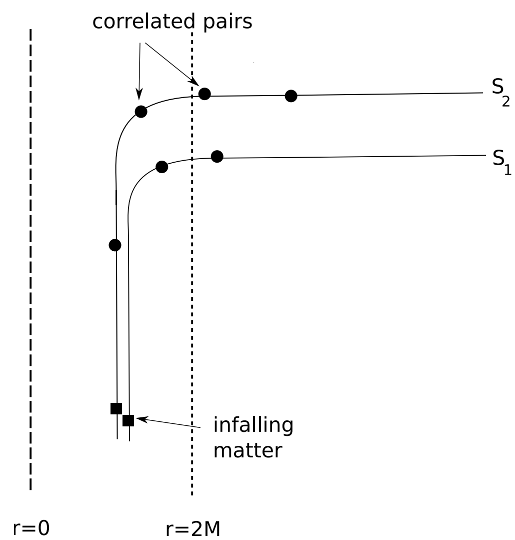

FIG. 1: Two spacelike slices near the Schwarzschild black hole. Outside the horizon, $t = \text{const}$ and inside the horizon, $r = \text{const}$; these sections are joined smoothly by a 'connector' that has to stretch as the slices evolve. Taken from Mathur's review [Mat11, Figure 2]

2. Evolution

To make a 'later' slice,

1. For $r > 4M$, set $t = t_1 + \Delta t$
2. For $r = r_1$ portion, set $r = r_1 + \delta$ for $\delta \ll M$
3. Again, a smooth connector $\mathcal{C}$ connects these portions. Because the $t = \text{const}$ portion of the slice has evolved forward, the $r = \text{const}$ part is longer, and so $\mathcal{C}$ has to stretch (see Figure 1).

Note that in the $t = \text{const}$ and $r = \text{const}$ portions of the slices, there is no change in the intrinsic geometry. The connector portion stretches however; thus, near the horizon is where all the changes between successive slices takes place. Then at the horizon, the Fourier modes of the field are stretched to longer wavelengths and particles are produced.

3. Pair creation

On an initial spacelike slice there is a matter state $|\psi\rangle_M$, which is the matter shell that collapse to make the black hole. On the next slice, as described, the portion of the slice near the horizon stretches; this creates correlated pairs b_1 and c_1 so that the state on the slice is:

$$|\Psi\rangle \approx |\psi\rangle_M \otimes \left(\frac{1}{\sqrt{2}}|0\rangle_{c_1}|0\rangle_{b_1} + \frac{1}{\sqrt{2}}|1\rangle_{c_1}|1\rangle_{b_1}\right) \tag{10}$$

Then the entanglement entropy is $S_{\text{ent}} = \ln 2$ (maximally entangled state).

On the next slice, $|\psi\rangle_M$ doesn't change because there in that part of the slice there is no evolution. The stretching occurs where the "connector" piece is, near the horizon. There, the pairs b_1, c_1 are moved away from the region of stretching, and new pairs b_2, c_2 are created. The state is then

$$|\Psi\rangle \approx |\psi\rangle_M \otimes \left(\frac{1}{\sqrt{2}}|0\rangle_{c_1}|0\rangle_{b_1} + \frac{1}{\sqrt{2}}|1\rangle_{c_1}|1\rangle_{b_1}\right) \otimes \left(\frac{1}{\sqrt{2}}|0\rangle_{c_2}|0\rangle_{b_2} + \frac{1}{\sqrt{2}}|1\rangle_{c_2}|1\rangle_{b_2}\right) \tag{11}$$

and the entanglement entropy is $S_{\text{ent}} = 2\ln 2$. After N slices, N pairs have been created and the entanglement entropy is $S_{\text{ent}} = N \ln 2$; see Figure 2.

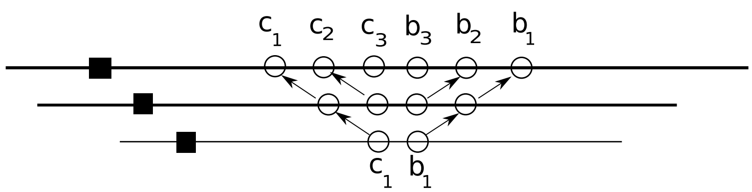

FIG. 2: Hawking pairs created by the evolution of spacelike slices. On each successive slice, pairs from the previous slice move further apart and a new pair is created near the horizon. Figure taken from Mathur's review [Mat11, Figure 4]

As the $\{b_i\}$ quanta accumulate at infinity, the mass of the black hole decreases; after $M_{\text{hole}} \sim m_{\text{planck}}$ the niceness conditions no longer hold and we stop evolving the slices. These quanta $\{b_i\}$ have entanglement $N \ln 2$ with the $\{c_i\}$. Either the black hole evaporates away completely; since the $\{b_i\}$ have entanglement entropy $N \ln 2$, the final state must be described by a density matrix. Otherwise, the evolution stops with $M_{\text{hole}} \sim m_{\text{remnant}}$, and we have remnants.

III. POSSIBLE RESOLUTIONS

A. Firewalls

One resolution is simply that condition A2 of Hawking's theorem (3) does not hold and instead an infalling observer encounters "firewalls" of high-energy quanta near the horizon. As presented in [AMPS13] and summarized in [HH13], the argument proceeds by considering an external observer Charlie and an infalling observer Alice.

1. Charlie's view

As seen by Charlie,

1. The formation and evaporation of the black hole is unitary, described by some pure state $|\Psi\rangle$ in a Hilbert space $\mathcal{H}_{\text{outside}}$.

2. We decompose $\mathcal{H}_{\text{outside}} = \mathcal{H}_H \otimes \mathcal{H}_B \otimes \mathcal{H}_R$, where $\mathcal{H}_R$ describes the radiation field outside the black hole, $\mathcal{H}_B$ the field theory modes just outside the horizon, and $\mathcal{H}_H$ the remaining modes in the black hole.

3. Setting $|H| = \dim \mathcal{H}_H$ and $|B| = \dim \mathcal{H}_B$, both $\log |H|$ and $\log |B|$ are proportional to the area of the horizon in Planck units, and so their size decreases with time. We consider $\mathcal{H}_R$ to only involve the portion involved with the black hole dynamics; hence, $|R| = \dim \mathcal{H}_R$ grows in time as more pairs are created.

Since $|R|$ grows and $|B|, |H|$ shrink in time, a distinction can be made between "young" black holes when $|R|$ is small compared to $|B||H|$ and "old" black holes when B and H are small compared to the whole space $\mathcal{H}_{\text{outside}}$. In the later case, the combined system BH is described by

$$\rho_{BH} \approx \frac{1}{|B||H|} I_B \otimes I_H \tag{12}$$

the maximally mixed state.

Furthermore, the BH system is maximally entangled with a subspace in R, as described in the proof of Hawking's theorem (section II C 3). Then we can write

$$\mathcal{H}_R = (\mathcal{H}_{R_H} \otimes \mathcal{H}_{R_B}) \oplus \mathcal{H}_{\text{other}} \tag{13}$$

where $|R_H| = |H|$, $|R_B| = |B|$, and we can write

$$|\Psi\rangle = \left(\frac{1}{\sqrt{|H|}} \sum_h |h\rangle_H |h\rangle_{R_H} \right) \otimes \left(\frac{1}{\sqrt{|B|}} \sum_b |b\rangle_B |b\rangle_{R_B} \right) \tag{14}$$

where h and b are bases for $\mathcal{H}_H$ and $\mathcal{H}_B$. Thus, the state is entangled between H and R_H as well as between B and R_B.

2. Alice's view

As seen by Alice,

1. Well before she hits the singularity, Alice observes an approximately quantum-mechanical system

2. Her Hilbert space is described by $\mathcal{H}_{\text{inside}} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_R \otimes \mathcal{H}_{H'}$, where $\mathcal{H}_B$ and $\mathcal{H}_R$ are shared with Charlie, $\mathcal{H}_A$ are the field modes just inside the horizon, and $\mathcal{H}_{H'}$ the remaining degrees of freedom.

3. Since $\mathcal{H}_R$ and $\mathcal{H}_B$ are spaces shared with Charlie, their states on the overlap agree: $\rho_{BR}^{\text{Alice}} = \rho_{BR}^{\text{Charlie}}$.

3. The contradiction

Since $\mathcal{H}_A$ and $\mathcal{H}_B$ describe field modes just outside and just inside the horizon, they are close to maximally entangled in the Minkowski vacuum [HH13]:

$$|\text{vac}\rangle \sim \sum_w e^{-\beta_\omega/2} |\omega\rangle_A |\omega\rangle_B \tag{15}$$

where β_ω is the Rindler energy. However, since Charlie sees the state on B as close to maximally entangled with R_B and since Alice sees the same as Charlie on the overlap, Alice must also see the state on B as close to maximally entangled with R_B. By the monogamy of entanglement, Alice cannot see the state on B as close to maximally entangled with both R_B and A.

One resolution here is that the horizon is not the Minkowski vacuum and instead there are high energy quanta that "annihilate" the infalling observer.

B. Quantum computation

As proposed by Harlow and Hayden [HH13], operational considerations may prohibit Alice and Charlie from observing discrepancies between their perceptions of R_B and B; that is, their states may disagree on the overlap (described in Alice's item 3), but they won't be able to do any measurement to observe this. This amounts to a new criteria for the validity of an effective field theory (and hence violates assumption A1 of Hawking's theorem):

> Two spacelike-separated low-energy observables which are not both computationally accessible to a single observer do not need to be realized even approximately as distinct and commuting observers on the same Hilbert space.

where *computationally inaccessible* means it would take more time or memory than is fundamentally available to the observer to measure both observables.

The idea is that Alice may not have time to measure the state on R_B sufficiently to measure the close-to-maximal entanglement between B and R_B before the black hole evaporates. In this case, R_B is computationally inaccessible from B, and we can say the effective field theory is not valid and so simply remove R_B from Alice's Hilbert space. Then B can be maximally entangled with A, and Alice can smoothly cross the horizon.

The entropy of a Schwarzschild black hole (9) is proportional to M^2 and evaporates in time proportional to M^3, so Alice would need to extract information from $n \sim M^2$ bits of Hawking radiation before $n^{3/2}$ time, so she would need to be able to do so very efficiently: the time needed scales as a low-order polynomial of the input size. Applying an arbitrary n-bit unitary transform requires exponential time in n; [HH13] argue that the decoding needed here is unlikely to be much faster. Hayden and Harlow also consider charged black holes with the Reissner-Nordstrom metric, and other complications, and show that they are unlikely to allow a significant speed-up of the computation.

IV. CONCLUSION

The black hole information paradox probes the limits of validity of our physical theories and presents holes in our understanding. The stretching of spacelike slices leads to entangled pair creation near the horizon of traditional black holes; ordinary local Hamiltonian evolution leads to either a violation of unitarity or remnants with unbounded degeneracy. Two resolutions are presented: ruling out traditional black holes by postulating firewalls at the horizons, or redefining the validity of our physical theories by considering fundamental computational limits as physically relevant. Results from string theory, not discussed here, also could point towards a resolution [Mat11].

[AMPS13] Ahmed Almheiri, Donald Marolf, Joseph Polchinski, and James Sully. Black Holes: Complementarity or Firewalls? *JHEP*, 1302:062, 2013.

[Haw75] S.W. Hawking. Particle creation by black holes. *Communications in Mathematical Physics*, 43(3):199–220, 1975.

[HH13] Daniel Harlow and Patrick Hayden. Quantum Computation vs. Firewalls. *JHEP*, 1306:085, 2013.

[Mat11] Samir D. Mathur. The Information paradox: A Pedagogical introduction. *Class.Quant.Grav.*, 26:224001, 2011.

[Wal84] Robert M. Wald. *General Relativity*. University Of Chicago Press, 1984.