The Horizon Is Special

Raphael Bousso

Berkeley Center for Theoretical Physics
University of California, Berkeley

KITP Firewall Workshop, August 19, 2013

Because of the Equivalence Principle.

1. Equivalence Principle $\rightarrow$ The vacuum looks the same everywhere.
2. No Hair Theorem $\rightarrow$ A black hole horizon is vacuous.

If we insist on (2), then firewalls violate (1).

Because of the Equivalence Principle.

1. Equivalence Principle $\rightarrow$ The vacuum looks the same everywhere.
2. No Hair Theorem $\rightarrow$ A black hole horizon is vacuous.

If we insist on (2), then firewalls violate (1).

This is their *only* objectionable feature.

Two implications:

(1) Nothing Else Matters

Why pursue approaches that reduce the drama,
but in which the horizon remains a special place?

(1) Nothing Else Matters

Why pursue approaches that reduce the drama,
but in which the horizon remains a special place?

- Modifications of EFT (“nonviolent nonlocality”)

(1) Nothing Else Matters

Why pursue approaches that reduce the drama,
but in which the horizon remains a special place?

- Modifications of EFT ("nonviolent nonlocality")
- Identification of Hilbert spaces ("$A = R_B$", "ER = EPR")

(1) Nothing Else Matters

Why pursue approaches that reduce the drama,
but in which the horizon remains a special place?

- Modifications of EFT (“nonviolent nonlocality”)
- Identification of Hilbert spaces (“$A = R_B$”, “ER $=$ EPR”)

I will mainly focus on identification of Hilbert spaces.

I will argue that if they reconstruct the vacuum, the horizon can still be locally detected.

(2) No Vacuum, No Problem

If the adiabatic vacuum can break down after a scrambling time, without conflict with causality, Lorentz invariance, and cosmology, then firewalls are acceptable.

Quantum Mechanics Argument

Strong subadditivity states:

$$S_{XYZ} + S_Y \leq S_{XY} + S_{YZ} \ .$$

Suppose that XY is in a pure, entangled state (the "vacuum"):

$$S_{XY} = 0 \ , \quad S_Y > 0 \ .$$

Then there cannot exist a third system Z, entirely distinct from X, such that YZ is in a pure state (the "out-state").
For otherwise we would have

$$S_{YZ} = 0 \ ,$$

and strong subadditivity would be violated.

Quantum Mechanics Argument

Equivalently:
Suppose that the "out-state" is pure (entangled or not):

$$S_{YZ} = 0 \; .$$

Then

$$S_{XY} = 0 \; , \quad S_Y > 0 \; .$$

cannot hold, i.e., Y cannot form an entangled pure state (the "vacuum") with any other system X.

Quantum Mechanics Argument

Equivalently:
Suppose that the "out-state" is pure (entangled or not):

$$S_{YZ} = 0 \ .$$

Then

$$S_{XY} = 0 \ , \quad S_Y > 0 \ .$$

cannot hold, i.e., Y cannot form an entangled pure state (the "vacuum") with any other system X.

I will argue that identifications of Hilbert spaces cannot overcome this difficulty.

Quantum Mechanics Argument

Equivalently:
Suppose that the "out-state" is pure (entangled or not):

$$S_{YZ} = 0 \; .$$

Then

$$S_{XY} = 0 \; , \quad S_Y > 0 \; .$$

cannot hold, i.e., Y cannot form an entangled pure state (the "vacuum") with any other system X.

I will argue that identifications of Hilbert spaces cannot overcome this difficulty.

But first, what exactly is Y, and why is only the horizon affected?

Minable modes

In the black hole geometry, the double purity conflict arises if we identify Y with any *minable mode b*, defined by the following two properties:

- In the infalling vacuum, the entropy of b is large, and b is (nearly) purified by an interior partner $\tilde{b}$ ("X"):

$$S_b \sim O(1) \; ; \quad S_{b\tilde{b}} \ll 1 \; .$$

- If b were extracted from near the black hole and transported to infinity, this would decrease the area of the black hole. Unitarity of the S-matrix demands

$$S_{b,F-b} = 0 \; ,$$

  where $F - b$ is the remainder ("Z") of the out-state (the Hawking radiation).

Minable zone

Minable modes

- live in the *zone*, which ranges from R to $3R/2$,
- at a time when the black hole horizon is a Killing horizon with respect to the asymptotic Killing vector field;
- additionally b must be well-localized with respect to the horizon
- and within the angular momentum barrier

The number of independent minable modes is $O(A)$, depending on the cutoff. Dominated by UV modes near horizon.

Example 1: No mining, no firewall

Accelerated detector at some Rindler horizon far from the black hole

- Does not decrease black hole mass
- Merely pair-creates particles
- If one of them falls into the black hole, Hawking radiation has to purify remaining particle
- But black hole mass goes up
- So the purification need not have appeared in the radiation, had the experiment not been performed
- Hence, consistent to declare that the modes were in the vacuum before they were disturbed

Example 2: No mining, no firewall

Accelerated detector at the event horizon, inside collapsing shell

- Not yet a Killing horizon
- Bigger black hole forms than if experiment had not been performed
- Again, consistent to declare that the modes were in the vacuum before they were disturbed

Successful Mining

- Grab a wavepacket that matches Rindler or Schwarzschild modes between the horizon and the barrier
- Imperfect mining may create additional energy outside the black hole, which may then fall into the black hole
- As long as this energy is very small, the black hole mass is not increased by enough to ascribe the purification of b to the resulting additional Hawking radiation
- But by unitarity, the Hawking radiation purifies b anyway
- So the extracted mode cannot have been purified by its local partner, independently of whether the experiment was carried out

Approximate Minability is Good Enough

Therefore,

- *b* can be taken to have support strictly away from the stretched horizon, where hidden degrees of freedom may reside and where *b* may not be free.
- We only need the angular momentum barrier to be far from the horizon, in Planck units.

Unitary Complementarity

- Infalling matter, inside the horizon, must be mapped to a subspace of F, the out-Hilbert space.
- If the map is unitary, this is consistent with a unitary S-matrix
- And it resolves the xeroxing paradox
- Extend this map to the infalling vacuum?

$$\tilde{\mathbf{b}}(t) \stackrel{!}{=} \hat{\mathbf{b}} \subset \mathbf{F}\ .$$

Unitary Complementarity

But let $B(t)$ be the collection of minable modes at time t. By unitarity, $\bigotimes'_t \mathbf{B}(t) \subset \mathbf{F}$. Every Hawking quantum was trivially minable at some point so $\bigotimes'_t \mathbf{B}(t) \supset \mathbf{F}$. Hence

$$\mathbf{F} = \bigotimes_t{}' \mathbf{B}(t) \ .$$

Unitary Complementarity

Exclude collapse era (at most, log R Hawking particles, which cannot carry away info). Then only vacuum:

$$|0\rangle_{\tilde{b}b} \propto \sum_{n=0}^{\infty} e^{-\beta n\omega/2} |n\rangle_{\tilde{b}} \otimes |n\rangle_b \ .$$

The collection of all b modes spans **F**, and with the unitary complementarity map, the collection of $\tilde{b}b$ pairs still only spans **F**. Mode by mode, the vacuum defines a unique state $|0\rangle_F$ in **F**. This out-state would be pure, unlike in Hawking's calculation. But it would be independent of the in-state, in violation of unitarity. Information would be lost.

Unitary Complementarity

- Note that unitary complementarity is consistent if applied to infalling matter only
- Unlike donkey map, causal and linear
- New firewall forms a scrambling time after most recent infall
- Note that pull-back-push-forward breaks down after scrambling time

Overlap condition

- Let's worry only about Alice, who falls in at time t
- The most general map one can consider is a many-to-one map from **F** to $\tilde{\mathbf{B}}(t)\mathbf{B}(t)$ that takes $|\Psi\rangle_F \rightarrow |0\rangle_{\tilde{B}B}$, for any $|\Psi\rangle_F$.
- However, both the total Hawking radiation F, and the neighborhood of the horizon $\tilde{B}B$ contain the minable zone B as subsystems.
- In general, such a map would assign two inconsistent states to B.

Pure Zone

- This problem arises for any state $|\Psi\rangle_F$ such that the reduced density matrix for $B(t)$ is not thermal.
- Then no identification of $\tilde{B}$ with any other degrees of freedom can reconstruct the vacuum.
- An extreme example is states where the zone modes are pure.
- Such states span the microcanonical ensemble
- Firewalls for such states alone make the horizon special

Canonical Ensemble

- Consider a black hole in a box, coupled to a heat bath.
- The black hole consists of the zone, B, and possibly other semiclassically inaccessible (“membrane”, “horizon”) degrees of freedom, H
- By assumption, H couples to B only through modes that get stretched or shrunk in and out of the semiclassical regime, i.e., locally at the stretched horizon
- Hence H cannot couple to distant zone modes, nor directly to the heat bath
- H is similar to a solid ball hovering in the center of a blackbody cavity whose outer wall is coupled to a heat bath

Canonical Entropy is Additive

- Black hole thermodynamics fixes S_{BH} ($= A/4$ but this is not relevant here)
- Subadditivity implies $S_B + S_H \geq S_{BH}$
- Minability implies $S_B + S_H = S_{BH}$
- Otherwise, mutual information would build up between the black hole and the heat bath

Microcanonical Ensemble Factorizes

- Hence, the zone (up to a Planck-size interaction region) can be treated independently of H, as a system coupled to a heat bath
- Can go over to the microcanonical ensemble for B alone
- As for any large system with canonical entropy S_B, the associated microcanonical ensemble spans a Hilbert space $\bar{\mathbf{B}}$ of finite dimension $\exp(S_B)$.
- If $S_B < S_{BH}$ then there exists another Hilbert space factor $\bar{\mathbf{H}}$ of dimension $\exp(S_H)$
- The full microcanonical ensemble for the black hole, $\bar{\mathbf{B}} \otimes \bar{\mathbf{H}}$, is spanned by a product state basis $\{|i\rangle_{\bar{H}} \otimes |j\rangle_{\bar{B}}\}$
- There is a firewall in every such basis state for the microcanonical ensemble of the black hole

Acceptable deviations from the vacuum

- UV detections in the adiabatic vacuum are exponentially suppressed, $e^{-O(\omega R)}$
- Satisfied in zone of black hole, away from horizon thanks to Boltzmann suppression of heavy emissions, $e^{-E/T}$
- Emissions of energy $E \gg T$ constitute a Hilbert space, but of dimension much smaller than $e^{S_{BH}}$.
- They do not span the full Hilbert space of the black hole
- In free fall, heavy emissions are equally likely to be encountered near or far from the horizon

A Complete Firewall Basis Is Not OK

- Product states are Haar-rare, not Boltzmann-rare
- Do not constitute a Hilbert space since product structure not preserved under linear superpositions
- But they do span the full Hilbert space of the black hole
- Firewalls sharply localized to horizon, nowhere else in zone
- In the infalling theory, the horizon is special

After Page

- So far the Page time played no special role in any arguments.
- Now I will assume that the black hole has less entropy than the radiation it has emitted.
- Then b can be approximately purified by a (highly non-unique) scrambled subsystem e_b of the early Hawking radiation alone:

$$S_b \sim O(1)\ , \quad S_{be_b} \lll 1$$

- To avoid conflict with $S_{b\tilde{b}} \ll 1$, identify b with e_b under a state-dependent “donkey map”
- I will begin by showing that the donkey map must self-correct for interactions

Donkey Map

Begin by treating e_b as unique. Also, don't worry yet about computational difficulty of extracting e_b. We can write

$$|\psi\rangle_{be_bp} \propto |0\rangle_p \otimes \sum_{n=0}^{\infty} x^n |n\rangle_b |n\rangle_{e_b} \ .$$

I have included a pointer p which has not yet interacted with either b or e_b.

The states of e_b have been labeled so as to make the donkey map look simple:

$$|0\rangle_{e_b} \to |0\rangle_{\tilde{b}} \ , \ \ |1\rangle_{e_b} \to |1\rangle_{\tilde{b}}, \quad \ldots$$

Measuring e_b

Suppose that the pointer measures e_b in the above basis. Now the state is

$$|\psi\rangle_{be_bp} \propto \sum_{n=0}^{\infty} x^n |n\rangle_b |n\rangle_{e_b} |n\rangle_p \ .$$

Tracing over pointer (=environment) decoheres the system:

$$\rho_{be_b} = (1 - x^2) \sum_{n=0}^{\infty} x^{2n} |n\rangle_b \, |n\rangle_{e_b} \ {}_{e_b}\langle n| \, {}_b\langle n| \ .$$

This state cannot be mapped to $|0\rangle_{b\tilde{b}}$ under any map from e_b to $\tilde{b}$, because it is not pure.

This is not ok: Hawking radiation will generically interact with an environment.

Adjusting the Map

To work with a pure state, consider entire be_bp system. With the updated map,

$$|0\rangle_{e_b}|0\rangle_p \to |0\rangle_{\tilde{b}} \ , \ \ |1\rangle_{e_b}|1\rangle_p \to |1\rangle_{\tilde{b}} \ \ \ldots \ , \tag{1}$$

the state

$$|\psi\rangle_{be_bp} \propto \sum_{n=0}^{\infty} x^n |n\rangle_b |n\rangle_{e_b} |n\rangle_p \ .$$

becomes the infalling vacuum.

But this means that the donkey map cannot ignore the environment, no matter whether the infalling observer even knows the interaction happened.

Measuring b

Now suppose that the pointer p measured b, not e_b. This results in the same state,

$$|\psi\rangle_{be_bp} \propto \sum_{n=0}^{\infty} x^n |n\rangle_b |n\rangle_{e_b} |n\rangle_p \ .$$

A larger environment (more pointers) could be used and the outcome, $N_b = 4$, can be recorded in classical records. Tracing over any pointer decoheres the state. For example Bob, who ran the experiment, knows that if he measures b again, he will find it in the pure state $|4\rangle_b$. Bob and the pointers stay outside the black hole but disperse.

Frozen vacuum

Much later, clueless Alice falls through the zone and encounters b.

Frozen vacuum

Much later, clueless Alice falls through the zone and encounters b.

Recall that her theory of black holes says that $\tilde{b}$ must be identified with whatever purifies b, whether or not Alice controls the purifying system or has any idea where it is.

Frozen vacuum

Much later, clueless Alice falls through the zone and encounters b.

Recall that her theory of black holes says that $\tilde{b}$ must be identified with whatever purifies b, whether or not Alice controls the purifying system or has any idea where it is.

In the above state, the purification happens to be a subspace of $e_b p$.

The associated donkey map is Eq. (1), and the result is the infalling vacuum.

Frozen vacuum

But this contradicts the fact that Alice could have met the pointer, or Bob.

Indeed, she could have arranged for the earlier measurement herself!

It follows that Alice is unable to find anything *but* the vacuum, even in cases where her own actions should have destroyed the vacuum.

This means that the vacuum near the horizon behaves differently from the vacuum everywhere else.

Alice can detect the location of the horizon by her inability to produce particles there. This violates the equivalence principle.

More complicated rule?

- Same $|\psi\rangle_{be_bp}$, both when we wanted the vacuum, and when we didn't!
  $\rightarrow$ rule cannot be based on the state

More complicated rule?

- Same $|\psi\rangle_{be_b p}$, both when we wanted the vacuum, and when we didn't!
  $\rightarrow$ rule cannot be based on the state
- Rule would have to be based on how the state came to be: interactions in zone vs. interactions with radiation?

More complicated rule?

- Same $|\psi\rangle_{be_b p}$, both when we wanted the vacuum, and when we didn't!
  $\rightarrow$ rule cannot be based on the state
- Rule would have to be based on how the state came to be: interactions in zone vs. interactions with radiation?
- Hard to make sharp, since we can interpolate between the two limits: take e_b, p to the zone, arrange three-way interaction

More complicated rule?

- Same $|\psi\rangle_{be_bp}$, both when we wanted the vacuum, and when we didn't!
  $\rightarrow$ rule cannot be based on the state
- Rule would have to be based on how the state came to be: interactions in zone vs. interactions with radiation?
- Hard to make sharp, since we can interpolate between the two limits: take e_b, p to the zone, arrange three-way interaction
- Limitations from computation? From backreaction?

Large systems

Truncate occupation numbers to $0, 1$. The vacuum is represented by the state

$$|0\rangle_{b\tilde{b}} \propto |0\rangle_b |0\rangle_{\tilde{b}} + |1\rangle_b |1\rangle_{\tilde{b}} \ .$$

Let e be a single computational qubit in the early Hawking radiation, i.e, a quantum such as a photon that can easily be measured.

Let f be the rest of the early Hawking radiation, consisting of an enormous number of qubits.

As before, p is a pointer that can be coupled to b, or to e.

Large systems

Initially, the state is

$$\begin{aligned} |\psi\rangle_{befp} \propto |0\rangle_p \otimes [|0\rangle_b &\otimes (|0\rangle_e|\alpha\rangle_f + |1\rangle_e|\beta\rangle_f) + \\ |1\rangle_b &\otimes (|0\rangle_e|\gamma\rangle_f + |1\rangle_e|\delta\rangle_f)] \ , \end{aligned}$$

where $|\alpha\rangle, \ldots, |\delta\rangle$ are four mutually orthogonal states in the vast system f.

Measure b

First suppose that the pointer interacts with b, resulting in the state

$$\begin{aligned}|\psi\rangle_{befp} &\propto |0\rangle_b|0\rangle_p \otimes (|0\rangle_e|\alpha\rangle_f + |1\rangle_e|\beta\rangle_f) \\ &+ |1\rangle_b|1\rangle_p \otimes (|0\rangle_e|\gamma\rangle_f + |1\rangle_e|\delta\rangle_f)\ .\end{aligned}$$

If we did not trace over the pointer, the donkey map would restore the vacuum, in violation of the equivalence principle. Hence, we must trace over p. This results in a mixed state for bef,

$$\begin{aligned}|\phi_1\rangle_{bef} &\propto |0\rangle_b \otimes (|0\rangle_e|\alpha\rangle_f + |1\rangle_e|\beta\rangle_f) \quad [50\%] \\ |\phi_2\rangle_{bef} &\propto |1\rangle_b \otimes (|0\rangle_e|\gamma\rangle_f + |1\rangle_e|\delta\rangle_f) \quad [50\%]\ ,\end{aligned}$$

to which no donkey map can be applied. The vacuum is destroyed by the measurement, as required.

Measure e

Now suppose that the pointer interacts with e instead, resulting in the state

$$\begin{aligned} |\psi\rangle_{befp} &\propto |0\rangle_b \otimes (|0\rangle_e|0\rangle_p|\alpha\rangle_f + |1\rangle_e|1\rangle_p|\beta\rangle_f) \\ &+ |1\rangle_b \otimes (|0\rangle_e|0\rangle_p|\gamma\rangle_f + |1\rangle_e|1\rangle_p|\delta\rangle_f) \ . \end{aligned}$$

If we traced over p, we would obtain a mixed state for bef, with equal probability for the two states

$$\begin{aligned} |\phi_1\rangle_{bef} &\propto |0\rangle_e \otimes (|0\rangle_b|\alpha\rangle_f + |1\rangle_b|\gamma\rangle_f) \quad [50\%] \\ |\phi_2\rangle_{bef} &\propto |1\rangle_e \otimes (|0\rangle_b|\beta\rangle_f + |1\rangle_b|\delta\rangle_f) \quad [50\%] \ . \end{aligned}$$

No single map from bef to $\tilde{b}$ can convert this mixed state into a pure state. To obtain the vacuum (now desired!), must retain p and apply the donkey map to the full system.

Again, it is hard to write down a general rule that distinguishes these two types of measurement, or interpolates the donkey map appropriately

Neither backreaction nor hard computation prevents e and p from being transported to the zone.

Dial ep and eb interactions to interpolate, or have ebp interaction.

Quantum Error Correction?

Exploit the redundancy of e_b to reconstruct $\tilde{b}$ even after interactions have taken place, in a spirit reminiscent of quantum error correction?

This breaks down in the generic case where all of the radiation interacts with a much larger environment.

An untouched portion of the original system, such as f above, contains an approximate purification of b only if f constitutes more than half of the qubits of the total system.

For an old black hole, the radiation alone constitutes more than half of the system, so a decoherent measurement of the radiation destroys *all* possible purifications of b.

The Challenge

The only option for recovering the vacuum is to apply a donkey map to (at least half of) the entire system consisting of black hole, the radiation, and the vastly larger environment it has interacted with.

But then we are back to the problem discussed above: why not apply the same map to recover the vacuum if b has *also* been measured?

Again, there is no sharp distinction between interactions that only involve most of the radiation, and interactions that in addition involve b.

No Vacuum, No Problem

If the adiabatic vacuum can break down after a scrambling time, without conflict with causality, Lorentz invariance, and cosmology, then firewalls are acceptable.

No Vacuum, No Problem

If the adiabatic vacuum can break down after a scrambling time, without conflict with causality, Lorentz invariance, and cosmology, then firewalls are acceptable.

It is time to constrain and construct the dynamics of firewalls.