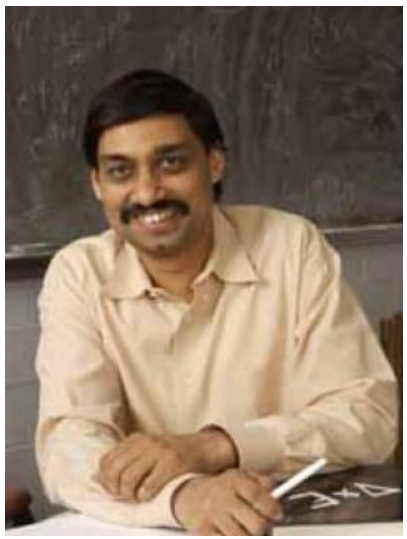
Why Black Hole Information Loss is Paradoxical

David Wallace (University of Pittburgh)

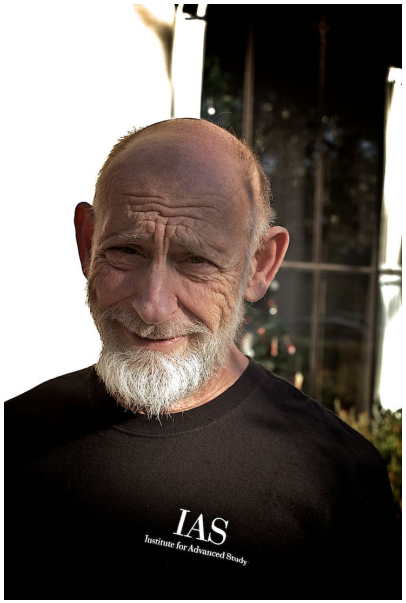
29th June 2022

Samir Mathur



We shall see that there is a very precise statement of the contradiction found by Hawking, and that bypassing the paradox needs a basic change in our understanding of how quantum effects operate in gravity.
(arxiv:0909.1038)

Leonard Susskind

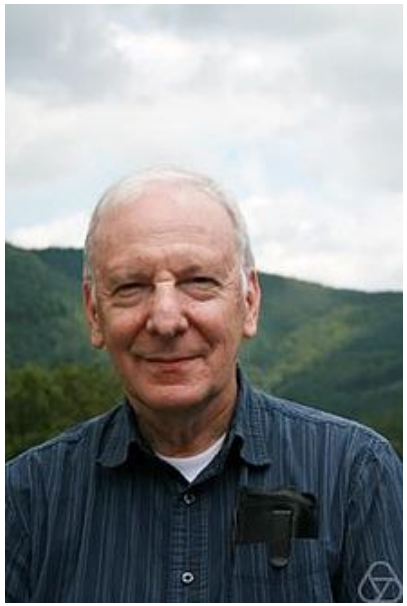


Whatever else Maldacena and Witten had done, they had proved beyond any doubt that information would never be lost behind a black hole horizon. The string theorists could understand this immediately; the relativists would take longer. (The Black Hole War)

Stephen Hawking

Because part of the information about the state is lost down the hole, the final situation is represented by a density operator rather than a pure quantum state. This means there is no S matrix for the process of black-hole formation and evaporation. (Phys Rev D 14, 2460)

Bill Unruh and Bob Wald



The proposals to [avoid information loss] typically require drastic violation of the local laws of physics . . . loss of information in black hole formation and evaporation does not violate any fundamental principle of physics and is not, in any way, a radical proposal. Thus, our strong inclination is to believe that there is loss of information in the process of black hole formation and evaporation. (arxiv:1703.02140)

Gordon Belot, John Earman, Laura Ruetsche

Assuming that black hole evaporation can be described by a spacetime of classical GR, that the evaporation is not of the thunderbolt type, and that the quantum aspects of the problem can be described by QFT on the resulting spacetime, then none of the escape routes discussed in the literature provides a plausible way to avoid the conclusion that the post-evaporation state is mixed. . . . information is lost, as Hawking originally maintained. (BJPS 50 (1999) pp.189-229)

Tim Maudlin

The so-called "information loss paradox" arises rather from the inaccurate application of foundational principles, involving both mathematical and conceptual errors. The resources for resolving the "paradox" are familiar and uncontroversial, and have been pointed out in the literature. The problem ought to have been dismissed 40 years ago. Recent radical attempts to "solve" the problem are blind alleys, solutions in search of a problem.
(arxiv:1705.03541)

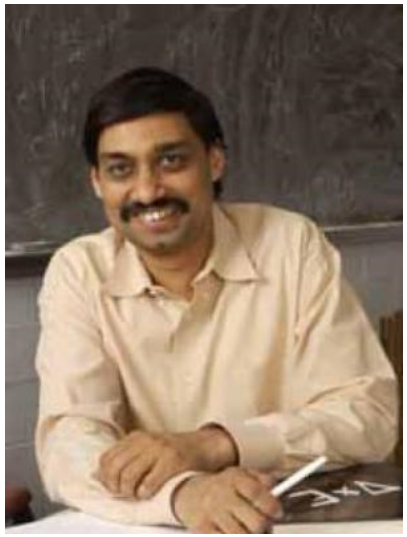
Maudlin again

How could a simple solution have gone unappreciated by so many theoretical physicists — among them very great physicists — over such a long period of time? Probably no completely satisfactory non-sociological explanation is possible. (ibid.)

Hawking again

[Q]uantum gravity is unitary and information is preserved in black hole formation and evaporation.. (arxiv:hep-th/0507171)

Mathur again



Not everyone understands Hawking's paradox the same way [.] (ibid.)

Classical black hole thermodynamics

Classical black hole thermodynamics

- Almost perfect analogy between BH thermodynamics and (self-gravitating) ordinary thermodynamics *as long as heat flow is not present* (Bardeen, Hawking, Christodolou, Ruffini, etc, etc), with BH area playing the role of entropy

Classical black hole thermodynamics

- Almost perfect analogy between BH thermodynamics and (self-gravitating) ordinary thermodynamics *as long as heat flow is not present* (Bardeen, Hawking, Christodolou, Ruffini, etc, etc), with BH area playing the role of entropy
- Extends to local description of BH surface as thermodynamic (via membrane paradigm of Thorne *et al*)

Classical black hole thermodynamics

- Almost perfect analogy between BH thermodynamics and (self-gravitating) ordinary thermodynamics *as long as heat flow is not present* (Bardeen, Hawking, Christodolou, Ruffini, etc, etc), with BH area playing the role of entropy
- Extends to local description of BH surface as thermodynamic (via membrane paradigm of Thorne *et al*)
- Extends to arbitrary diffeomorphism-covariant theory of gravity (at cost of complicating the formula for entropy)

Classical black hole thermodynamics

- Almost perfect analogy between BH thermodynamics and (self-gravitating) ordinary thermodynamics *as long as heat flow is not present* (Bardeen, Hawking, Christodolou, Ruffini, etc, etc), with BH area playing the role of entropy
- Extends to local description of BH surface as thermodynamic (via membrane paradigm of Thorne *et al*)
- Extends to arbitrary diffeomorphism-covariant theory of gravity (at cost of complicating the formula for entropy)
- Completely fails to reproduce those features of thermodynamics that concern multiple thermally-interacting systems

Hawking radiation

Hawking radiation

- Given QFT, a BH is surrounded by an almost-thermal atmosphere of radiation, mostly trapped behind the angular-momentum barrier, which gradually radiates away

Hawking radiation

- Given QFT, a BH is surrounded by an almost-thermal atmosphere of radiation, mostly trapped behind the angular-momentum barrier, which gradually radiates away
- Multiple independent derivations (Hawking, Candelas *et al*, Wilczek, etc etc)

Hawking radiation

- Given QFT, a BH is surrounded by an almost-thermal atmosphere of radiation, mostly trapped behind the angular-momentum barrier, which gradually radiates away
- Multiple independent derivations (Hawking, Candelas *et al*, Wilczek, etc etc)
- Completes black hole thermodynamic description by permitting heat transfer between black holes and other TD systems

Hawking radiation

- Given QFT, a BH is surrounded by an almost-thermal atmosphere of radiation, mostly trapped behind the angular-momentum barrier, which gradually radiates away
- Multiple independent derivations (Hawking, Candelas *et al*, Wilczek, etc etc)
- Completes black hole thermodynamic description by permitting heat transfer between black holes and other TD systems
- Very strong theoretical evidence that it carries away energy

Black hole statistical mechanics I

Black hole statistical mechanics I

- QFT calculations in effective-field GR:

Black hole statistical mechanics I

- QFT calculations in effective-field GR:
  - Tree-order calculations reproduce classical entropy for arbitrary diffeomorphism-covariant action

Black hole statistical mechanics I

- QFT calculations in effective-field GR:
  - Tree-order calculations reproduce classical entropy for arbitrary diffeomorphism-covariant action
  - One-loop corrections introduce divergent terms that have exactly the right form to renormalise the gravitational constant, plus finite terms in log M that are negligible in the classical limit

Black hole statistical mechanics I

- QFT calculations in effective-field GR:
  - Tree-order calculations reproduce classical entropy for arbitrary diffeomorphism-covariant action
  - One-loop corrections introduce divergent terms that have exactly the right form to renormalise the gravitational constant, plus finite terms in log M that are negligible in the classical limit
- String theory calculations for certain (near-)extremal black holes: very precise reproduction of entropy formula including higher-derivative and quantum corrections, e.g.:

Black hole statistical mechanics I

- QFT calculations in effective-field GR:
  - Tree-order calculations reproduce classical entropy for arbitrary diffeomorphism-covariant action
  - One-loop corrections introduce divergent terms that have exactly the right form to renormalise the gravitational constant, plus finite terms in $\log M$ that are negligible in the classical limit
- String theory calculations for certain (near-)extremal black holes: very precise reproduction of entropy formula including higher-derivative and quantum corrections, e.g.:

$$S = 2\pi(Q_1 Q_5 N - J^2)(1+3/2N) - \frac{1}{12}(n_v - 3)\ln(Q_1 Q_5 N - J^2).$$

Black hole statistical mechanics II

Calculations via AdS/CFT correspondence:

Black hole statistical mechanics II

Calculations via AdS/CFT correspondence:

- Correspondence itself isn't proven but has very large amount of calculational evidence, e.g. Konishi operator anomalous dimension:

Black hole statistical mechanics II

Calculations via AdS/CFT correspondence:

- Correspondence itself isn't proven but has very large amount of calculational evidence, e.g. Konishi operator anomalous dimension:

$$\gamma_K(a) = 3a$$

Black hole statistical mechanics II

Calculations via AdS/CFT correspondence:

- Correspondence itself isn't proven but has very large amount of calculational evidence, e.g. Konishi operator anomalous dimension:

$$\gamma_K(a) = 3a - 3a^2$$

Black hole statistical mechanics II

Calculations via AdS/CFT correspondence:

- Correspondence itself isn't proven but has very large amount of calculational evidence, e.g. Konishi operator anomalous dimension:

$$\gamma_K(a) = 3a - 3a^2 + \frac{21}{4}a^3$$

Black hole statistical mechanics II

Calculations via AdS/CFT correspondence:

- Correspondence itself isn't proven but has very large amount of calculational evidence, e.g. Konishi operator anomalous dimension:

$$\gamma_K(a) = 3a - 3a^2 + \frac{21}{4}a^3 - \left(\frac{39}{4} - \frac{9}{4}\zeta(3) + \frac{45}{8}\zeta(5)\right)a^4$$

Black hole statistical mechanics II

Calculations via AdS/CFT correspondence:

- Correspondence itself isn't proven but has very large amount of calculational evidence, e.g. Konishi operator anomalous dimension:

$$\gamma_K(a) = 3a - 3a^2 + \frac{21}{4}a^3 - \left(\frac{39}{4} - \frac{9}{4}\zeta(3) + \frac{45}{8}\zeta(5)\right)a^4$$

$$+ \left(\frac{237}{16} - \frac{81}{16}\zeta(3)^2 + \frac{27}{4}\zeta(3) - \frac{135}{16}\zeta(5) + \frac{945}{32}\zeta(7)\right)a^5$$

Black hole statistical mechanics II

Calculations via AdS/CFT correspondence:

- Correspondence itself isn't proven but has very large amount of calculational evidence, e.g. Konishi operator anomalous dimension:

$$\gamma_K(a) = 3a - 3a^2 + \frac{21}{4}a^3 - \left(\frac{39}{4} - \frac{9}{4}\zeta(3) + \frac{45}{8}\zeta(5)\right)a^4$$
$$+ \left(\frac{237}{16} - \frac{81}{16}\zeta(3)^2 + \frac{27}{4}\zeta(3) - \frac{135}{16}\zeta(5) + \frac{945}{32}\zeta(7)\right)a^5$$

- CFT description has right qualitative features to reproduce thermodynamic description of black holes

Black hole statistical mechanics II

Calculations via AdS/CFT correspondence:

- Correspondence itself isn't proven but has very large amount of calculational evidence, e.g. Konishi operator anomalous dimension:

$$\gamma_K(a) = 3a - 3a^2 + \frac{21}{4}a^3 - \left(\frac{39}{4} - \frac{9}{4}\zeta(3) + \frac{45}{8}\zeta(5)\right)a^4$$

$$+ \left(\frac{237}{16} - \frac{81}{16}\zeta(3)^2 + \frac{27}{4}\zeta(3) - \frac{135}{16}\zeta(5) + \frac{945}{32}\zeta(7)\right)a^5$$

- CFT description has right qualitative features to reproduce thermodynamic description of black holes
- In the special case of ADS_3/CFT_2, the *Cardy formula* demonstrates quantitative, exact reproduction

Black hole statistical mechanics II

Calculations via AdS/CFT correspondence:

- Correspondence itself isn't proven but has very large amount of calculational evidence, e.g. Konishi operator anomalous dimension:

$$\gamma_K(a) = 3a - 3a^2 + \frac{21}{4}a^3 - \left(\frac{39}{4} - \frac{9}{4}\zeta(3) + \frac{45}{8}\zeta(5)\right)a^4$$

$$+ \left(\frac{237}{16} - \frac{81}{16}\zeta(3)^2 + \frac{27}{4}\zeta(3) - \frac{135}{16}\zeta(5) + \frac{945}{32}\zeta(7)\right)a^5$$

- CFT description has right qualitative features to reproduce thermodynamic description of black holes
- In the special case of ADS_3/CFT_2, the *Cardy formula* demonstrates quantitative, exact reproduction
- The extremal-hole calcuations can be reinterpreted as ADS_3/CFT_2 calculations

The Quantum Membrane Paradigm (QMP)

The Quantum Membrane Paradigm (QMP)

- A stationary black hole, from the point of view of any observer remaining outside the hole, can be treated *in all respects* as a quantum-mechanical system in the form of a thin membrane, at thermodynamic equilibrium

The Quantum Membrane Paradigm (QMP)

- A stationary black hole, from the point of view of any observer remaining outside the hole, can be treated *in all respects* as a quantum-mechanical system in the form of a thin membrane, at thermodynamic equilibrium
- In particular, the dimension of the energy-$\sim E$ subspace of states of the membrane $\mathcal{H}(E)$ satisfies

$$\mathrm{Dim}\mathcal{H}(E) \propto \exp(A/4) \tag{1}$$

where $A = 16\pi E^2$ is the surface area of a black hole of that energy (simplifying by assuming $J = Q = 0$).

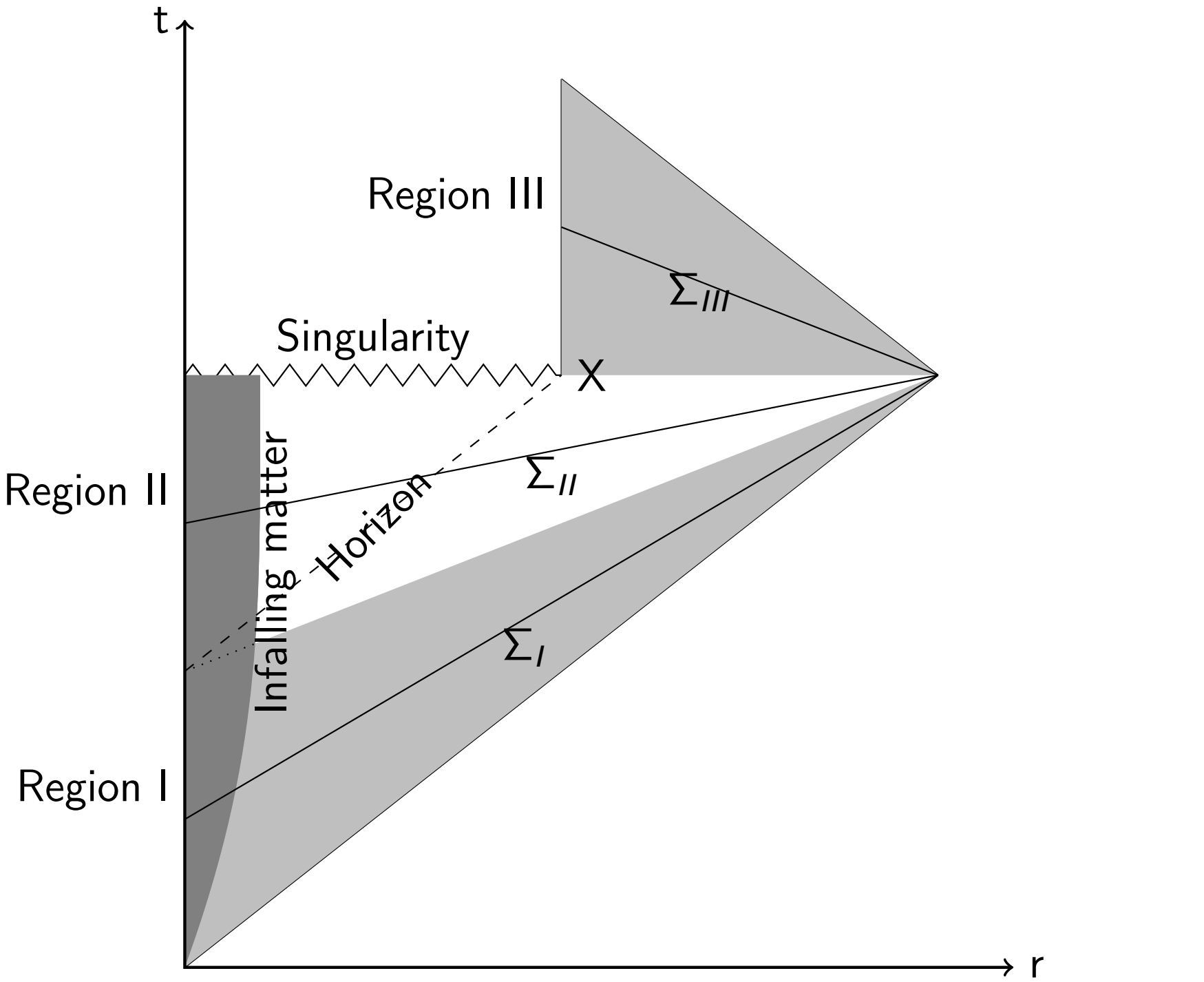
The Quantum Membrane Paradigm (QMP)

- A stationary black hole, from the point of view of any observer remaining outside the hole, can be treated *in all respects* as a quantum-mechanical system in the form of a thin membrane, at thermodynamic equilibrium
- In particular, the dimension of the energy-$\sim E$ subspace of states of the membrane $\mathcal{H}(E)$ satisfies

$$\mathrm{Dim}\mathcal{H}(E) \propto \exp(A/4) \tag{1}$$

where $A = 16\pi E^2$ is the surface area of a black hole of that energy (simplifying by assuming $J = Q = 0$).

- Hawking radiation is ordinary thermal cooling of that membrane


t
r
Region III
Σ_{III}
Singularity
X
Region II
Σ_{II}
Horizon
Infalling matter
Σ_{I}
Region I

The evaporation-time paradox

The evaporation-time paradox

- Information loss is a pretty straightforward consequence of failure of global hyperbolicity on the evaporating-BH spacetime

The evaporation-time paradox

- Information loss is a pretty straightforward consequence of failure of global hyperbolicity on the evaporating-BH spacetime
- No really persuasive arguments *against* information loss in this situation

The evaporation-time paradox

- Information loss is a pretty straightforward consequence of failure of global hyperbolicity on the evaporating-BH spacetime
- No really persuasive arguments *against* information loss in this situation
  - General requirement for QM? (But path integrals and the AQFT framework both seem to make sense)

The evaporation-time paradox

- Information loss is a pretty straightforward consequence of failure of global hyperbolicity on the evaporating-BH spacetime
- No really persuasive arguments *against* information loss in this situation
  - General requirement for QM? (But path integrals and the AQFT framework both seem to make sense)
  - Mechanisms to avoid information loss turn on details of Planck-scale quantum gravity and so are highly speculative at best
- Non-unitarity leads to Planck-scale energy dissipation in flat space? (Controversial, speculative)

The evaporation-time paradox

- Information loss is a pretty straightforward consequence of failure of global hyperbolicity on the evaporating-BH spacetime
- No really persuasive arguments *against* information loss in this situation
  - General requirement for QM? (But path integrals and the AQFT framework both seem to make sense)
  - Mechanisms to avoid information loss turn on details of Planck-scale quantum gravity and so are highly speculative at best
- Non-unitarity leads to Planck-scale energy dissipation in flat space? (Controversial, speculative)

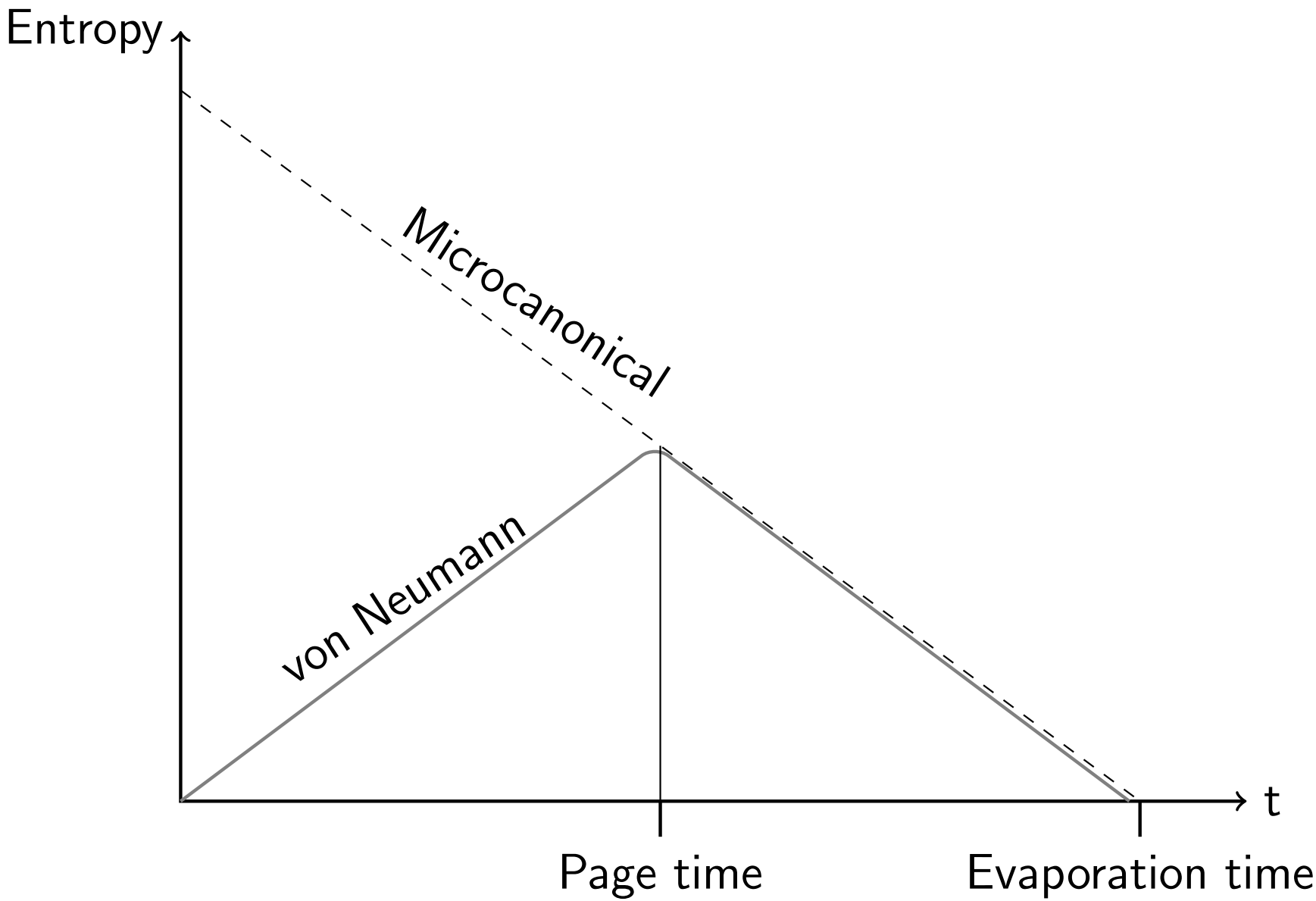


Figure: The Page curve

The Page-time paradox

The Page-time paradox

- If BHs are thermodynamic objects (more accurately: if we can assume QMP), and if BH radiation is exactly thermal:

The Page-time paradox

- If BHs are thermodynamic objects (more accurately: if we can assume QMP), and if BH radiation is exactly thermal:
  - BH *von Neumann* entropy increases the more radiation is emitted

The Page-time paradox

- If BHs are thermodynamic objects (more accurately: if we can assume QMP), and if BH radiation is exactly thermal:
  - BH *von Neumann* entropy increases the more radiation is emitted
  - BH *microcanonical* entropy ($=A/4$) decreases as the black hole energy decreases

The Page-time paradox

- If BHs are thermodynamic objects (more accurately: if we can assume QMP), and if BH radiation is exactly thermal:
  - BH *von Neumann* entropy increases the more radiation is emitted
  - BH *microcanonical* entropy ($=A/4$) decreases as the black hole energy decreases
- But von Neumann entropy is bounded by microcanonical entropy, so the radiation must eventually stop being thermal

The Page-time paradox

- If BHs are thermodynamic objects (more accurately: if we can assume QMP), and if BH radiation is exactly thermal:
  - BH *von Neumann* entropy increases the more radiation is emitted
  - BH *microcanonical* entropy ($=A/4$) decreases as the black hole energy decreases
- But von Neumann entropy is bounded by microcanonical entropy, so the radiation must eventually stop being thermal
- The *Page time* is the point when the radiation must stop being thermal — it's approximately half the BH total evaporation time, i. e. long before Planck-level curvature is reached

The Page-time paradox

- If BHs are thermodynamic objects (more accurately: if we can assume QMP), and if BH radiation is exactly thermal:
  - BH *von Neumann* entropy increases the more radiation is emitted
  - BH *microcanonical* entropy ($=A/4$) decreases as the black hole energy decreases
- But von Neumann entropy is bounded by microcanonical entropy, so the radiation must eventually stop being thermal
- The *Page time* is the point when the radiation must stop being thermal — it's approximately half the BH total evaporation time, i. e. long before Planck-level curvature is reached
- Clash between QMP and QFT

Information loss for non-evaporating black holes

Evidently whatever the information loss paradox is supposed to be, it does not arise for a non-evaporating black hole. (Maudlin)

Information loss for non-evaporating black holes

Evidently whatever the information loss paradox is supposed to be, it does not arise for a non-evaporating black hole. (Maudlin)

Information loss for non-evaporating black holes

> *Evidently whatever the information loss paradox is supposed to be, it does not arise for a non-evaporating black hole. (Maudlin)*

But the Page time paradox arises (and has been discussed) for:

Information loss for non-evaporating black holes

> *Evidently whatever the information loss paradox is supposed to be, it does not arise for a non-evaporating black hole. (Maudlin)*

But the Page time paradox arises (and has been discussed) for:

- Large (stable) black holes in AdS spacetime, where QFT predicts exponential decay of correlation functions in time, and this contradicts general results about discrete quantum systems;

Information loss for non-evaporating black holes

> *Evidently whatever the information loss paradox is supposed to be, it does not arise for a non-evaporating black hole. (Maudlin)*

But the Page time paradox arises (and has been discussed) for:

- Large (stable) black holes in AdS spacetime, where QFT predicts exponential decay of correlation functions in time, and this contradicts general results about discrete quantum systems;
- Extremal (charged and/or rotating) black holes, perturbed away from extremality and then allowed to decay back down again, where QMP predicts the decay radiation is in a pure state

Evidence for unitarity from AdS/CFT I

Evidence for unitarity from AdS/CFT I

- Far from a BH, Hawking radiation is quite far from black-body radiation - *grey-body* factors parametrise this and can be calculated by solving the wave equation on the BH background

Evidence for unitarity from AdS/CFT I

- Far from a BH, Hawking radiation is quite far from black-body radiation - *grey-body* factors parametrise this and can be calculated by solving the wave equation on the BH background
- For certain near-extremal black holes, these grey-body factors can be calculated explicitly (they have a fairly complicated form)

Evidence for unitarity from AdS/CFT I

- Far from a BH, Hawking radiation is quite far from black-body radiation - *grey-body* factors parametrise this and can be calculated by solving the wave equation on the BH background
- For certain near-extremal black holes, these grey-body factors can be calculated explicitly (they have a fairly complicated form)
- Maldacena and Strominger reproduced these grey-body factors *exactly* in the (manifestly unitary) quantum decay of the excited string state corresponding to the near-extremal BH (retrospectively the calculation can be reinterpreted as an AdS/CFT calculation)

Evidence for unitarity from AdS/CFT I

- Far from a BH, Hawking radiation is quite far from black-body radiation - *grey-body* factors parametrise this and can be calculated by solving the wave equation on the BH background
- For certain near-extremal black holes, these grey-body factors can be calculated explicitly (they have a fairly complicated form)
- Maldacena and Strominger reproduced these grey-body factors *exactly* in the (manifestly unitary) quantum decay of the excited string state corresponding to the near-extremal BH (retrospectively the calculation can be reinterpreted as an AdS/CFT calculation)

The string decay rates, d to the large black hole region, agree precisely with the semiclassical Hawking decay rates in a wide variety of circumstances. However, the string method not only supplies the decay rates, but it also gives a set of unitary amplitudes underlying the rates. We find it tempting to conclude that these extrapolated amplitudes are also correct. It is hard to imagine a mechanism which corrects the amplitudes, but somehow conspires to leave the rates unchanged.

The string decay rates, d to the large black hole region, agree precisely with the semiclassical Hawking decay rates in a wide variety of circumstances. However, the string method not only supplies the decay rates, but it also gives a set of unitary amplitudes underlying the rates. We find it tempting to conclude that these extrapolated amplitudes are also correct. It is hard to imagine a mechanism which corrects the amplitudes, but somehow conspires to leave the rates unchanged. The robust nature of the string picture is very significant because it allows us to directly confront the black hole information puzzle . . . According to Hawking information is lost as a large excited black hole decays to extremality. On the other hand the string analysis . . . gives a manifestly unitary answer. (Maldacena and Strominger; emphasis theirs)

Evidence for unitarity from AdS/CFT II

- For a large black hole in AdS space, multi-time correlation functions outside the BH initially decay exponentially but then oscillate randomly and eventually return to their original value, as per the QMP description

Evidence for unitarity from AdS/CFT II

- For a large black hole in AdS space, multi-time correlation functions outside the BH initially decay exponentially but then oscillate randomly and eventually return to their original value, as per the QMP description
- For a small black hole in AdS space formed by infalling matter, the Poincaré recurrence theorem ensures that the BH will eventually decay completely.

Firewalls

Firewalls

- From the point of view of a freely falling observer, nothing very special happens at the horizon

Firewalls

- From the point of view of a freely falling observer, nothing very special happens at the horizon
- This requires modes of the radiation field just inside, and just outside, the horizon to be strongly entangled

Firewalls

- From the point of view of a freely falling observer, nothing very special happens at the horizon
- This requires modes of the radiation field just inside, and just outside, the horizon to be strongly entangled
- But after the Page time, if black hole radiation is a unitary process then the just-outside modes must be maximally entangled with radiation emitted much earlier

Firewalls

- From the point of view of a freely falling observer, nothing very special happens at the horizon
- This requires modes of the radiation field just inside, and just outside, the horizon to be strongly entangled
- But after the Page time, if black hole radiation is a unitary process then the just-outside modes must be maximally entangled with radiation emitted much earlier
- Monogamy of entanglement makes these two requirements inconsistent

Daniel Harlow

Thus we find ourselves in the enviable situation of having an interesting problem with no really satisfying answer; if we are lucky this means that we will learn something deep. (arxiv:1409.1231)