On information paradox and the fate of black holes

A.Morozov

MIPT, ITEP & IITP, Moscow, Russia

ABSTRACT

A sketchy review of the "island" paradigm in black hole evaporation theory, which actually brings us back to the old idea that interior of black hole decouples from our universe after Page time, so that Hawking radiation is entangled with emerging new universe, thus leaving no room for the information paradox. Instead this provides a self-consistent description of multiverse, where every black hole in a parent universe is a white hole – the origin – of a new one.

Information paradox is a long-standing puzzle in the theory of black hole (BH) evaporation [1]. Recently it attracted a new attention [2] because of the new source of evidence, provided by AdS/CFT-correspondence. Surprisingly or not, it brought some of us back to the old point of view, that there is no paradox at all. This is a brief review of what we understand about the present state of affairs in this field.

The essence of the paradox [3] is that starting from a *pure* quantum state we can go through a formation of black hole to its complete evaporation, leaving one with a thermal ensemble for emitted Hawking radiation – what drastically contradicts unitarity of quantum evolution, which can not convert *pure* states into *mixed*.

There was a variety of radical attempts to resolve the paradox, going deep into different projects of quantum gravity. However, in full accordance with Okkam's razor principle, one does not need to go so far – the paradox is resolved at the same level as it is formulated, i.e. without any reference to quantum gravity.

1 The basic scenario

The old/new resolution is that the final state is very different from above assumption. Namely, at the Page time [4] a part of the black hole interior starts to decouple from our universe, forming the Wheeler's "bag of gold" [5] or a wormhole to a new universe. In modern language this is called an *island* inside the black hole, the domain which can not be seen in our universe, even by entropic considerations. At the very end, the "observable" part of black hole fully disappears (evaporates), and thus entanglement entropy of this "nothing" with radiation vanishes – and the true entanglement is between *thermal* Hawking radiation **in** our universe and the island **beyond** it. Together they continue to be in the *pure* state, but separately they look *mixed*. The archetypical example is the *pure* thermodouble state

$$\rho_{TD} = |\Psi_{TD}\rangle \langle \Psi_{TD}| \tag{1}$$

with

$$|\Psi_{TD}\rangle = \sum_i e^{-E_i/2T} |i\rangle_I \otimes |i\rangle_R \tag{2}$$

and *mixed* states

$$\rho_R = \text{Tr}_I \, \rho_{TD} = \sum_i e^{-E_i/T} |i\rangle_R \langle i|_R \tag{3}$$

$$\rho_I = \text{Tr}_R \, \rho_{TD} = \sum_i e^{-E_i/T} |i\rangle_I \langle i|_I \tag{4}$$

This interpretation of BH evaporation implies that every black hole in a given universe is a white-hole source of a new one, thus providing a tree model of multiverse, see s.3 below. Each decoupling *bag of gold = black hole for us and a white-hole for some emerging new world* has all the quantum numbers (charges), associated with gauge symmetries, vanishing, but can have non-trivial entanglement with the parent universe and also non-trivial values of ungauged charges, like baryon number. what can provide new options to resolve the baryon asymmetry problem.
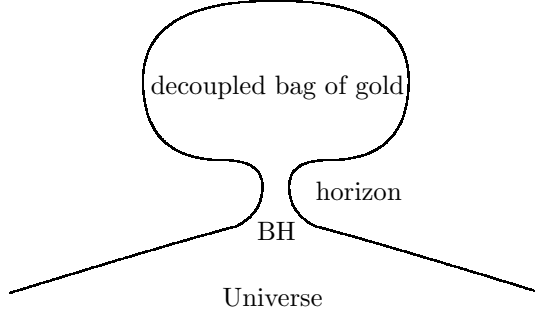
Figure 1: Wheeler's "bag of gold" is a solution to Einstein equations where interior of the black hole is substituted by geometry with a fixed area (equal to horizon area) and arbitrarily large volume. We need a quasiclassical treatment of this kind of geometries, when they coexist with Hawking radiation around in the main part of the Universe, outside the Schwarzschild radius.

2 Relevant theses

The following comments are very simple, still they can be useful for thinking about the information paradox.

- Entropy growth is consistent with unitary evolution.

  Of course, if ρ represents the state of the whole system, then

  $$\mathrm{tr}U\rho U^\dagger \log(U\rho U^\dagger) = \mathrm{tr}\rho\log\rho \tag{5}$$

  i.e. the entropy does not change. However, the entropy of some part 1 of the system $1 \cup 2$, which is described by $\rho_1 = \mathrm{tr}_2\rho$, is not conserved:

  $$\rho' = U\rho U^\dagger, \qquad\qquad \rho_1' = \mathrm{tr}_2\,\rho'$$
  $$\mathrm{tr}\rho'\log\rho' = \mathrm{tr}\rho\log\rho, \qquad\qquad \mathrm{tr}_1\rho_1'\log\rho_1' \neq \mathrm{tr}_1\rho_1\log\rho_1. \tag{6}$$

  Thus, entropy growth of a part is consistent with unitary evolution of the whole system.

  - In particular, burning a fire is consistent with unitary evolution.
    The whole system containing the fire and its radiation is assumed to evolve unitary. However, when we trace out information that runs away with the radiation, the entropy of remaining system (fire) can – and it does so.

- Entropies of both subsystems 1 and 2 can increase simultaneously, while entire system remains in the pure state, i.e. has vanishing entropy: see (12) below.

- Entropy and entanglement entropy are intimately related notions.

  - Entropy of a subsystem can be defined only as entanglement entropy.
    If an incompleteness of our knowledge about a subsystem is only due to absence of knowledge about complementary subsystem, the entropy of subsystem coincides with entanglement entropy. In general, entanglement entropy is a lower bound on entropy of subsystem.

  - Ordinary entropy is equal to entanglement entropy *with* a pure state.
    Alternatively it may seem that "entanglement entropy *with* a pure state" is an ill defined object, and entanglement entropy is well-defined only in the case when the state of whole system is pure. For example, tensor product of the density matrix of a mixed state and that of a pure state is not pure:

    $$\tilde\rho = \rho \otimes \big(|\psi\rangle\langle\psi|\big), \qquad \mathrm{tr}\tilde\rho\log\tilde\rho = \mathrm{tr}\rho\log\rho \neq 0 \tag{7}$$

  As a consequence entropy of reduced density matrices are not equal to each other

  $$\rho_1 = \mathrm{tr}_2\tilde\rho = \rho, \qquad \mathrm{tr}_1\rho_1\log\rho_1 \neq 0$$
  $$\rho_2 = \mathrm{tr}_1\tilde\rho = |\psi\rangle\langle\psi|, \qquad \mathrm{tr}_2\rho_2\log\rho_2 = 0 \tag{8}$$

– If entire system is in pure state, entanglement entropies of its parts sum to zero.
This depends on the meaning of "sum to zero". The situation is the following

1. Take the whole system in pure state, whose parts 1 and 2 are (possibly) entangled

$$|\psi\rangle = \sum_i c_i |i\rangle_1 |\tilde{i}\rangle_2 \tag{9}$$

2. Now, forget about (for example) subsystem 2, so that we see the subsystem 1 in the state

$$\rho_1 = \text{tr}_2 |\psi\rangle\langle\psi| = \sum |c_i|^2 |i\rangle_1 \langle i|_1 \tag{10}$$

Incompleteness of our knowledge about subsystem 1 results into the entropy

$$S(\rho_1) = -\text{tr}_1 \rho_1 \log \rho_1 \tag{11}$$

3. Moreover, incompleteness of our knowledge about the subsystem 2 after forgetting about the subsystem 1 equals to incompleteness of our knowledge about the subsystem 1 after forgetting about the subsystem 2

$$S(\rho_1) = S(\rho_2) \tag{12}$$

4. Thus, if one remembers everything about complement subsystem the entropies of the subsystem "collapse into zero".

Note: The only role of the word "entanglement" before "entropy" is to emphasize that the *nature* of incompleteness is that the whole system state was entangled state, namely it's subsystems 1 and 2 were somehow correlated.

- Complement of the thermal state to a pure one should be at least as big:

$$\left(\sum_{i=1}^N e^{-\beta E_i} |i><i| \right) \otimes \left(\sum_{a=1}^n u_a |a><a| \right) = |\psi><\psi| \implies n \geq N \tag{13}$$

The easiest way to purify a mixed state

$$\rho = \sum_i p_i |i\rangle\langle i|, \qquad \langle i|j\rangle = \delta_{ij} \tag{14}$$

is the following. Take the double-copied Hilbert space $\mathcal{H} \otimes \mathcal{H}$ and consider the pure state

$$|\psi\rangle = \sum_i \sqrt{p_i} \, |i\rangle \otimes |i\rangle \tag{15}$$

Tracing of corresponding density matrix $\rho_{\text{pure}} = |\psi\rangle\langle\psi|$ over the second copy of the Hilbert space leads to the initial density matrix

$$\text{tr}_{\text{copy}} \rho_{\text{pure}} = \sum_{i,j} \sqrt{p_i p_j} \, \langle j|i\rangle \, |i\rangle\langle j| = \sum_i p_i |i\rangle\langle i| = \rho \tag{16}$$

In the case of thermodynamic Gibbs distribution density matrix

$$\rho_{\text{th}} = \sum_i e^{-\beta E_i} |i\rangle\langle i| \tag{17}$$

the purified state is so-called thermofield double state

$$|TFD\rangle = \sum_i e^{-\beta E_i/2} |i\rangle_1 \otimes |i\rangle_2, \qquad \text{tr}_2 |TFD\rangle\langle TFD| = \rho_{\text{th}} \tag{18}$$

which is said to be holographic dual to eternal AdS-Schwarzschild black hole.

In the case, where some of probabilities in (14) vanish, dimension of additional Hilbert space, used for the state purification, can be reduced. Indeed, introducing Hilbert space

$$\tilde{\mathcal{H}} = \text{Span}(|\tilde{i}\rangle), \quad \tilde{i} = 1, \dots (\# \text{ of } p_i \neq 0) \tag{19}$$

one can purify (14) as

$$|\psi\rangle = \sum_{i:\, p_i \neq 0} \sqrt{p_i} \, |i\rangle \otimes |\tilde{i}\rangle. \tag{20}$$

3

- Evaporation of one-half of the black hole content defines the Page time – and this is the last moment at which its interior (back of gold or "island") *should* decouple. Of course, this can happen earlier, but this depends on the still unknown details of the black hole evolution. It is even unclear what are the factors, which can affect the answer to this question.

- Horizon area calculates only a "part" of the states inside the black hole. These are the states "visible" from the outside and their number decreases to zero in the process of evaporation.

- Essential part of evaporating black hole decouples from our universe at around Page time (as a "bag of gold" or as a wormhole to somewhere else). This can be interpreted as formation of an *island* [2].

- Their can be various assumptions about the fate of evaporating black hole, we mention just a few:

  - It partly disappears from our universe and can become an origin of a new one.
  - Absorbed *energy* is fully radiated back to us, new universe is born with net zero energy.
  - Absorbed information is partly transferred to the new universe.
  - Perhaps, it is born with non-vanishing information *or* it is born, with vanishing info, but entangled with our universe.
  - Entropy of our universe gets increased after evaporation.
  - A new universe is born with vanishing entropy.

  It is not clear which of these theses are fully correct and what are the factors which can regulate the possible deviations.

3   History of the multiverse

Since in this scenario Hawking radiated is only a part (at most half) of what falls into the black hole, its interior remains non-trivial and decouples from our Universe. Afterwards it lives its own life, i.e. becomes a separate new universe. If its history is similar to ours, we get a typical multiverse pattern:
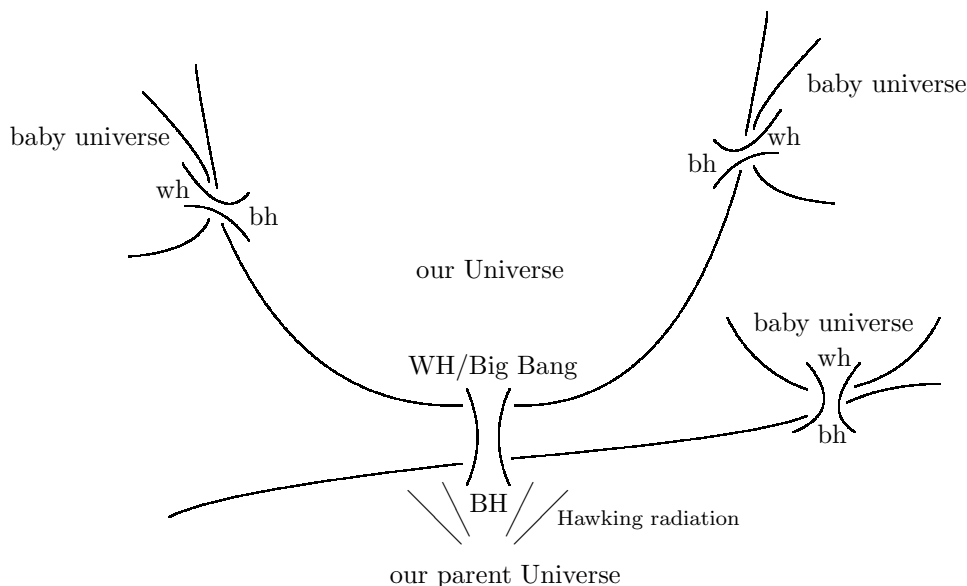

Figure 2: Schematic picture of the history of Universe. It is created as a white hole from a black hole in the parent Universe. Other black holes give rise to other baby universes – some of them are sisters of our, others are descendants and so on. Each black hole is surrounded by Hawking radiation, which stays in its own universe, while the other content of the black hole decouples from it at around the Page time and gives birth to a new baby universe. Inside the new universe the origin look like a white hole or a Big Bang.

4   Open questions

It is an almost unavoidable prediction of General Relativity that every universe has a simple white hole at its origin (big bang) and many black holes, formed during its evolution — with each black hole giving birth to a new baby/descendant universe with its own fate.
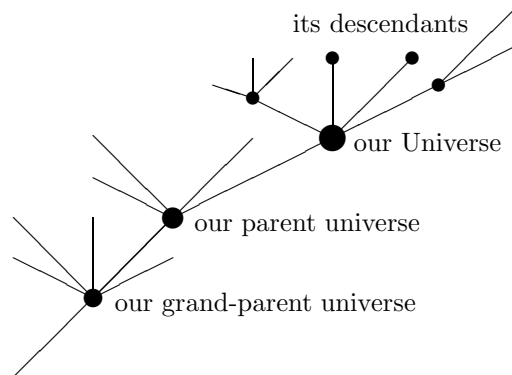

Figure 3: Another schematic picture of the multiverse. Black dots show ancestors of our Universes, lines are the black-white hole transitions between the universes. Many other universes, created from various black holes and their further descendants are not shown. Each universe has a single white-hole origin and many black holes, which give rise to many descendent baby-universes. Some of them can collapse before giving birth to anything new, some can have a lot of descendants.

This *prediction* should be distinguished from an additional *possibility* that there can be wormholes, some looking like black-white hole pairs, which can connect different universes or different regions in the same universe — these are still only speculations, depending on various assumptions beyond General Relativity *per se*.

The main technical problem is the quasiclassical description of Wheeler's bag-of-gold geometries, which would provide a consistent description of Hawking radiation and new-born baby-universes. In particular, this should answer the question, which part of matter, falling into the black hole, is actually radiated (at most half, but it can actually be much smaller) and what are the parameters, which regulate this fraction. We hope that more attention will be paid to critical analysis of this kind of naive scenario for black hole evaporation and this will finally lead to resolution of paradoxes and to healthy theories of multiverse dynamics.

Acknowledgements

References

[1] S. W. Hawking, *Particle Creation by Black Holes*, Commun. Math. Phys. 43 (1975) 199–220

[2] G.Penington, arXiv:1905.08255
A.Almheiri, N.Engelhardt, D.Marolf, and H.Maxfield, arXiv:1905.08762
A. Almheiri, R. Mahajan, and J. Maldacena, arXiv:1910.11077
G. Penington, S. H. Shenker, D. Stanford, and Z. Yang, arXiv:1911.11977
G.Grimaldi, J.Hernandez and R.C.Myers, JHEP 03 (2022) 136, arXiv:2202.00679
and many more references therein

[3] S. W. Hawking, *Breakdown of Predictability in Gravitational Collapse*, Phys. Rev. D14 (1976) 2460–2473
S. D. Mathur, Class. Quant. Grav. 26 (2009) 224001, arXiv:0909.1038; arXiv:1108.0302

[4] D. N. Page, Phys. Rev. Lett. 71 (1993) 1291; Phys. Rev. Lett. 71 (1993) 3743

[5] J. A. Wheeler, in Relativity, Groups, and Fields, edited by B. S. DeWitt and C. M. DeWitt (Gordon and Breach, New York, 1964)
S. D. H. Hsu and D. Reeb, *Unitarity and the Hilbert space of quantum gravity*, arXiv:0803.4212
D.Marolf, *Black Holes, AdS, and CFTs. General Relativity and Gravitation*, 41 (2009) 903, arXiv:0810.4886