Quantum Gravity and Strings

Joseph Polchinski
Kavli Institute for Theoretical Physics
UC Santa Barbara

A Century of General Relativity
Harnack House, Berlin, Nov. 30, 2015

1915: General Relativity

1899: Planck's units

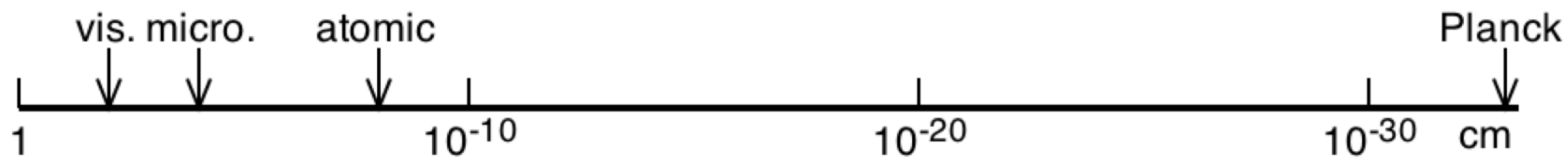
From $\hbar$, c, and G one can define a natural system of units:

Planck length: $l_P = \sqrt{\hbar G/c^3} = 1.6\times 10^{-33}\,\mathrm{cm}$

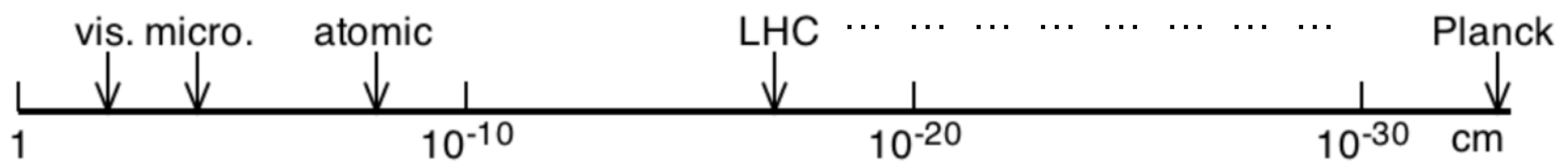
Planck time: $t_P = \sqrt{\hbar G/c^5} = 5.4\times 10^{-44}\,\mathrm{sec}$

Planck mass: $M_P = \sqrt{\hbar c/G} = 2.1\times 10^{-5}\,\mathrm{g}$

These necessarily retain their meaning for all times and for all civilizations, even extraterrestrial and non-human ones, and can therefore be designated as *natural units*.



Eddington: this length “must be the key to some essential structure.”



16 orders of magnitude still to go.

Theory has sometimes been able to leap a smaller gap (thought experiments!):

Maxwell, Dirac (antiparticles), the Higgs.

So we should try.

Strings and quantum gravity

I. Successes

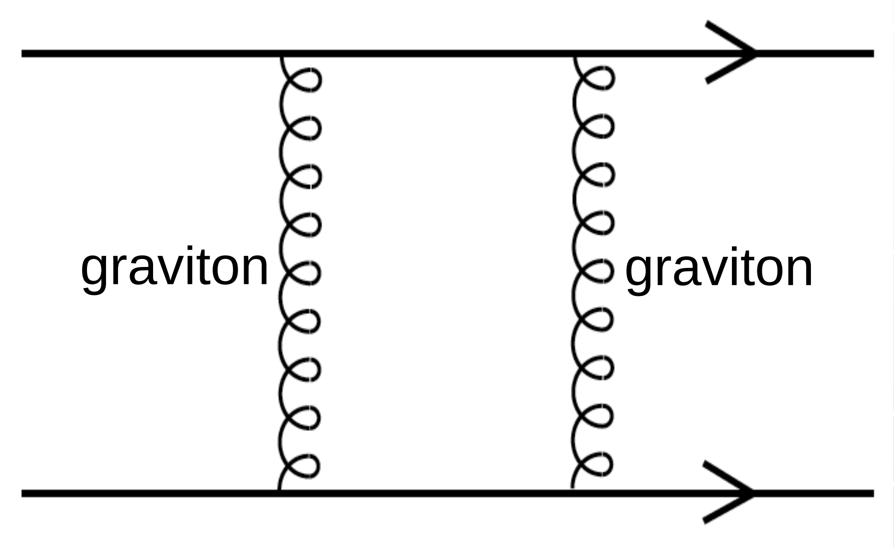
- The short distance problem
- Uniqueness of dynamics
- Physics from geometry
- Duality between gauge fields and strings

II. The multiverse and the cosmological constant

III. Black holes, firewalls, entanglement

The short distance problem

What do existing theories (GR + QM) say about short distance?



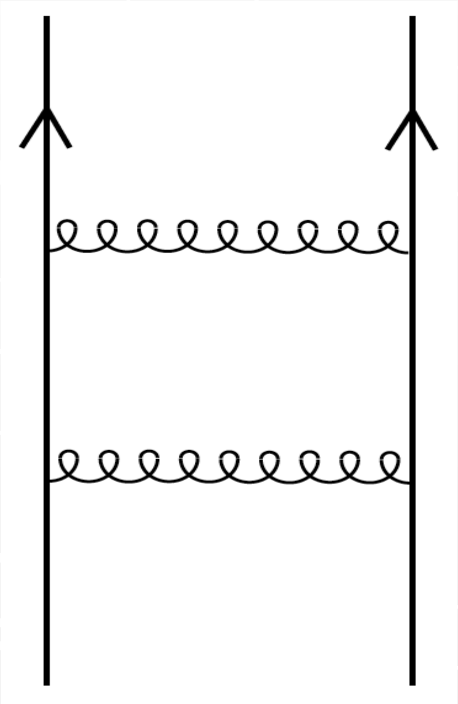
Infinities! Nonrenormalizable (= spacetime foam)

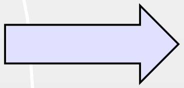
Dimensional analysis, $[G] = m^{-2}$ so GE^2 is dimensionless.

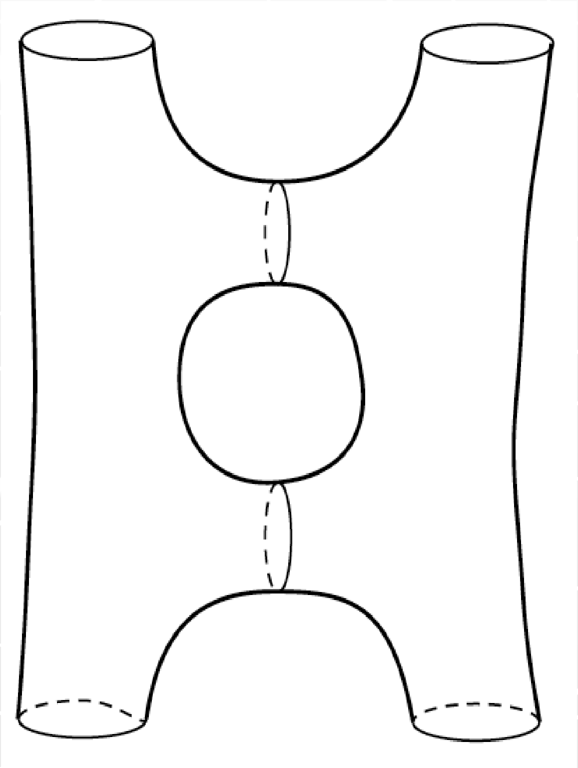
Not unique to GR. In the Fermi theory of the weak interaction, $[G_F] = m^2$. This clue $\rightarrow$ W, Z, higgs.

Finiteness + Lorentz invariance + unitarity/causality is strongly constraining.

For gravity, here's one thing that works:







Uniqueness of dynamics

Einstein's equation: $R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu}$

LHS determined by equivalence principle. But:

- What about RHS?
- Cosmological constant
- Higher derivative (e.g. $R_{\mu\nu\sigma\rho}{}^{2,3,4,\square}$) terms

Uniqueness of dynamics

Einstein's equation: $R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu}$

LHS determined by equivalence principle. But:

What about RHS?

Cosmological constant

Higher derivative (e.g. $R_{\mu\nu\sigma\rho}{}^{2,3,4,\ldots}$) terms

A theory whose dynamical equations are fully determined by general principle is an attractive goal, but as this example shows it is hard to attain.

String theory does this.

In string theory, the dynamical equations are fully determined by general principle.

Full disclosure: we do not know what the equations are, or what the principle is. String theory is a discovery in progress.

One approximate form:

$$T_{ab} = 0$$

Principle: world-sheet (two-dimensional) conformal invariance).

Physics from geometry

In GR, gravity is the curvature of spacetime. So it is an attractive idea that the other interactions come from the geometry of spacetime as well. But GR seems to have used up the spacetime that we know, so we need more spacetime!

String theory requires extra dimensions, plus branes, of a form nicely compatible with the rest of physics.

Duality between gauge fields and strings

Some important equations:

Einstein (1915): $R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu}$

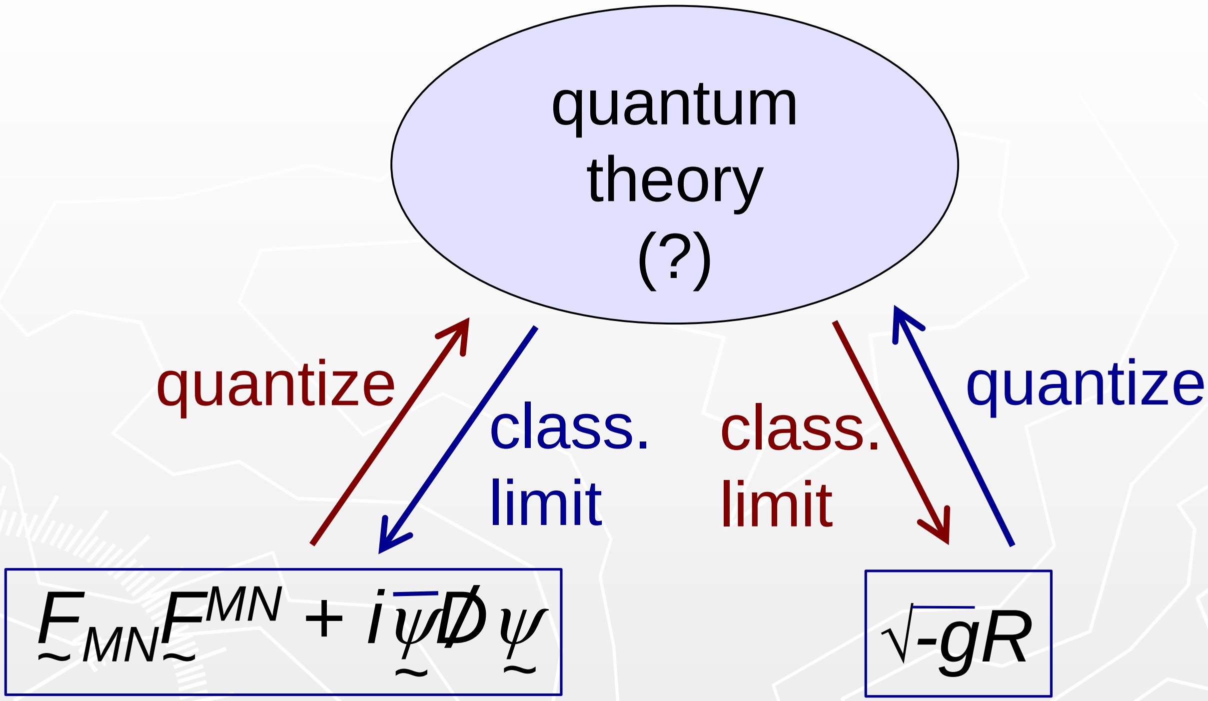
Maxwell (~1861): $\partial^\mu F_{\mu\nu} = j_\nu$

Yang-Mills (1954): $D^\mu \underset{\sim}{F}_{\mu\nu} = \underset{\sim}{j}_\nu$

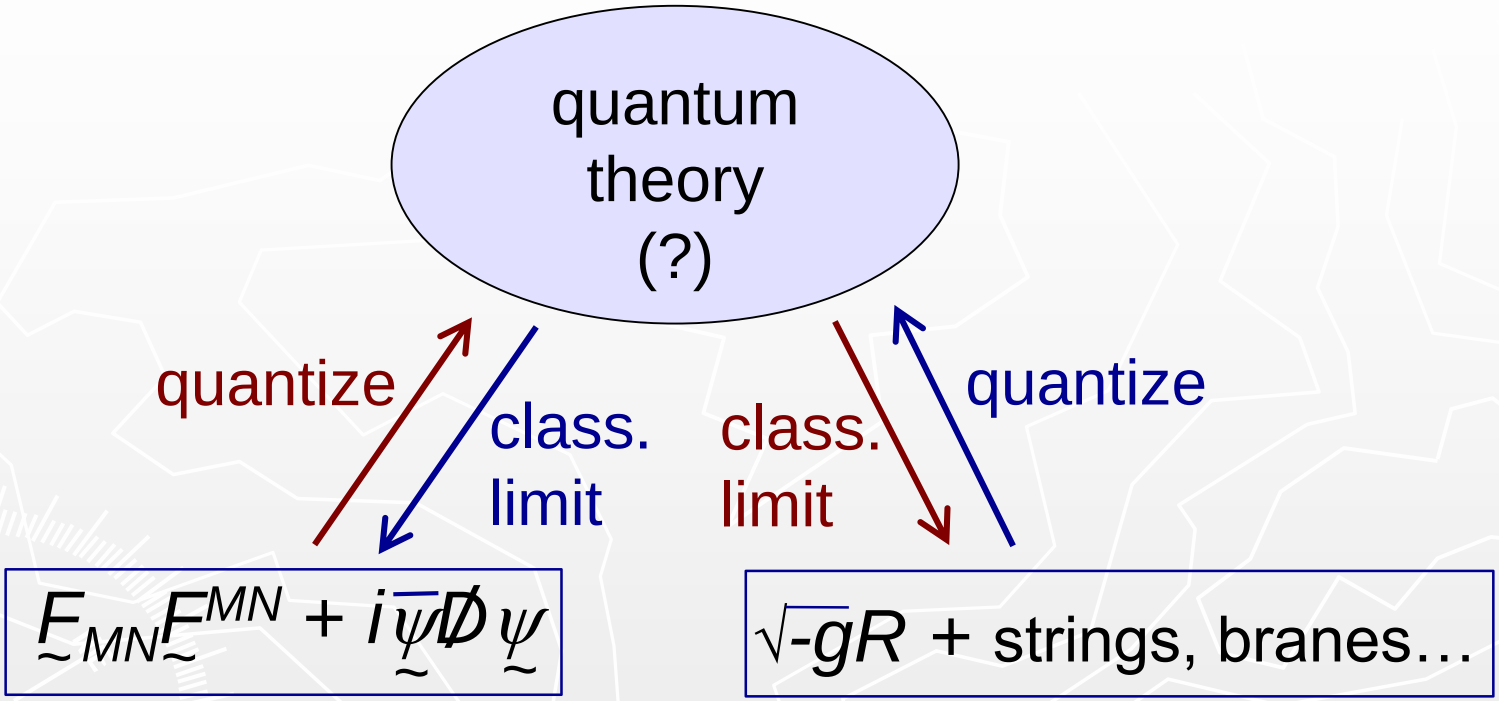
Dirac (1928): $i\not{D}\psi = m\psi$

How are these related?

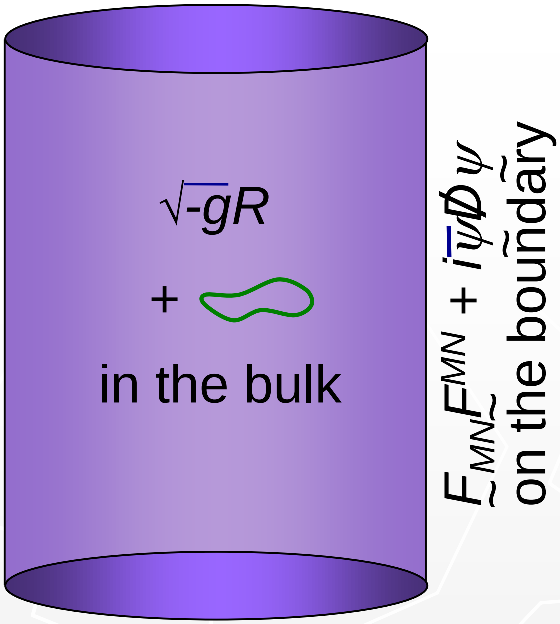
The great surprise:



The great surprise:



Holographic:



- GR, strings, branes, bulk spacetime are emergent.
- Gauge degrees of freedom are highly nonlocal from the bulk point of view.
- Limited, for now, to special boundary conditions. High energy scattering, black hole evaporation, some topology change are included. Cosmology, not.

Summary, part I

- The short distance problem
- Uniqueness of dynamics
- Physics from geometry
- Duality between gauge fields and strings

IIa. The multiverse

The dark side of *physics from geometry*: If the physics that we see depends on the shape of extra dimensions, what determines that shape?

How many solutions does the 10D Einstein equation have that look minimally like our world
(4D Minkoski x 6D compact)?

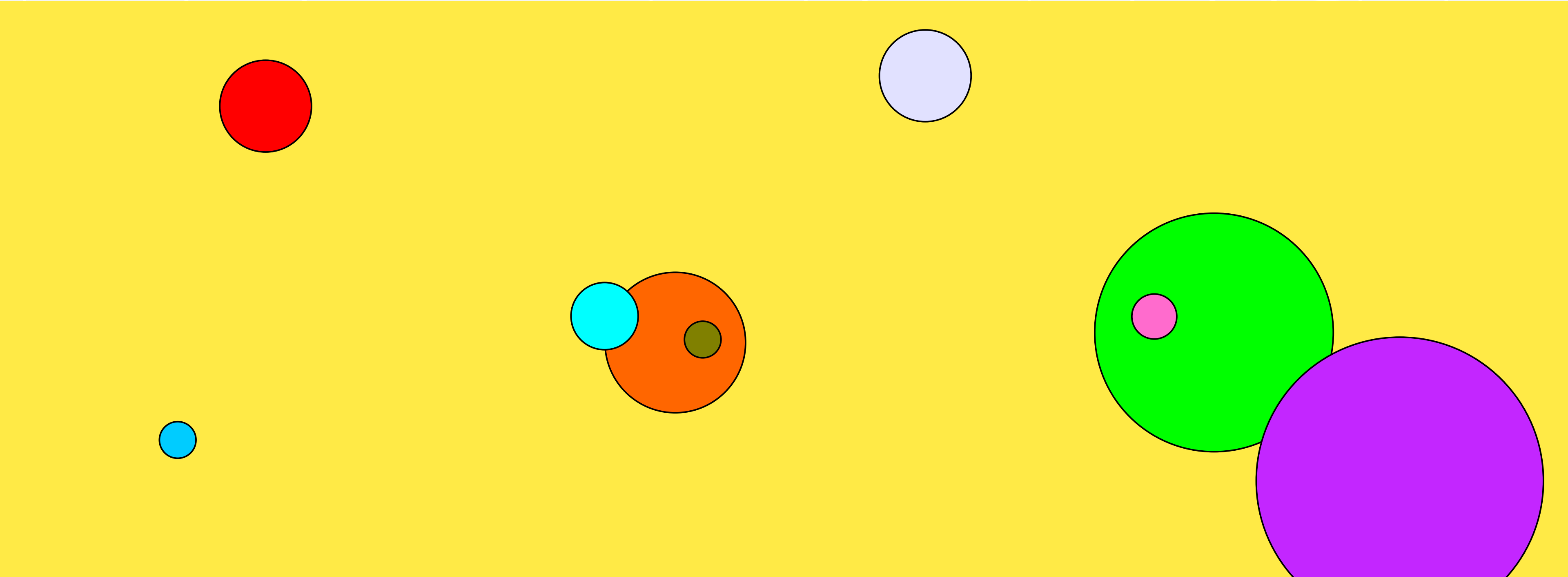
Addressed by Calabi, Yau: perhaps $O(10^{9})$ vacuum solutions (times moduli), combinatorically more with matter. $O(10^{500})$? $O(10^{272,000})$? The landscape.

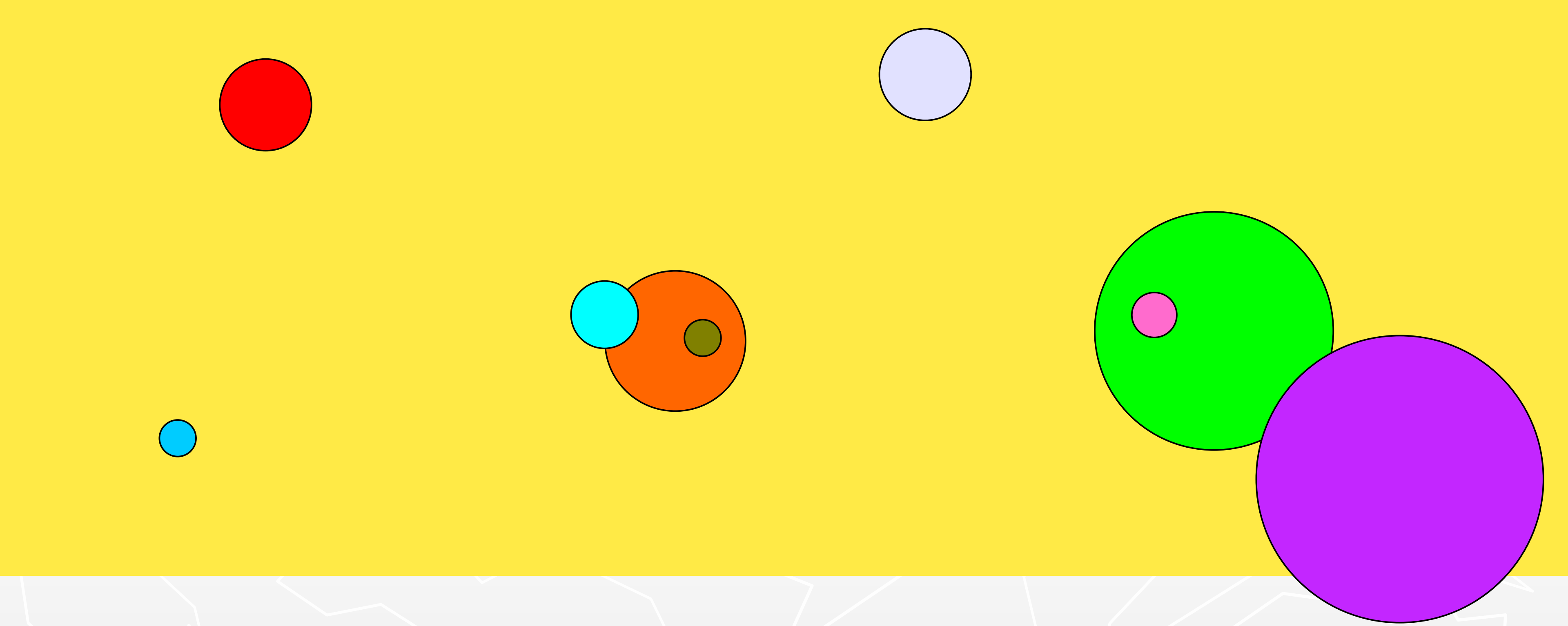
Einstein encounted a tiny version of the landscape, the radius of the Kaluza-Klein circle.

If there is a landscape, what determines the geometry, and therefore the physics, near us? Initial conditions?

Dynamics erases much of the initial conditions: given and positive c.c. solution, we get *expansion*, then localized *tunneling* to a new vacuum, and repeat:

GR + QM + landscape → multiverse





A multiverse seems to be the price we pay for *physics from geometry*, and also for *uniqueness of dynamics*.

Yet there is evidence that we live in just such a universe.

IIb. The cosmological constant

The modern point of view: the vacuum is a complicated place, with zero point energies, Higgs fields, quark condensates, color fluctuations.

All of these have direct physical effects, and all should gravitate. Why don't they?

Various kinds of theory:

Theories that predict a definite value for the c.c.:

No supersymmetry, enormous c.c.

Unbroken supersymmetry, zero c.c.

No supersymmetry, zero c.c.

Theories that do not predict a definite value for the c.c.

Theories in which the c.c is a free parameter in the dynamical equations.

The multiverse.

Various kinds of theory:

Theories that predict a definite value for the c.c.:

No supersymmetry, enormous c.c.

conflicts with observation

Unbroken supersymmetry, zero c.c.

conflicts with observation

No supersymmetry, zero c.c.

not yet found, despite much effort

Theories that do not predict a definite value for the c.c.

Theories in which the c.c is a free parameter in the dynamical equations.

why small c.c. (fine tuning?). ultimately → 1,2,3, or 5.

The multiverse.

The multiverse:

The formation of complex structures requires

- Many degrees of freedom
- Large volume
- Long times

These will only be available in bubbles where the c.c. is far below its `natural' scale (Weinberg, following Sakurai, Banks).

Of the five kinds of theory, only the multiverse predicts that observers see an unnaturally small c.c.

But these conditions do not require c.c. = 0 identically, so a small nonzero value is predicted (Weinberg).

It is often said that the discovery of the `dark energy' was a complete surprise, but it was expected by two groups:

• Those who paid attention to the data without theoretical prejudice.

• Those who had tried hard to solve the c.c. problem by conventional means, knew how hard it was, and were aware of Weinberg's prediction.

It is often said that multiverse is unpredictive, but in a sense the opposite is true:

The c.c. is arguably the most important discovery about the nature of spacetime since Hubble (alongside the CMB), and only the multiverse got it right.

Further, Weinberg's prediction requires that a dynamics that gives rise to a multiverse. String theory provides it.

III. Black hole QM and entanglement

I began with the challenge provided by Planck's units.

If there is a multiverse, then we face another challenge just as great, because the physical laws that we see are partly, maybe largely, random. What to do?

III. Black hole QM and entanglement

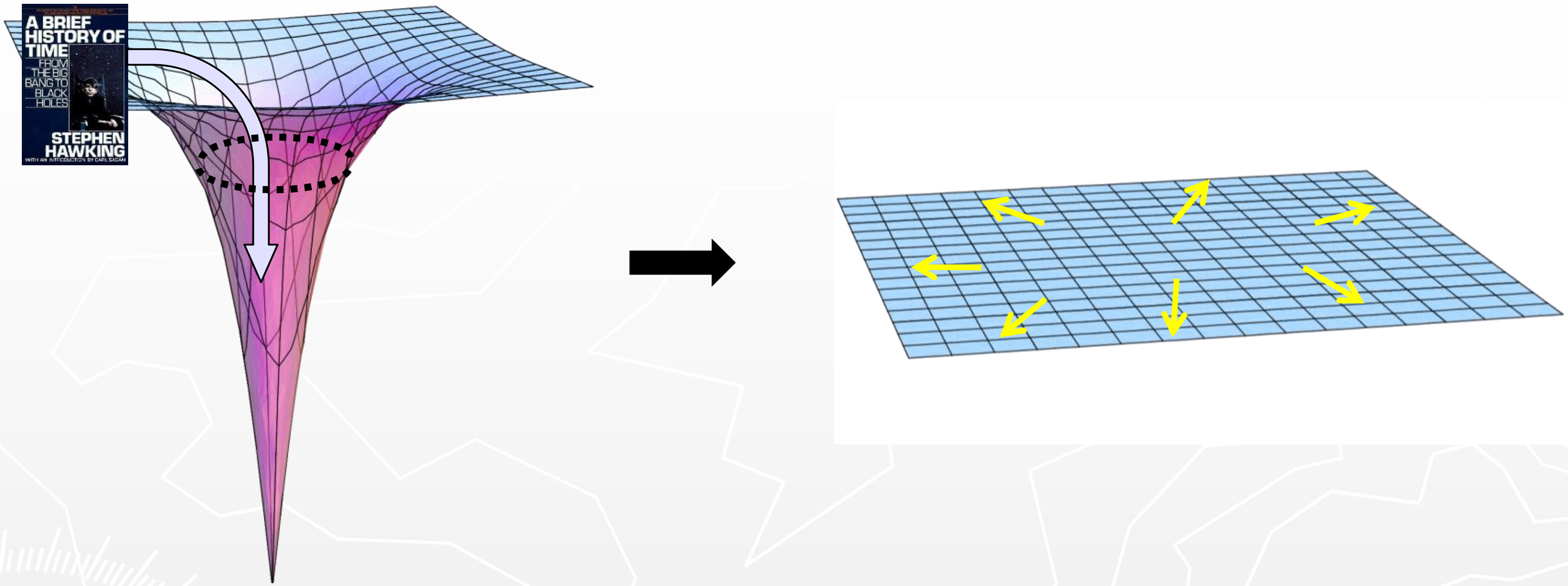
I began with the challenge provided by Planck's units.

If there is a multiverse, then we face another challenge just as great, because the physical laws that we see are partly, maybe largely, random. What to do?

There is still a well-defined problem ahead, to complete the discovery of `string theory,' and new ideas may come from unexpected directions.

The kind of theoretical reasoning that we have followed has brought us quite far. It thrives on conflict and paradox, and black hole quantum mechanics provides it.

Information paradox (naïve version)

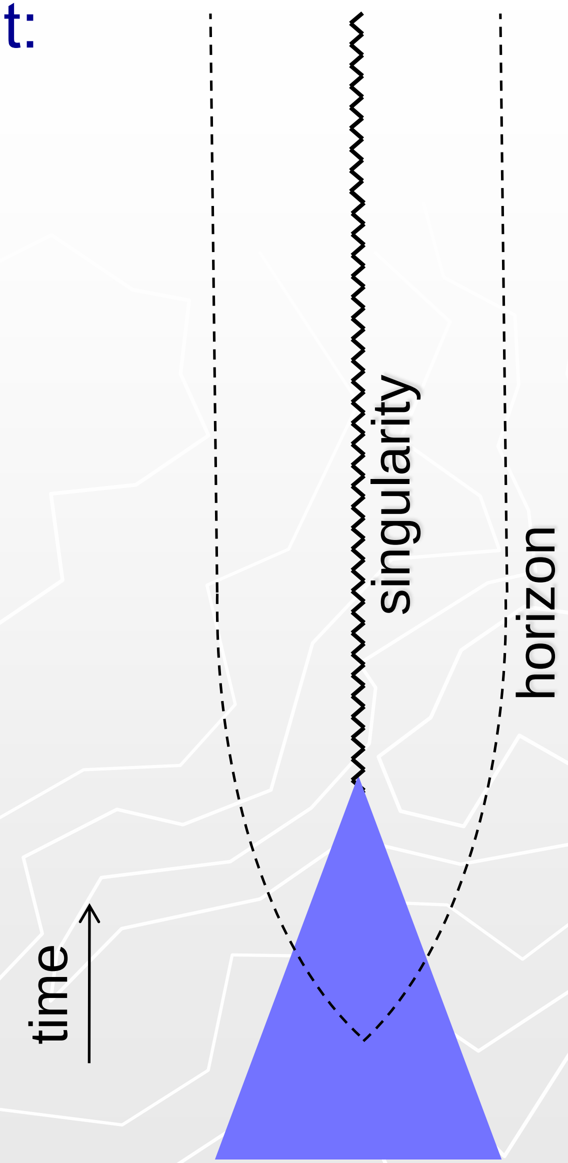


Throw a message into a black hole. Once it is behind the horizon, it can't influence the later Hawking radiation, so the final state does not depend on the message.

Not consistent with $i\hbar \dfrac{d\psi}{dt} = H\psi\,!$

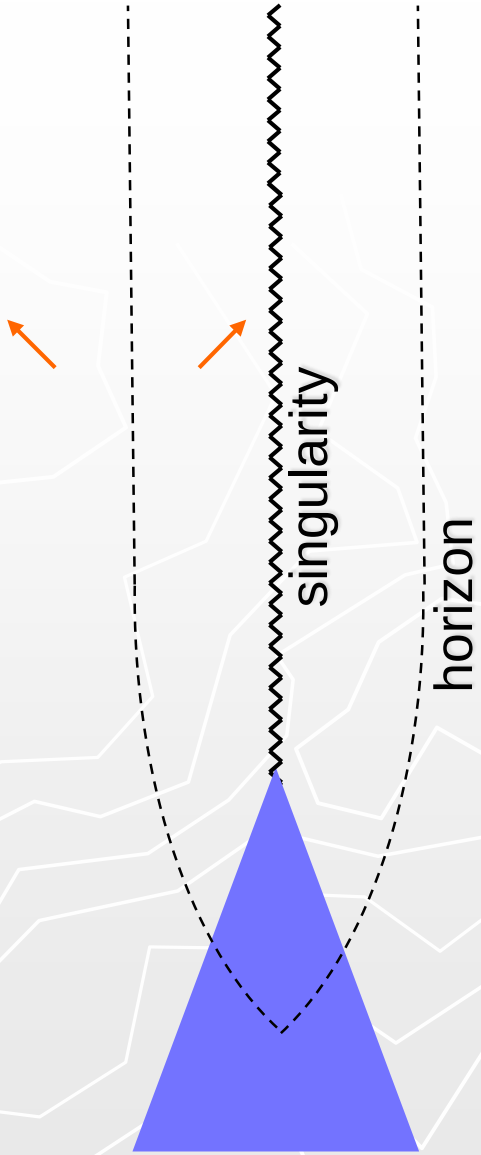
Hawking's thought experiment:

Formation of a classical black hole:

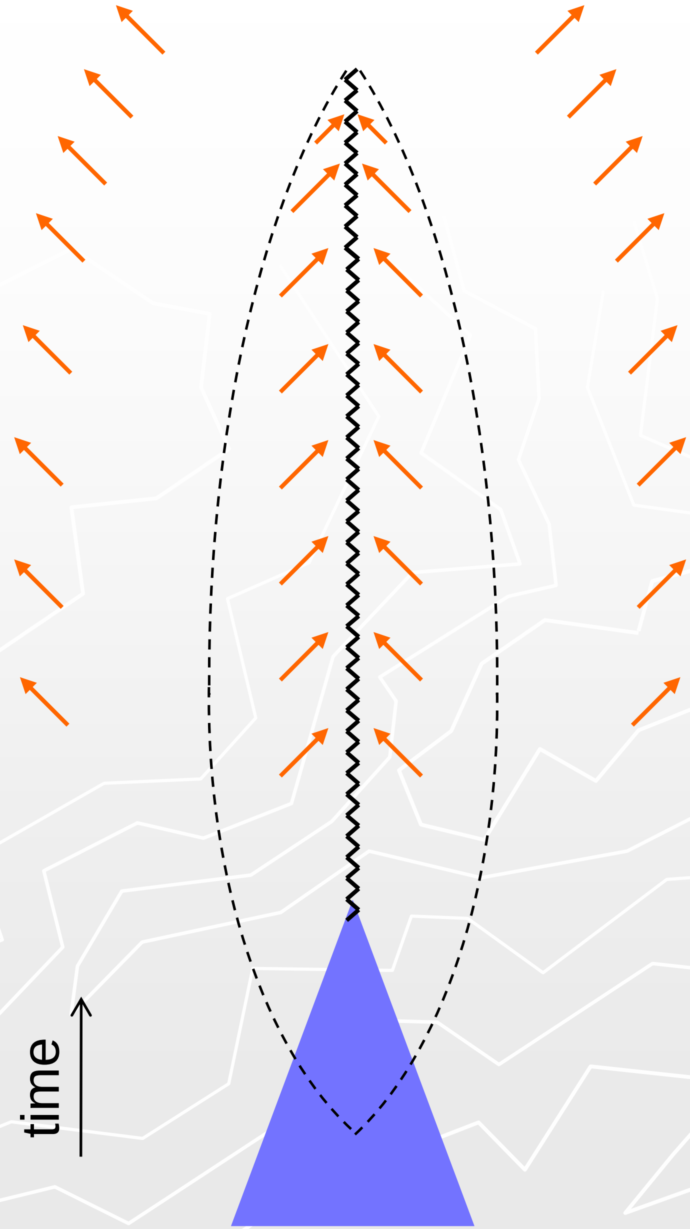


Hawking radiation, producing a black-body spectrum:

The horizon is a region of low curvature, so this calculation should be reliable.



As a result, the black hole eventually emits all of its energy and disappears, leaving only the outgoing Hawking radiation.

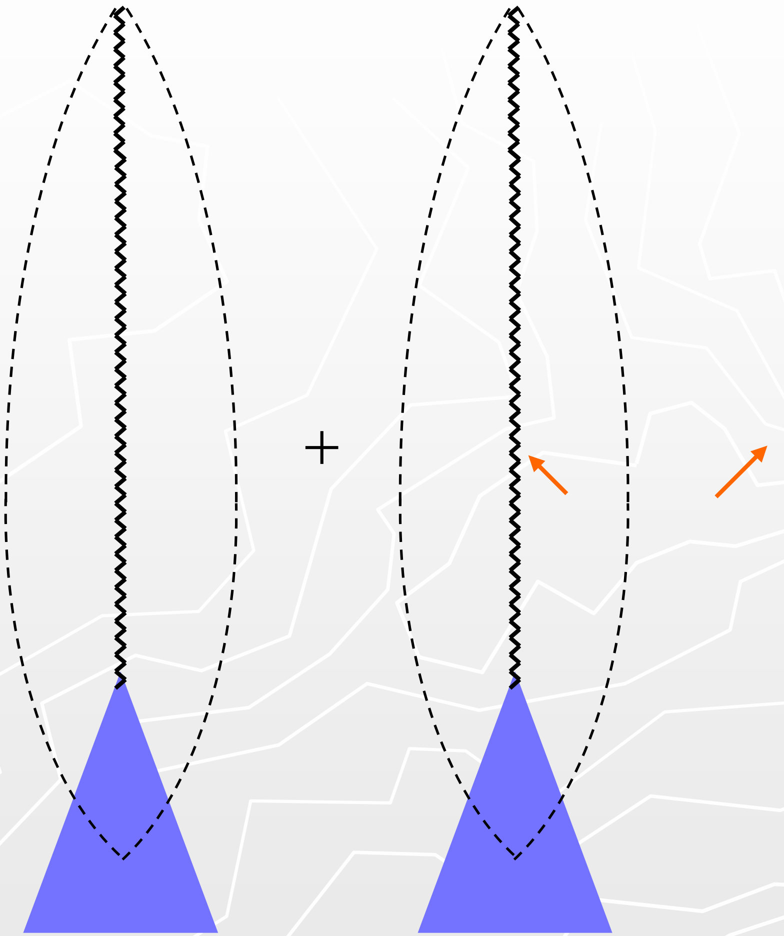


Hawking's paradox:

The Hawking process is a quantum effect, and produces a superposition,

$$\approx |0',0\rangle + |1',1\rangle$$

The two photons are *entangled*; the outside photon by itself is in a mixed state.



Hawking's argument:

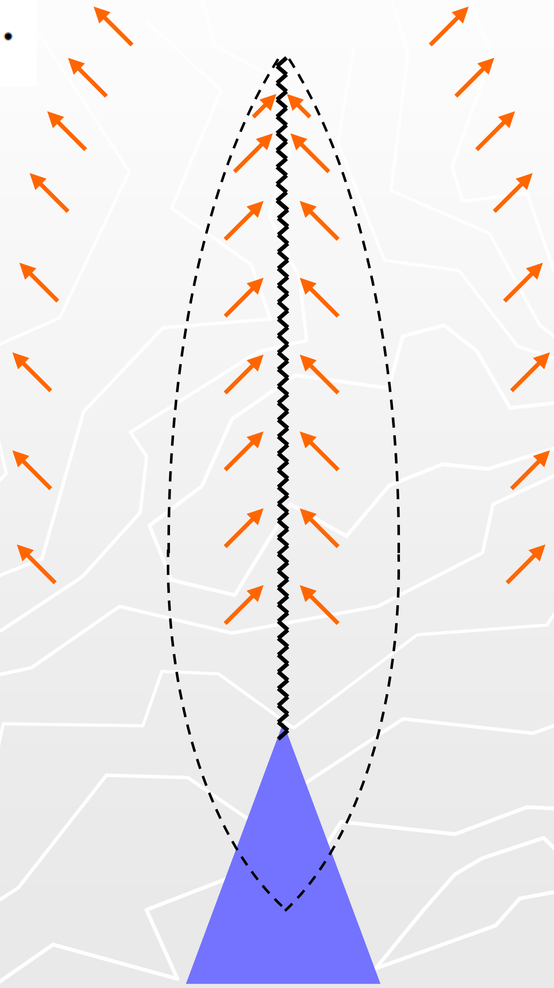
The net result is a highly *entangled* state,

$$\sim\big(|0',0\rangle + |1',1\rangle\big)\big(|0',0\rangle + |1',1\rangle\big)\ldots$$

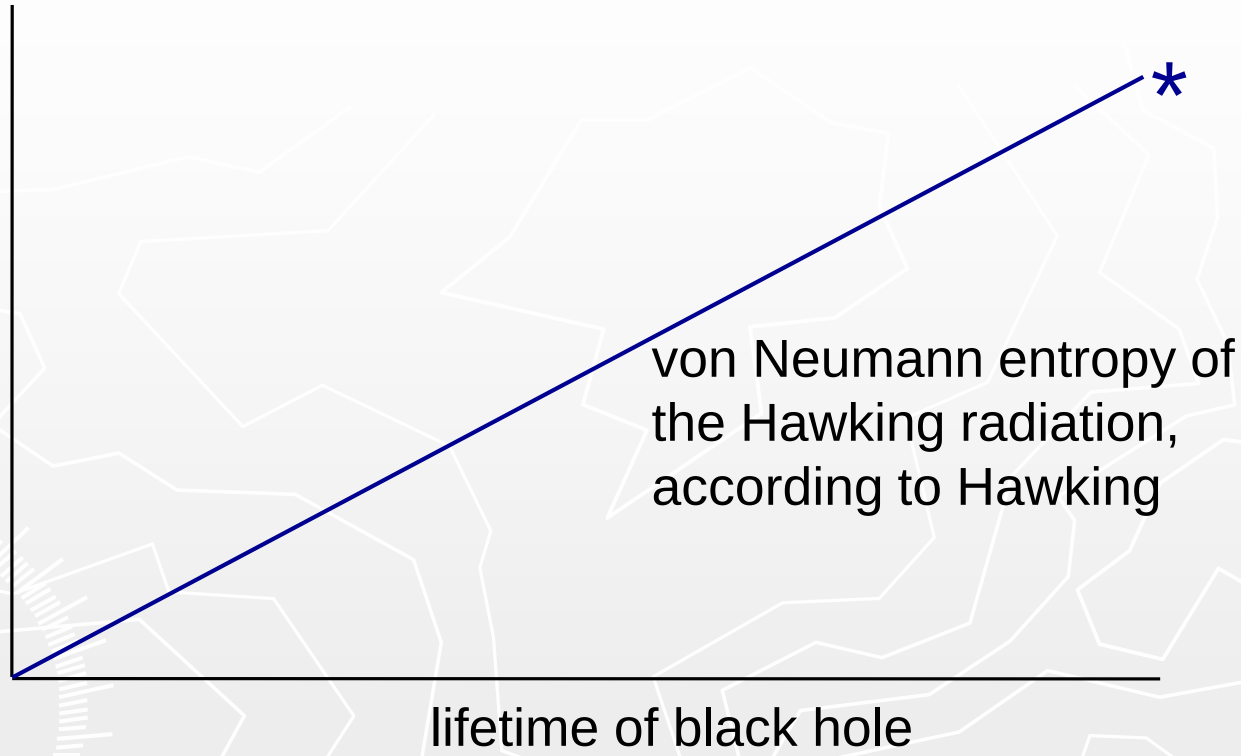
When the evaporation is completed, the inside (primed) degrees of freedom are gone, leaving the Hawking radiation in a highly mixed state.

Pure state evolves to mixed state.

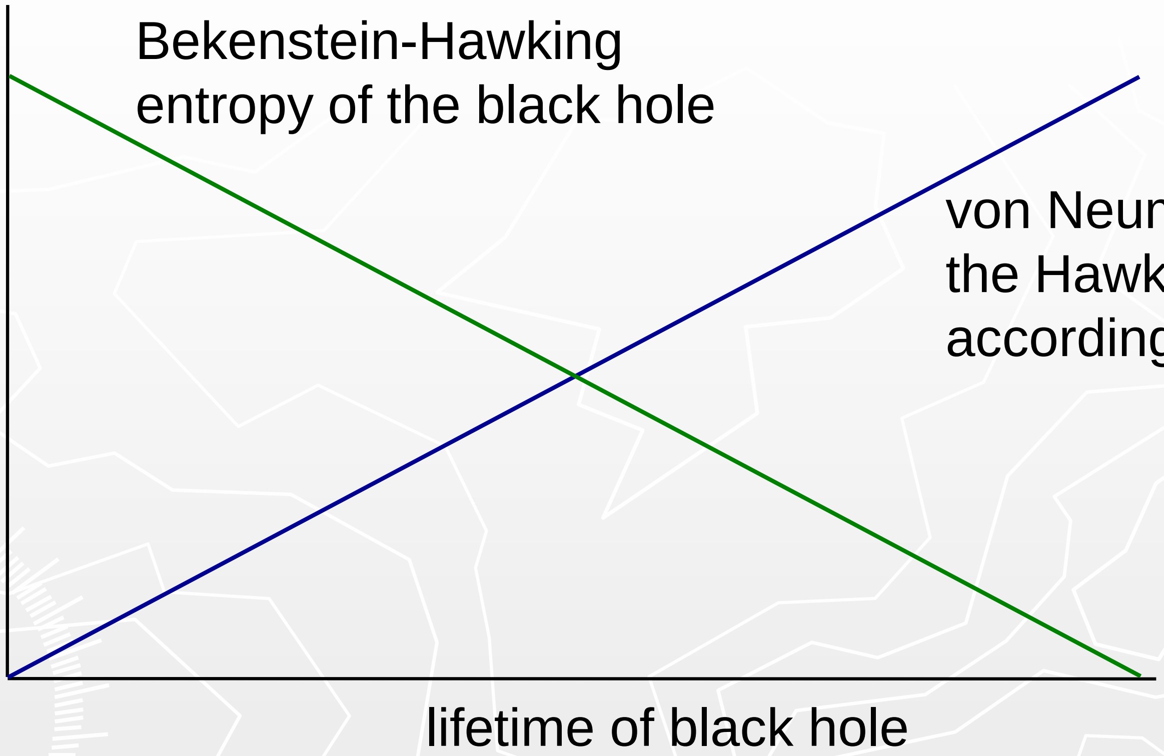
Not sensitive to small corrections.



The Page curve:

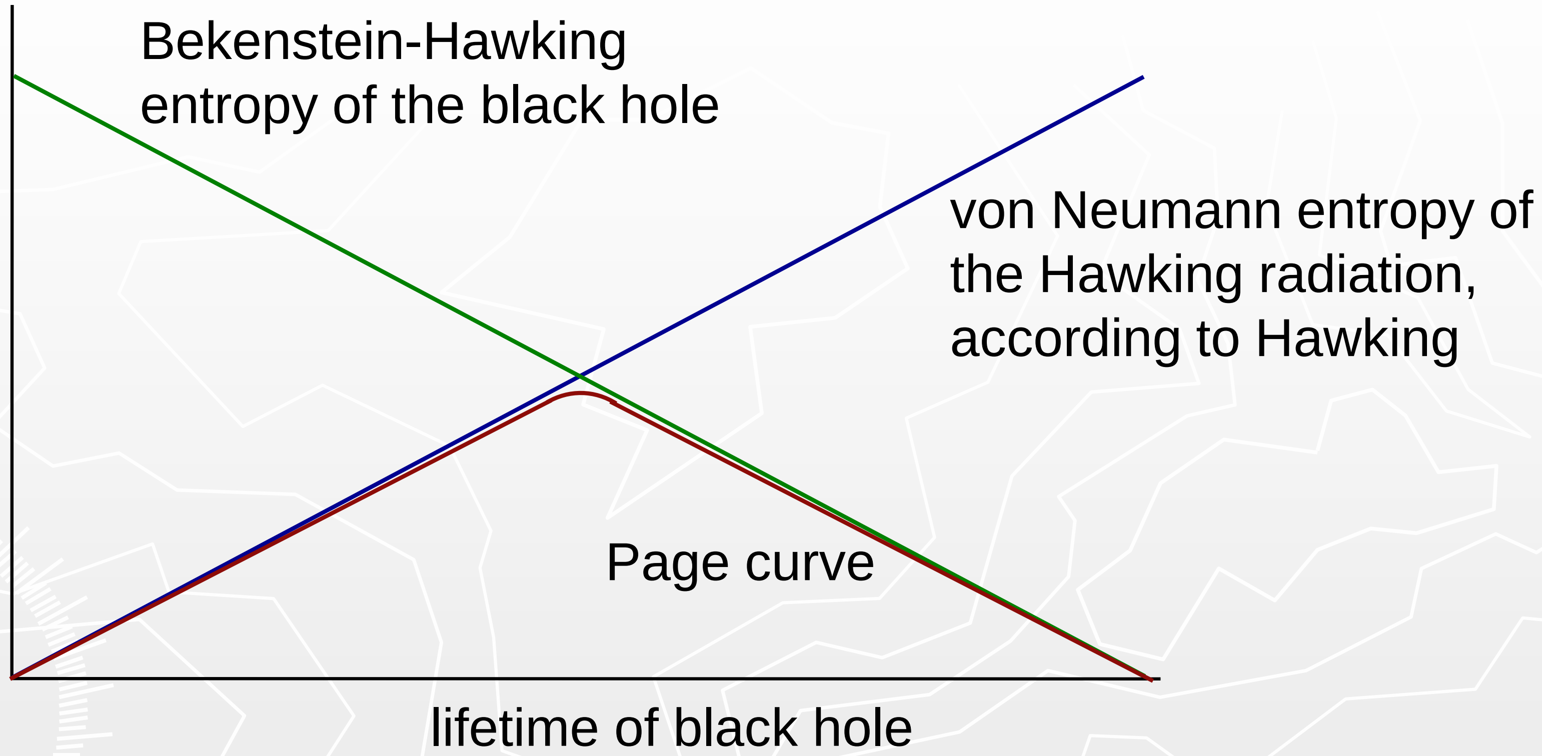


The Page curve:



There is a problem when the curves cross: fine-grained entropy > coarse-grained entropy.

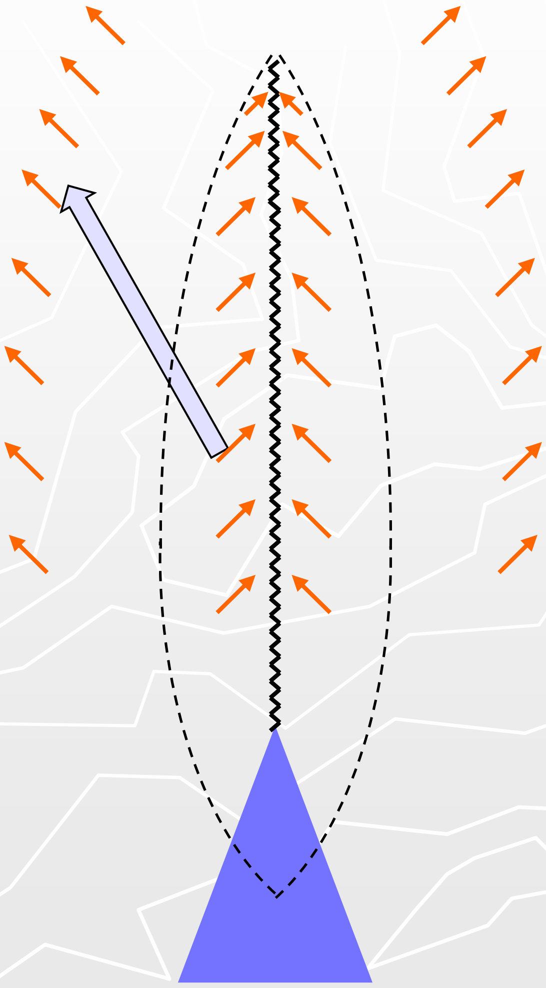
The Page curve:



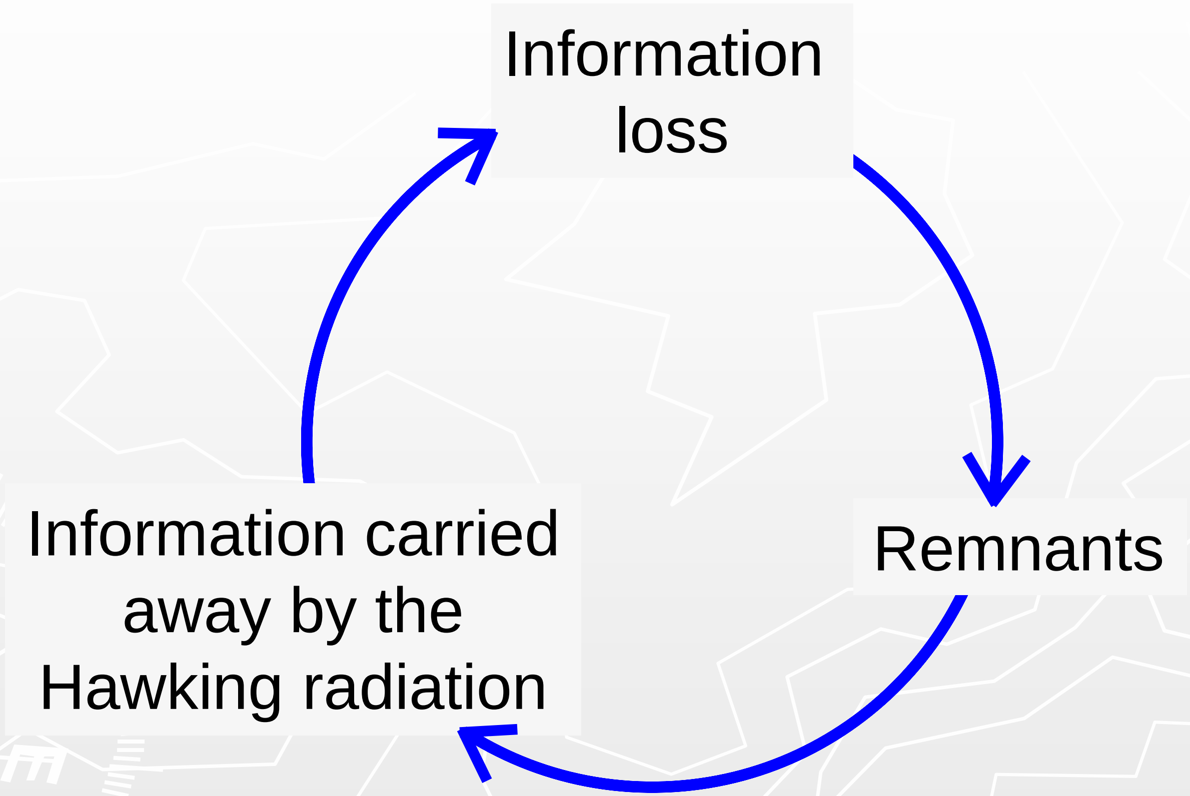
Need $O(1)$ correction, when the black hole is still large, else information loss or remnants.

Ordinary thermal systems, like burning coal, follow the Page curve.

But coal doesn't have a horizon.
If the Hawking radiation is to be in a pure state (information is not lost, but carried away by the radiation) it seems that somehow information must travel *faster than light*…



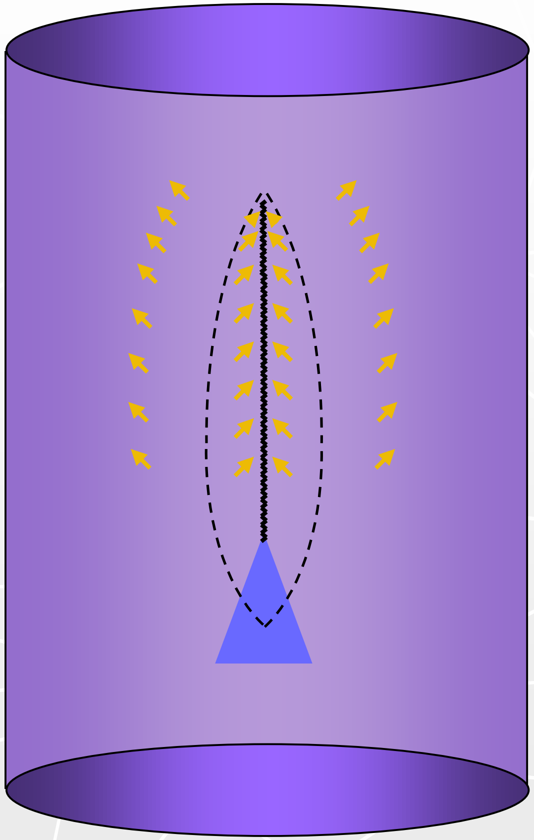
Going around in circles (1976-97):



New insight from AdS/CFT duality

We can consider the Hawking experiment in an AdS box.

Since the dual quantum field theory is described by ordinary QM, pure states must evolve to pure states.



But *how* does the information get out? And how does AdS/CFT really work, with its emergent gravity, emergent space, emergent strings?

A new thought experiment, and a new paradox.

Ahmed Almheiri, Don Marolf, JP, James Sully, arXiv 1207.3123

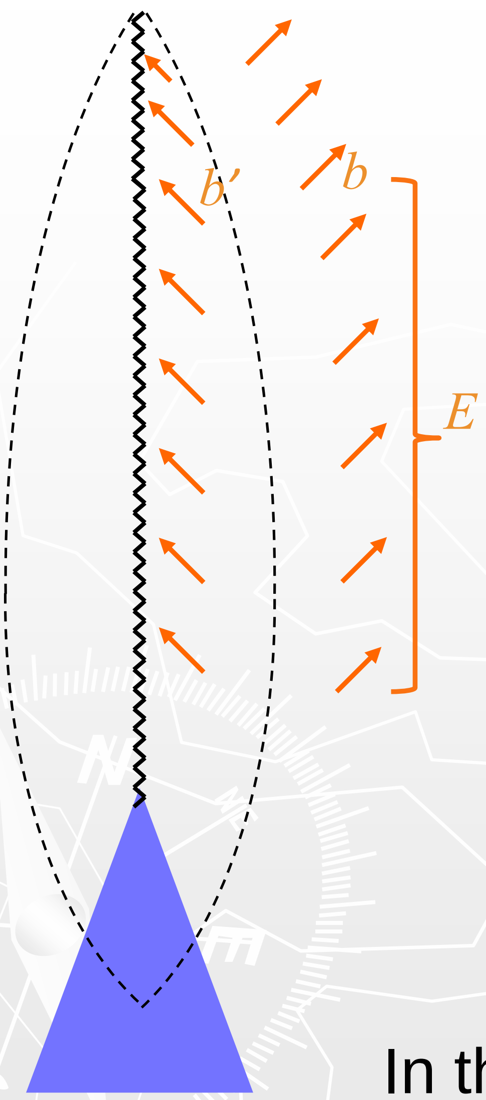
Widely believed after AdS/CFT:

- The Hawking radiation is in a pure state
- An outside observer sees ordinary low energy physics
- An infalling observer sees nothing unusual

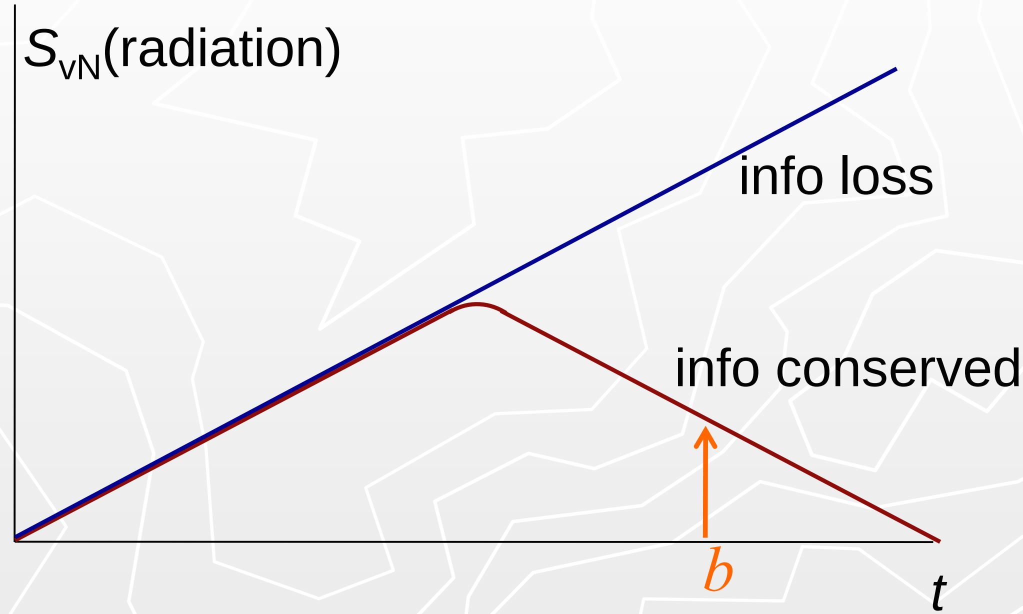
Perhaps a new relativity principle (Black Hole Complementary): the outside and infalling observers see the same information in different places.

Claim: these are incompatible.

Late Hawking photon b, its inside partner b', and the earlier Hawking radiation E.



If information is not lost, b must be entangled with E:



In the Hawking process, b is entangled with b'.

This polygamous entanglement is inconsistent with QM.

Strong subadditivity (Mathur): $S_{b'b} + S_{bE} \geq S_b + S_{b'bE}$

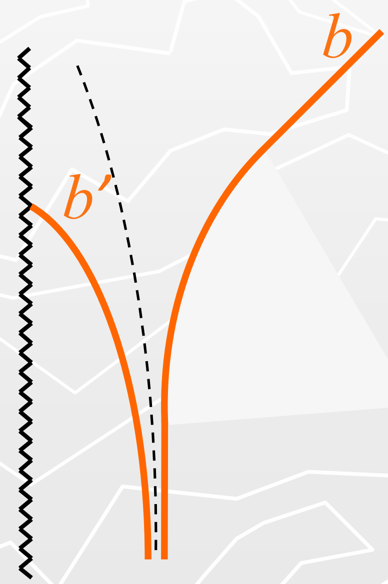
If information is not lost:

$$|b',b,E\rangle \sim |\psi\rangle(|0\rangle|0\rangle + |1\rangle|1\rangle).$$

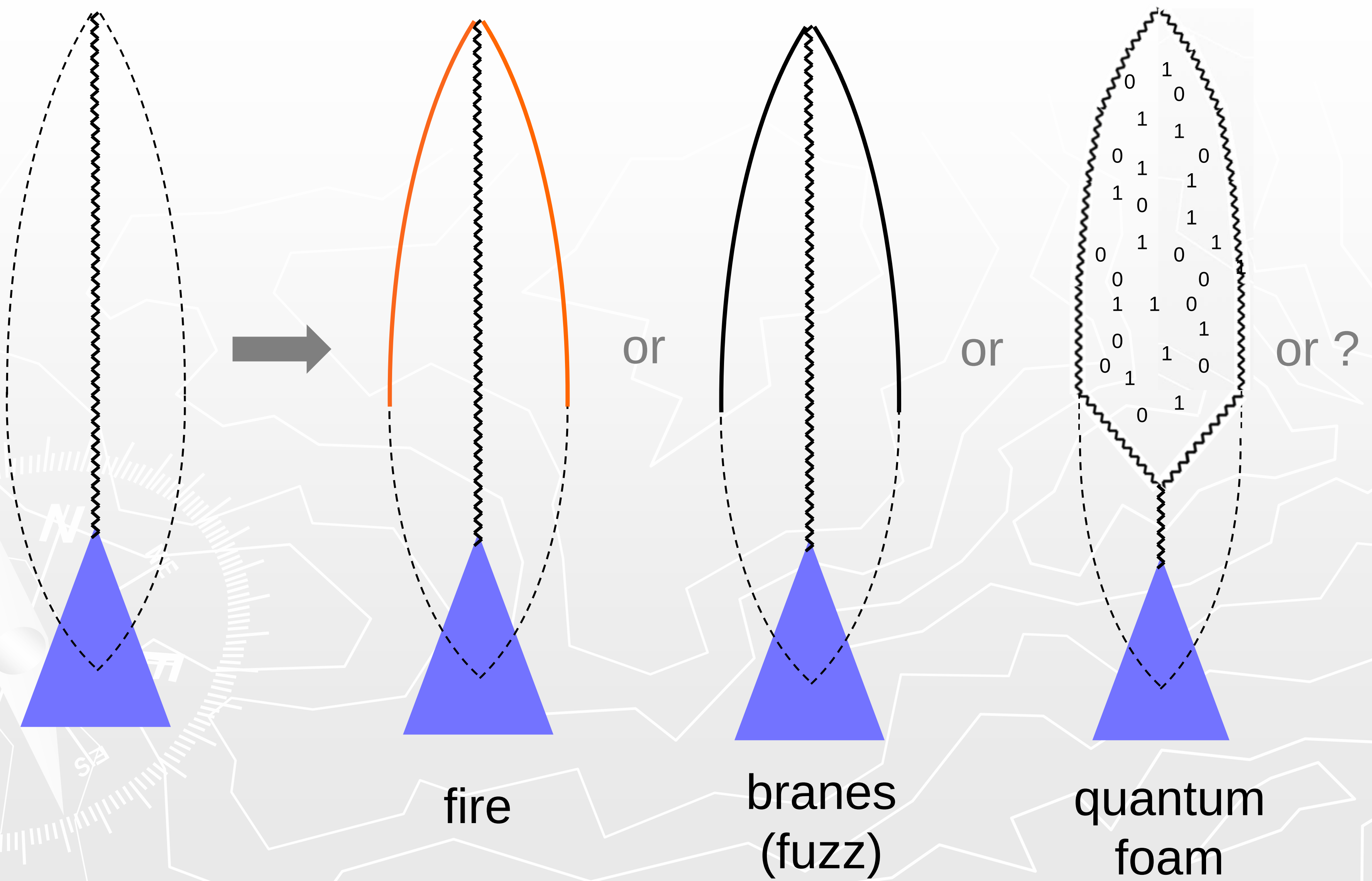
Hawking:

$$|b',b,E\rangle \sim (|0\rangle|0\rangle + |1\rangle|1\rangle)|\psi\rangle.$$

If information is not lost, then the entanglement between b and b' is lost. But nearby modes must be entangled or else there is a large cost in energy: the *firewall*.



An infalling observer, instead of the smooth spacetime predicted by Einstein's theory, hits a shell of high energy particles, or perhaps even the end of space entirely.



or
fire
or
branes
(fuzz)
or ?
quantum
foam

This is a radical conclusion, but what are the alternatives?

- Information is lost after all, or there are remnants
- Nonlocal physics reaches well outside the event horizon
- Quantum mechanics is modified or reinterpreted, perhaps in a subtle way.

Modifications of QM:

Strong complementarity (no global Hilbert space)

Limits on quantum computation (Harlow & Hayden ‘12)

Final state boundary condition at the black hole singularity (Horowitz & Maldacena ’03; Preskill & Lloyd ’13)

EPR = ER (Spacetime from entanglement, Maldacena & Susskind ’13)

Nonlinear observables (Papadodimas & Raju ‘12, Verlinde^2 ’12)

All of these are preliminary frameworks, not theories.
We need a new insight.

Modifications of QM:

Strong complementarity (no global Hilbert space)

Limits on quantum computation (Harlow & Hayden ‘12)

Final state boundary condition at the black hole singu-larity (Horowitz & Maldacena ’03; Preskill & Lloyd ’13)

EPR = ER (Spacetime from entanglement, Maldacena & Susskind ’13)

Nonlinear observables (Papadodimas & Raju ‘12, Verlinde^2 ’12)

All of these are preliminary frameworks, not theories.
We need a new insight.

The idea that the emergence of spacetime is connected with entanglement is arising from multiple directions:

• Einstein’s equation from thermodynamics (see Jacobson’s talk).

• Ryu-Takayanagi: entanglement = area in AdS/CFT

• Almheiri-Dong-Harlow-Pastawski-Preskill-Yoshida: quantum error correction and emergent locality

• Shenker-Stanford-Kitaev-Maldacena: chaos as a signature of the black hole horizon

These ideas seem to be pointing at something, but what?

Conclusions I, Firewalls

- Are there any observational effects for black holes or cosmological horizons?

Too early to say.

- Bigger picture: spacetime from entanglement as the key to unifying QM and GR.

Conclusions II, String theory and quantum gravity

- We have found evidence for a remarkable structure, `string theory,' that unifies quantum mechanics and general relativity.
- This subject has taken surprising twists and turns, and more can be expected.
- In the past, leaping a gap of scales has led to unexpected discoveries:

  Maxwell $\rightarrow$ light

  Dirac $\rightarrow$ antimatter

  Glashow-Weinberg-Salam $\rightarrow$ *W*, *Z*, higgs

Leaping the gap to the Planck scale should be no less fruitful.