Gerard ’t Hooft

How the Quantum Black hole Replies to Your Messages

Centre for Extreme Matter and Emergent Phenomena,
Science Faculty, Utrecht University,
POBox 80.089, 3508 TB, Utrecht

Qui Nonh, Viet Nam, July 24, 2017
Introduction

– Einstein’s theory of gravity, based on General Relativity, and
– Quantum Mechanics, as it was developed early 20th century,
are both known to be valid at high precision.

But combining these into one theory still leads to problems today.

Existing approaches:
– Superstring theory, extended as $M$ theory
– Loop quantum gravity
– Dynamical triangulation of space-time
– Asymptotically safe quantum gravity
are promising but not (yet) understood at the desired level.

In particular when black holes are considered.
We do not claim that these theories are incorrect, but they are not fool-proof. The topology of space and time is not (yet) handled correctly in these theories.

This is why
We do not claim that these theories are incorrect, but they are not fool-proof. The *topology of space and time* is not (yet) handled correctly in these theories.

This is why one encounters problems with:

- information loss
- incorrectly entangled states
- firewalls

We shall show that fundamental new ingredients in all these theories are called for:

- the gravitational back reaction cannot be ignored,
- one must expand the momentum distributions of in- and out-particles in spherical harmonics, and
- one must apply antipodal identification in order to avoid double counting of pure quantum states.

This we will explain.
We do not claim that these theories are incorrect, but they are not fool-proof. The topology of space and time is not (yet) handled correctly in these theories.

This is why one encounters problems with:

- information loss
- incorrectly entangled states
- firewalls

We shall show that fundamental new ingredients in all these theories are called for:

- the gravitational back reaction cannot be ignored,
- one must expand the momentum distributions of in- and out-particles in spherical harmonics, and
- one must apply antipodal identification in order to avoid double counting of pure quantum states.

This we will explain.
\[ ds^2 = \frac{1}{1 - 2GM/r} dr^2 - \left(1 - \frac{2GM}{r}\right) dt^2 + r^2 d\Omega^2; \]
\[
\begin{cases}
\Omega \equiv (\theta, \varphi), \\
\ d\Omega \equiv (d\theta, \sin\theta \, d\varphi)
\end{cases}
\]
\[ ds^2 = \frac{1}{1-2GM/r} dr^2 - (1 - \frac{2GM}{r}) dt^2 + r^2 d\Omega^2; \]

\[ \Omega \equiv (\theta, \varphi), \quad d\Omega \equiv (d\theta, \sin \theta \, d\varphi) \]

\[ t + \log \frac{2GM}{t} \rightarrow t \]
\[ ds^2 = \frac{1}{1-\frac{2GM}{r}} dr^2 - \left(1 - \frac{2GM}{r}\right) dt^2 + r^2 d\Omega^2 ; \]
\[
\begin{cases}
\Omega \equiv (\theta, \varphi), \\
d\Omega \equiv (d\theta, \sin \theta d\varphi)
\end{cases}
\]

\[ t + \log \frac{2GM}{t} \quad \rightarrow t \]
Penrose diagrams
Discovery by S, Hawking:

Quantized field operators are composed of annihilation operators \( a \) and creation operators \( a^\dagger \). If these fields transform locally under general coordinate transformations, then, in the new coordinates, annihilation and creation operators mix, getting partially interchanged. Consequently, the vacuum state \( |\Omega\rangle \) as seen by one observer, defined by the equation \( a|\Omega\rangle = 0 \), will obey \( (a + \varepsilon a^\dagger)|\Omega\rangle = 0 \), which means that this is not the vacuum state as seen by the other observer.

\[ \Rightarrow \text{ Black holes emit particles.} \]
Discovery by S, Hawking:

Quantized field operators are composed of annihilation operators \((a)\) and creation operators \((a^\dagger)\). If these fields transform locally under general coordinate transformations, then, in the new coordinates, annihilation and creation operators mix, getting partially interchanged. Consequently, the vacuum state \(|\Omega\rangle\) as seen by one observer, defined by the equation \(a|\Omega\rangle = 0\), will obey \((a + \varepsilon a^\dagger)|\Omega\rangle = 0\), which means that this is not the vacuum state as seen by the other observer.

\[ \Rightarrow \text{Black holes emit particles.} \]

But the *Hartle Hawking state* stretches over regions I and region II. Data in region II end up in III, so these data are lost ⇒ information gets lost! How to cure this situation?

Hawking suggests that one has to average over the unseen states in region II / III ⇒ one thus obtains a *thermal* state, with temperature \(T_{\text{Hawking}} = 1/8\pi GM\)

This would violate quantum mechanics!
A theory is needed that blends black holes with other, ordinary forms of matter.

To do this correctly, we need a description of black holes in terms of pure quantum states (as opposed to thermodynamical objects).

The set of pure quantum states must be time reversal symmetric.

Here we’ll show how all this is done.

This is not a model, and (in the beginning) not even a theory – we simply apply known laws of physics as far as we can . . . – no strings attached –

One learns that “information” does not need to “get lost”; however, one then discovers that small, but important, amendments must be made on the known formulation of Nature’s laws:
A theory is needed that blends black holes with other, ordinary forms of matter.

To do this correctly, we need a description of black holes in terms of pure quantum states (as opposed to thermodynamical objects).

The set of pure quantum states must be time reversal symmetric.

Here we’ll show how all this is done.

This is not a model, and (in the beginning) not even a theory – we simply apply known laws of physics as far as we can . . . – no strings attached –

One learns that “information” does not need to “get lost”; however, one then discovers that small, but important, amendments must be made on the known formulation of Nature’s laws:

What is matter, and what is the space-time topology for a black hole.
Our prototype is the Schwarzschild black hole. No serious complications when generalised to tortoise coordinates near on the region close to the horizon: \( r \approx 2GM \). There:

\[
\begin{align*}
  u^+ u^- & \to \frac{r}{2GM} - 1, \quad \tau = \frac{t}{2GM} \\
  u^- (\tau) &= u^- (0) e^\tau \\
  u^+ (\tau) &= u^+ (0) e^{-\tau}
\end{align*}
\]

As time goes forwards, \( u^+ \) approaches the horizon asymptotically; as time goes backwards, \( u^- \) approaches the past horizon asymptotically (tortoises).
Hartle-Hawking vacuum:

\[ |HH\rangle = C \sum_{E,n} e^{-\frac{1}{2} \beta E} |E, n\rangle_I |E, n\rangle_{II} \]

Time boost for distant observer = Lorentz boost for local observer.

Usual interpretation:

\[ I = \text{outside} \]
\[ III = \text{inside} \]

\[ \rightarrow \text{quantum entanglement becomes entropy:} \]
\[ \rightarrow \text{a thermal state . . .} \]

In- and out-going particles: energies \( E \) for distant observer stay small.

But for the local observer, energies of in-particles in distant past, as well as the out-particles in distant future, rapidly tend to infinity.
How to do this right:

Hawking particles are seen as being formed by a *local vacuum state*. There, they have no effect on the metric at all. *By time reversal invariance*, we must now also first regard the states of the incoming particles as having no effect on the metric as well.

Thus, consider first the black hole metric without the effects of matter – the *eternal black hole*. Penrose diagram:
Hard and soft particles

Now consider particles. Their energies, in this coordinate frame, may go way beyond $M_{\text{Planck}}$. If so, we call them hard particles: their effects on space-time curvature cannot be ignored.

This curvature is strong and chaotic: no observer trying to cross such a curtain of particles can survive: firewalls

Almheiri, Marolf, Polchinski, Sully (2013)

Particles whose energies, in a given Lorentz frame, are small compared to $M_{\text{Planck}}$ will be called soft particles.

Their effects on curvature are small compared to $L_{\text{Planck}}$, and will be ignored (or taken care of in perturbative Qu.Gravity).

During its entire history, a black hole has in-going matter (grav. implosion) and out-going matter (Hawking).
If we want to express these in terms of pure quantum states, we must expect firewalls both on the future and past event horizon.

(The pure quantum theory must be CPT symmetric)
Such firewalls would form a natural boundary surrounding region I

That can't be right

Derivation of Hawking radiation asks for analytic extension to region III
Time reversal symmetry then asks for analytic extension to region IV
In combination, you then also get region II
Such firewalls would form a natural boundary surrounding region I

That can’t be right

Derivation of Hawking radiation asks for analytic extension to region III

Time reversal symmetry then asks for analytic extension to region IV

In combination, you then also get region II
Such firewalls would form a natural boundary surrounding region I

That can't be right

Derivation of Hawking radiation asks for analytic extension to region III

Time reversal symmetry then asks for analytic extension to region IV

In combination, you then also get region II
Such firewalls would form a natural boundary surrounding region I.

That can't be right.

Derivation of Hawking radiation asks for analytic extension to region III.

Time reversal symmetry then asks for analytic extension to region IV.

In combination, you then also get region II.
All this suggests that firewalls can be switched on and off: 

*the firewall transformation.*

(1.) Note: Hawking’s wave function seems to form a single quantum state (*if we assume both regions I and II of the Penrose diagram to be physical! – see later*). A firewall would form infinitely many quantum states. What kind of mapping do we have? Aren’t we dealing with an information problem here??

(2.) Region II would have its own asymptotic regions: $\infty^{'}, \infty^{'+},$ and $\infty^{'-}$. What is their physical significance?

Wait and see ...
The gravitational effect of a fast, massless particle is easy to understand: Schwarzschild metric of a particle with tiny rest mass $m \ll M_{\text{Planck}}$:
The gravitational effect of a fast, massless particle is easy to understand:

Schwarzschild metric of a particle with tiny rest mass $m \ll M_{\text{Planck}}$:

And now apply a strong Lorentz boost, so that $E/c^2 \gg M_{\text{Planck}}$:

![Diagram showing flat space with curvature](image-url)
This gives us the gravitational backreaction:

Lorentz boosting the light (or massless) particle gives the *Shapiro time delay* caused by its grav. field:

$$\delta x = \delta \hat{x}$$

$$u - p - \delta u - \frac{1}{2}$$

$$\delta u^- (\tilde{x}) = -4G \rho^- (\tilde{x}') \log |\tilde{x} - \tilde{x}'| .$$

We start with **only soft particles** populating space-time in the Penrose diagram.

We first have to understand the evolution operator for **short time intervals** only. Firewalls have no time to develop.
The evolution law for the soft particles is fully dictated by QFT on curved space-time.

At $|\tau| = \mathcal{O}(1)$, particles going in, and Hawking particles going out, are soft. However, during our short time interval, some soft particles might pass the borderline between soft and hard: they now interact with the out-particles. The interaction through QFT forces stay weak, but the gravitational forces make that (early) in-particles interact strongly with (late) out-particles.

Effect of gravitational force between them easy to calculate ...
Calculate Shapiro shift,

Every in-particle with momentum $p^-$ at solid angle $\Omega = (\theta, \varphi)$ causes a shift $\delta u^-$ of all out-particles at solid angles $\Omega' = (\theta', \varphi')$:

$$\delta u^-(\Omega') = 8\pi G f(\Omega', \Omega)p^- ; \quad (1 - \Delta_W)f(\Omega', \Omega) = \delta^2(\Omega', \Omega) .$$

Many particles: $p^-(\Omega) = \sum_i p_i^- \delta^2(\Omega, \Omega_i) \rightarrow$

$$\delta u^-(\Omega') = 8\pi G \int d^2\Omega f(\Omega', \Omega)p^-(\Omega) .$$

Small modification: replace $\delta u^-_{\text{out}}$ by $u^-_{\text{out}}$, then:

$$u^-_{\text{out}}(\Omega) = 8\pi G \int d^2\Omega' f(\Omega, \Omega')p^-_{\text{in}}(\Omega')$$

adding an in-going particle with momentum $p^-_{\text{in}}$, corresponds to displacing all out-going particles by $u^-_{\text{out}}$ as given by our equation.

All $u^-_{\text{out}}$ are generated by all $p^-_{\text{in}}$. 
Expand in Spherical harmonics:

\[ u^\pm (\Omega) = \sum_{\ell, m} u_{\ell m} Y_{\ell m}(\Omega) , \]

\[ p^\pm (\Omega) = \sum_{\ell, m} p_{\ell m} Y_{\ell m}(\Omega) ; \]

\[ [u^\pm (\Omega), p^\mp (\Omega')] = i \delta^2 (\Omega, \Omega') , \]

\[ [u^\pm_{\ell m}, p^\mp_{\ell' m'}] = i \delta_{\ell \ell'} \delta_{mm'} ; \]

\[ u^-_{\text{out}} = \frac{8 \pi G}{\ell^2 + \ell + 1} p^-_{\text{in}} , \]

\[ p^+_\ell m = \text{total momentum in of } \text{out}-\text{particles in } (\ell, m)\text{-wave} , \]

\[ u^+_{\ell m} = (\ell, m)\text{-component of c.m. position of } \text{in out}-\text{particles} . \]

Because we have linear equations, all different \( \ell, m \) waves decouple, and for one \( (\ell, m) \)-mode we have just the variables \( u^\pm \) and \( p^\pm \). They represent only one independent coordinate \( u^+ \), with \( p^- = -i \partial/\partial u^+ \).
The basic, explicit, calculation

Our algebra generates the scattering matrix, by giving us the boundary condition that replaces $|\text{in}\rangle$-states by $|\text{out}\rangle$-states.

NOT a model, theory, or assumption . . .

Apart from the most basic assumption of unitary evolution, this is nothing more than applying GR and quantum mechanics!

Spherical harmonics diagonalises the total $S$-matrix into 1 dimensional partial diff. equations!
Commutator equation for $u$ and $p$:

$$[u, p] = i,$$

so that

$$\langle u | p \rangle = \frac{1}{\sqrt{2\pi}} e^{ipu}.$$

In tortoise coordinates, split $u$ and $p$ in a positive part and a negative part:

$$u \equiv \sigma_u e^{\varrho_u}, \quad p = \sigma_p e^{\varrho_p}; \quad \sigma_u = \pm 1, \quad \sigma_p = \pm 1,$$

Then

$$\tilde{\psi}_{\sigma_p}(\varrho_p) = \sum_{\sigma_u = \pm 1} \int_{-\infty}^{\infty} d\varrho \ K_{\sigma_u \sigma_p}(\varrho) \tilde{\psi}_{\sigma_u}(\varrho - \varrho_p),$$

with

$$K_\sigma(\varrho) \equiv \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2} \varrho} e^{-i\sigma \varrho}.$$

If $\varrho_p \to \varrho_p + \lambda$, then simply $u \to u e^{-\lambda}$, $p \to p e^{\lambda}$, a symmetry of the Fourier transform.
Use this symmetry to write plane waves:

\[ \tilde{\psi}_{\sigma u}(\varrho_u) \equiv \psi_{\sigma u}(\kappa) e^{-i\kappa\varrho_u} \quad \text{etc.} \]

Fourier transform becomes:

\[ \tilde{\psi}_\sigma(\kappa) = \sum_{\sigma_p = \pm 1} F_{\sigma\sigma_p}(\kappa) \tilde{\psi}_{\sigma u}(\kappa) ; \quad F_\sigma(\kappa) \equiv \int_{-\infty}^{\infty} K_\sigma(\varrho) e^{-i\kappa\varrho} d\varrho . \]

Thus, we see in-going waves produce out-going waves. One finds (just do the integral):

\[ F_\sigma(\kappa) = \int_0^{\infty} \frac{dy}{y} y^{\frac{1}{2} - i\kappa} e^{-i\sigma y} = \Gamma(\frac{1}{2} - i\kappa) e^{-\frac{i\sigma \pi}{4}} - \frac{\pi}{2} \kappa \sigma . \]

Matrix \( \begin{pmatrix} F_+ & F_- \\ F_- & F_+ \end{pmatrix} \) is unitary: \( F_+ F_-^* = -F_- F_+^* \) and \( |F_+|^2 + |F_-|^2 = 1 \).
The in-particles never get the opportunity to become truly hard particles. Like a “soft wall”-boundary condition near the origin of the Penrose diagram. Wave functions going in reflect as wave functions going out. Soft in-particles emerge as soft out-particles.

No firewall, ever.

The total of the in-particles in regions $I$ and $II$ are transformed (basically just a Fourier transform) into out-particles in the same two regions.

Regions $III$ and $IV$ are best to be seen as lying somewhere on the time-line where time $t$ is somewhere beyond infinity.
The in-particles never get the opportunity to become truly hard particles. Like a “soft wall”-boundary condition near the origin of the Penrose diagram. Wave functions going in reflect as wave functions going out. Soft in-particles emerge as soft out-particles.

No firewall, ever.

The total of the in-particles in regions $I$ and $II$ are transformed (basically just a Fourier transform) into out-particles in the same two regions.

Regions $III$ and $IV$ are best to be seen as lying somewhere on the time-line *where time* $t$ *is somewhere beyond infinity*

This is how the black hole reads and answers to your messages: the in-going particle (your message) leaves a gravitational footprint in the out particles. The footprints, all taken together, generate the complete spectrum of out-particles. But only if we expand in spherical harmonics, we see that this is unitary.

But there is a problem: Where do the states in region $II$ go?
The antipodal identification
Sanchez(1986), Whiting(1987)

Regions $I$ and $II$ of the Penrose diagram are exact copies of one another. Does region $II$ represent the “inside” of the black hole? NO! There are asymptotic regions. Region $I$ is carbon copy of region $II$.

We must assume that region $II$ describes the same black hole as region $I$. It represents some other part of the same black hole. Which other part? The local geometry stays the same, while the square of this $O(3)$ operator must be the identity.

Search for $A \in O(3)$ such that: $A^2 = I$, and $Ax = x$ has no real solutions for $x$.

$\Rightarrow$ All eigenvalues of $A$ must be $-1$. Therefore: $A = -I$:

the antipodal mapping.
We stumbled upon a new restriction for all general coordinate transformations:

**Amendment # 2 for Nature’s Constitution”**

*For a curved space-time background to be used to describe a region in the universe, one must demand that every point on our space-time region represent exactly one point in the universe (not two, as in analytically extended Schwarzschild metric)*

The emergence of a non-trivial topology needs not be completely absurd, as long as no signals can be sent around. We think that this is the case at hand here. It is the absence of singularities in the physical domain of space-time that we must demand.

Note that, now, $\ell$ has to be odd!
Black emptiness: blue regions are the accessible part of space-time; dotted lines indicate identification.

The white sphere within is not part of space-time. Call it a ‘vacuole’.

At given time $t$, the black hole is a 3-dimensional vacuole. The entire life cycle of a black hole is a vacuole in 4-d Minkowski space-time: an instanton

N.Gaddam, O.Papadoulaki, P.Betzios (Utrecht PhD students)

Space coordinates change sign at the identified points
– and also time changes sign
(Note: time stands still at the horizon itself).
Virtual black holes must be everywhere in space and time. Due to vacuum fluctuations, amounts of matter that can contract to become black holes, must occur frequently. They also evaporate frequently, since they are very small. This produces small vacuoles in the space-time fabric. How to describe multiple vacuoles is not evident. The emerging picture could be that of “space-time foam”:
Virtual black holes and space-time foam  (Summary)

Virtual black holes must be everywhere in space and time. Due to vacuum fluctuations, amounts of matter that can contract to become black holes, must occur frequently. They also evaporate frequently, since they are very small. This produces small vacuoles in the space-time fabric. How to describe multiple vacuoles is not evident. The emerging picture could be that of “space-time foam”:

![Space-time foam diagram](image-url)
Virtual black holes must be everywhere in space and time. Due to vacuum fluctuations, amounts of matter that can contract to become black holes, must occur frequently. They also evaporate frequently, since they are very small. This produces small vacuoles in the space-time fabric. How to describe multiple vacuoles is not evident. The emerging picture could be that of “space-time foam”:
A timelike Möbius strip

Draw a spacelike closed curve as follows:

Begin at a point \( r_0 = 2GM \), \( t_0 = 0 \), \( (\theta_0, \varphi_0) = \Omega_0 \) on the horizon. Move to larger \( r \) values, then travel to the antipode: \( r_0 = 2GM \), \( t_0 = \tilde{\Omega}_0 = (\pi - \theta_0, \varphi_0 + \pi) \). You arrived at the same point, so the (space-like) curve is closed.

Now look at the environment \( \{dx\} \) of this curve. Continuously transport \( dx \) around the curve. The identification at the horizon demands

\[
 dx \leftrightarrow -dx ,
\]

both for the space coordinates and for time. If we keep \( dt = 0 \) we have a three dimensional curve, but the identification at the horizon then has negative parity. If we would keep \( dx \) timelike, then we see that time changes sign at the horizon, and we cannot undo this using small deformations, as in the inhomogeneous part of the Lorentz group. So this is a Möbius strip, in particular in the time direction.
There are no direct contradictions, but take in mind that the Hamiltonian switches sign as well. Demanding that the external observer chooses the point where the Hamilton density switches sign as being on the horizon, gives us a good practical definition for the entire Hamiltonian.

Note that the soft particles near the horizon adopted the dilaton operator as their Hamiltonian. That operator leaves regions I and II invariant. Also, the boundary condition, our “scattering matrix”, leaves this Hamiltonian invariant. Therefore, indeed, breaking the Hamiltonian open exactly at the horizon still leaves the total Hamiltonian conserved. So indeed, there are no direct contradictions.

However, this is a peculiarity that we have to take into consideration.
Entanglement of Hawking particles

The hartle-Hawking state,

\[ |HH\rangle = C \sum_{E,n} e^{-\frac{1}{2} \beta E} |E, n\rangle_I |E, n\rangle_{II} \]

\[ I = \text{antipode of } II \]

is now a pure quantum state, where regions I and II are entangled.

It is not a thermal state.

Only if we do not look at states II, the states in I seem to form a perfect thermal mixture.
Perfect entanglement:

String theory did not warn us about these features . . . !

We only used perturbative QFT and General Relativity to reach this result.
More to be done. Searching for like-minded colleagues.

A mapping of the momenta $p^-_{in}$ of the in-particles onto the positions $u^-_{out}$ of the out-particles. Agrees with time evolution:

$$p^-_{in} \rightarrow p^-_{in}(0)e^{\tau}, \quad u^+_{in} \rightarrow u^+_{in}(0)e^{-\tau};$$

$$p^+_{out} \rightarrow p^+_{out}(0)e^{-\tau}, \quad u^-_{out} \rightarrow u^-_{out}(0)e^\tau.$$ 

What we calculated is the **footprint** of in-particles onto the out-particles, caused by gravity.

And then: $\delta u^- \rightarrow u^-$ implies that now the in-particles are to be replaced by the out-particles. **The particles are their footprints!**

This makes sense: $u^-$ is the position of the out-particles – if we Fourier transform that, we get the momentum distribution $p^+$ of the out-particles!

Footprints promoted to the status of particles themselves. Avoids double counting: only describe the in-particle or its footprint (the out-particle), not both, as in the older equations.

Note: **hard** in-particles generate **soft** out-particles and **vice versa**.

This way, replace all hard particles by soft ones. This removes the firewalls: the **firewall transformation**.
Authors of older papers, when encountering “firewalls”, did not take into account that neither in-going nor out-going particles should be followed for time intervals $\delta \tau$ with $|\delta \tau| \gg O(1)$.

This evolution law involves soft particles only. Is it unitary?
Authors of older papers, when encountering “firewalls”, did not take into account that neither in-going nor out-going particles should be followed for time intervals $\delta \tau$ with $|\delta \tau| \gg O(1)$.

This evolution law involves soft particles only. Is it unitary?

Only in the variables $p^{\pm}(\Omega)$ and $u^{\pm}(\Omega)$ are involved.

No quantum numbers like baryon, lepton . . .

$p^{\pm}(\Omega)$ are like vertex insertions in string theories.
Postulating that this respects unitarity makes sense . . .

First amendment on Nature's Constitution:

A particle may be replaced by its gravitational footprint: At a horizon, out-particles are the Fourier transforms of in-particles.

Second problem:

What is the relation between regions I and II? 
Both have asymptotic domains: two universes!
a) Wave functions $\psi(u^+)$ of the in-particles live in region I, therefore $u^+ > 0$.  
b) Out-particles in region I have $\psi(u^-)$ with $u^- > 0$.

c) In region II, the in-particles have $u^+ < 0$ and the out-particles $u^- < 0$.  

The physical picture