

Gerard 't Hooft

Quantum Black Holes, Firewalls and the Topology of Space-Time

A pedagogical treatment of a new approach

version #3

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

February 2, 2017



This is the ‘Lecture Notes’ version of my latest paper on black hole microstates. These notes highlight the essentials, such that they should be easier to read and understand. My plan is to keep this updated. If new insights arise, they will be added.

In this version, v3, a chapter, with figures, on “virtual black holes and space-time foam”, was added at the end.

These notes must be seen as a personal view of the author. Most importantly, we try to avoid any wild assumptions, as we usually see in the literature. Our claim is that everything here is essentially derived from basic laws that are well-established. Not even String Theory is used (this might change in the future).

These notes are not intended for use in talks (too much information on single “slides”) with apologies for not always following the same notation as used elsewhere.

There is a discussion site on <https://www.researchgate.net/>



Black holes are an essential feature of the gravitational force. Without a thorough understanding of their properties one cannot understand quantum gravity. See: [G. 't Hooft, arxiv:1612.08640 \[gr-qc\] + references there.](#)

String theory could have been a good tool as well, and I don't go as far as saying that string theory would be wrong, but this lecturer does not understand string theory well enough.

The assumption that “particles are pieces of string” is just an assumption. I think we do not need to make such assumptions. If string theory is right, we will be able to derive it from other first principles.

Anyway, I here describe a path towards understanding without dubious assumptions at all. Just derivations from known principles, such as Quantum Mechanics and General Relativity.

Note, that these theories are also assumptions, but they are much more compelling. Dropping or loosening these assumptions is much more difficult than keeping them as long as we can. No theory is sacred, and some sense of taste for what to keep and what not, is a useful attire for a theoretical physicist.

We claim that topics such as microstates can be understood very well in our theory.

The tortoise coordinates

Our prototype is the Schwarzschild black hole. No serious complications are expected when generalized to Kerr-Newman or such. The *extreme* black hole would not serve our purposes, because it is a limiting case, and its horizon is fundamentally different from the more general black holes.

Consider the Schwarzschild metric:

$$ds^2 = \frac{1}{1 - \frac{2GM}{r}} dr^2 - \left(1 - \frac{2GM}{r}\right) dt^2 + r^2 d\Omega^2 ; \quad \begin{cases} \Omega & \equiv (\theta, \varphi) , \\ d\Omega & \equiv (d\theta, \sin\theta d\varphi) . \end{cases}$$

Go to Kruskal-Szekeres coordinates x, y , defined by

$$\begin{aligned} xy &= \left(\frac{r}{2GM} - 1\right) e^{r/2GM} ; \\ y/x &= e^{t/2GM} . \end{aligned}$$

$$ds^2 = \frac{32(GM)^3}{r} e^{-r/2GM} dx dy + r^2 d\Omega^2 .$$

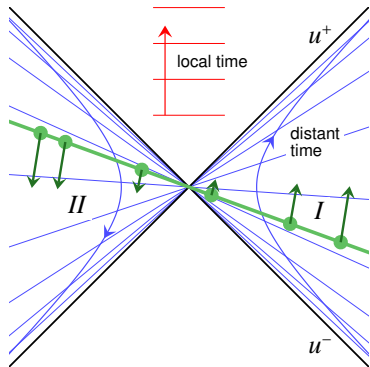
At $r = 2GM$, we have $x = 0$: past event horizon, and
 $y = 0$: future event horizon.

We concentrate on the region close to the horizon: $r \approx 2GM$. There:

$$x = \frac{\sqrt{e/2}}{2GM} u^+ ; \quad y = \frac{\sqrt{e/2}}{2GM} u^- ; \quad 2GM \equiv R$$

$$ds^2 \rightarrow 2du^+ du^- + R^2 d\Omega^2 .$$

$$\text{time } t/4GM = \tau$$



$$u^-(\tau) = u^-(0)e^\tau$$

$$u^+(\tau) = u^+(0)e^{-\tau}$$

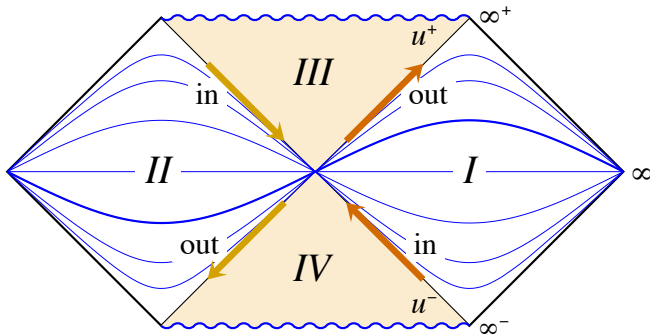
As time goes forwards, u^+ approaches the horizon asymptotically;

as time goes backwards, u^- approaches the past horizon asymptotically (tortoises).

If we redefine $u^+ \rightarrow f(u^+)$, $u^- \rightarrow g(u^-)$, then the metric keeps the form $ds^2 = 2A(u)du^+du^- + r^2(u)d\Omega^2$. \rightarrow *Map u^\pm on compact domains:*

The Penrose diagram.

For pure Schwarzschild (without matter responsible for the formation of a black hole, or representing its final decay):



Now, we begin to deviate from standard practice:

This is the “eternal black hole” (it never formed, it will never decay).
We keep it. Why?

We put the “microstates” on a Cauchy surface. As yet, we are only interested in how they evolve for a short amount of time. For the quantum properties of the BH, the distant past, and the distant future, should be irrelevant. They are always irrelevant in quantum physics. Late in the future, or long ago in the past, the states we are looking at now, will be highly entangled states. Not suitable for generating a well-defined, curved space-time. Leave all that nonsense out.

First, consider the particles near the horizon(s).

$$\text{Mass shell: } 2p^+p^- + \tilde{p}^2 + \mu^2 = 0.$$

Here, \tilde{p} is the transverse part of the momentum, $\mu = \text{mass}$;

$$|\tilde{p}| \approx L/R \text{ and } \mu \text{ are basically constant.}$$

$$\text{But } p^-(\tau) = p^-(0)e^\tau, \quad p^+(\tau) = p^+(0)e^{-\tau}.$$

Define *soft* particles: $|\tilde{p}|, \mu \ll M_{\text{Planck}}$ Negligible effect on space-time curvature.

As $\tau \rightarrow \infty$, $p^+ \rightarrow 0$, $p^- \rightarrow \infty$: that's a *hard* in-particle;

As $\tau \rightarrow -\infty$, $p^- \rightarrow 0$, $p^+ \rightarrow \infty$: that's a *hard* out-particle.

$|\tilde{p}|$ and μ stay small.

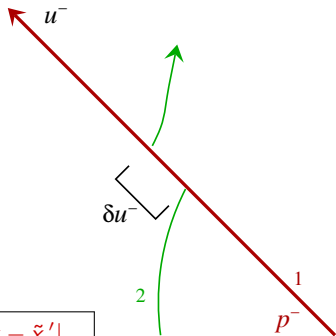
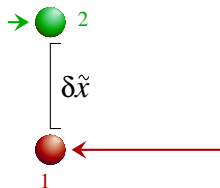
To understand what happens with the evolution at longer time intervals, we have to understand what the hard in- and out- particles do.

As $\tau \gg 1$, the in-particles become hard. Their interactions with other in-particles are negligible (they basically move in parallel orbits), but they do 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.

The gravitational force between them can easily be calculated

The gravitational backreaction:

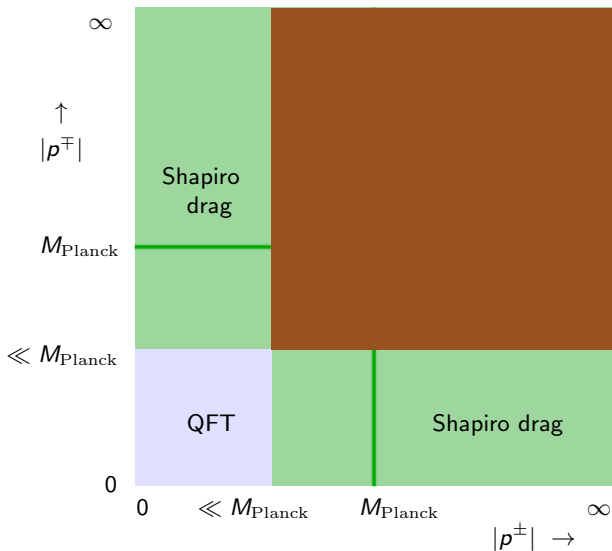
Calculate the *Shapiro time delay* caused by the grav. field of a fast moving particle:
simply Lorentz boost the field of a particle at rest:



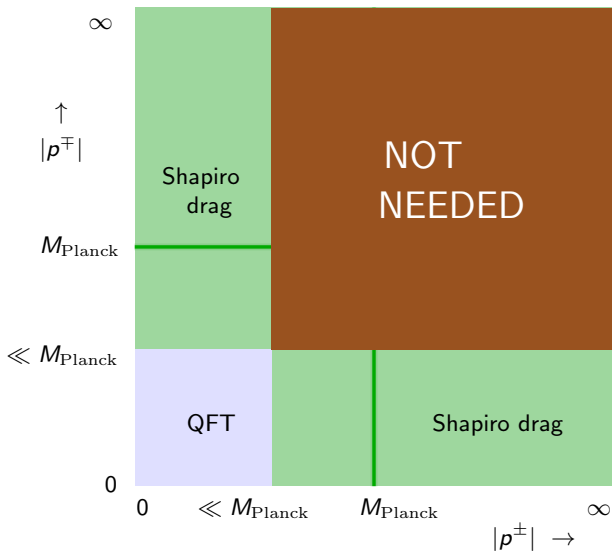
$$\delta u^-(\tilde{x}) = -4G p^-(\tilde{x}') \log |\tilde{x} - \tilde{x}'| .$$

P.C. Aichelburg and R.U. Sexl, *J. Gen. Rel. Grav.* **2** (1971) 303,
W.B. Bonnor, *Commun. Math. Phys.* **13** (1969) 163,
T. Dray and G. 't Hooft, *Nucl. Phys.* **B253** (1985) 173.

What we need to know is the interactions between in- and out-particles.



What we need to know is the interactions between in- and out-particles.



The fact that the mutual interactions between particles at the Planck mass or beyond, will not be needed, is a very important aspect of this work. As you will see.

We start with only soft particles on a Cauchy surface of the Penrose diagram. These will define all quantum microstates of the black hole.

Now, the question is how do these evolve with time.

The soft particles won't stay soft; their longitudinal momenta will quickly explode. To see what happens, calculate the Shapiro shift.

in-particle with momentum p^- at solid angle $\Omega' = (\theta', \varphi')$
causes a shift δu^- at solid angle $\Omega = (\theta, \varphi)$:

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

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

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

Now start with one unique “reference” black hole, yielding out-particles with given positions at the origin of some coordinate system. Suppose $p^-(\Omega)$ represents all in-particles needed to describe any *other* black hole.

Then the out-particles will be at positions $u^-(\Omega)$ given by

$$u_{\text{out}}^-(\Omega) = 8\pi G \int d^2\Omega' f(\Omega, \Omega') p_{\text{in}}^-(\Omega') .$$

Spherical wave expansion:

$$\begin{aligned}
 u^\pm(\Omega) &= \sum_{\ell, m} u_{\ell m} Y_{\ell m}(\Omega) , & p^\pm(\Omega) &= \sum_{\ell, m} p_{\ell m}^\pm Y_{\ell m}(\Omega) ; \\
 [u^\pm(\Omega), p^\mp(\Omega')] &= i\delta^2(\Omega, \Omega') , & [u_{\ell m}^\pm, p_{\ell' m'}^\mp] &= i\delta_{\ell\ell'}\delta_{mm'} ; \\
 u_{\text{out}}^- &= \frac{8\pi G}{\ell^2 + \ell + 1} p_{\text{in}}^- , & u_{\text{in}}^+ &= -\frac{8\pi G}{\ell^2 + \ell + 1} p_{\text{out}}^+ ,
 \end{aligned} \tag{1}$$

$p_{\ell m}^\pm$ = total momentum in of in^{out} -particles in (ℓ, m) -wave ,

$u_{\ell m}^\pm$ = (ℓ, m) -component of c.m. position of in^{out} -particles .

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

3 more steps to be taken:

1. Momentum density $p^\pm(\Omega)$ for every quantum state in a QFT is well-defined. However, our evolution law will only be unitary if we can find the QFT quantum state back from the momentum density. That's difficult.

Use vertex insertions as in string theory?

2. What is the physical interpretation of region II ?

We demand *locality*. This means that commutators of space-like separated operators should vanish. Then, since in-going signals in region I produce signals in the out-going objects in region II , we must guarantee that I and II are not space-like separated.

Postulate that region II refers to the same black hole as region I , but *not at the same solid angle* $\Omega = (\theta, \varphi)$. Only one possibility:

The antipodal identification: $\Omega \rightarrow \tilde{\Omega} = (\pi - \theta, \varphi + \pi)$

3. The “firewalls”. Soft particles become hard particles. Must be ‘removed’.

Obvious suggestion: all information carried by the in- particles is now present in the out-particles. The in-particles are redundant (“quantum clones”). Leave the hard ones out. Hilbert space is completely specified by the coordinates $u^-(\Omega)$ of the out-particles. Note: as soon as $|u^-| > L_{\text{Planck}}$, these out-particles are soft.

This is the “firewall transformation”; it removes firewalls.

The basic, explicit, calculation

The algebra (1) generates the scattering matrix, by giving us the *boundary condition* that replaces $|\text{in}\rangle$ -states by $|\text{out}\rangle$ -states. This boundary condition *replaces the old brick wall model* and it is embarassingly easy to derive

(The catch will be exposed later) .

All of this is NOT a model, or a theory, or an assumption ...

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

Let there be two operators, u and p , obeying the commutator equation

$$[u, p] = i, \quad \text{so that} \quad \langle u|p\rangle = \frac{1}{\sqrt{2\pi}} e^{ipu}.$$

and a wave function $|\psi\rangle$, defined by $\psi(u) \equiv \langle u|\psi\rangle$. Then

$$\hat{\psi}(p) \equiv \langle p|\psi\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} du e^{-ipu} \psi(u).$$

Now introduce **tortoise coordinates**, and split both u and p in a **positive part** and a **negative part**:

$$u \equiv \sigma_u e^{\varrho_u}, \quad p \equiv \sigma_p e^{\varrho_p}; \quad \sigma_u = \pm 1, \quad \sigma_p = \pm 1, \quad \text{and} \\ \tilde{\psi}_{\sigma_u}(\varrho_u) \equiv e^{\frac{1}{2}\varrho_u} \psi(\sigma_u e^{\varrho_u}), \quad \tilde{\psi}_{\sigma_p}(\varrho_p) \equiv e^{\frac{1}{2}\varrho_p} \hat{\psi}(\sigma_p e^{\varrho_p});$$

$$\text{normalized: } |\psi|^2 = \sum_{\sigma_u=\pm} \int_{-\infty}^{\infty} d\varrho_u |\tilde{\psi}_{\sigma_u}(\varrho_u)|^2 = \sum_{\sigma_p=\pm} \int_{-\infty}^{\infty} d\varrho_p |\tilde{\psi}_{\sigma_p}(\varrho_p)|^2.$$

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

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

Notice the symmetry under $\varrho_u \rightarrow \varrho_u + \lambda$, $\varrho_p \rightarrow \varrho_p - \lambda$, which is simply the symmetry $u \rightarrow u e^{\lambda}$, $p \rightarrow p e^{-\lambda}$, a property of the Fourier transform, and an invariance of our algebra.

We now use this symmetry to write plane waves:

$$\tilde{\psi}_{\sigma_u}(\varrho_u) \equiv \check{\psi}_{\sigma_u}(\kappa) e^{-i\kappa\varrho_u} \quad \text{and} \quad \tilde{\psi}_{\sigma_p}(\varrho_p) \equiv \check{\psi}_{\sigma_p}(\kappa) e^{i\kappa\varrho_p} \quad \text{with}$$

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

Thus, we see left-going waves produce right-going waves. On finds (just do the integral):

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

Matrix $\begin{pmatrix} F_+ & F_- \\ F_- & F_+ \end{pmatrix}$ is unitary: $F_+ F_-^* = -F_- F_+^*$ and $|F_+|^2 + |F_-|^2 = 1$.

Look at how our soft particle wave functions evolve with time τ , slide # 8.
 Their Hamiltonian is the dilaton operator:

$$H = -\frac{1}{2}(u^+ p^- + p^- u^+) = \frac{1}{2}(u^- p^+ + p^+ u^-) =$$

$$i \frac{\partial}{\partial \varrho_{u^+}} = -i \frac{\partial}{\partial \varrho_{u^-}} = -i \frac{\partial}{\partial \varrho_{p^-}} = i \frac{\partial}{\partial \varrho_{p^+}} = \kappa ,$$

(if we apply the previous slide to the coordinates u^+ and p^-).

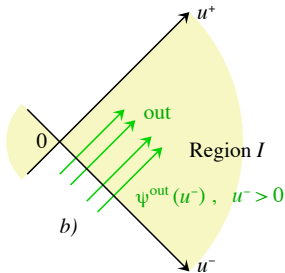
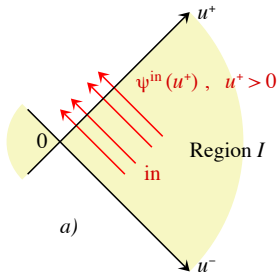
Add the scale factor $\frac{8\pi G}{\ell^2 + \ell + 1}$, to get, if $u^\pm = \sigma_\pm e^{\varrho^\pm}$,

$$\psi_{\sigma_+}^{\text{in}} e^{-i\kappa \varrho^+} \rightarrow \psi_{\sigma_-}^{\text{out}} e^{i\kappa \varrho^-} , \quad (2)$$

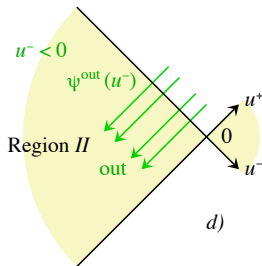
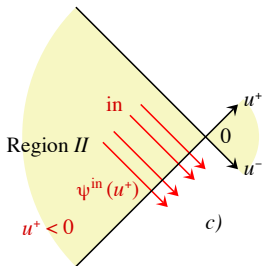
$$\psi_{\sigma_-}^{\text{out}} = \sum_{\sigma_+} F_{\sigma_+ \sigma_-}(\kappa) e^{-i\kappa \log(8\pi G / (\ell^2 + \ell + 1))} \psi_{\sigma_+}^{\text{in}}$$

These equations generate the contributions to the scattering matrix from all (ℓ, m) sectors of the system, where $|m| \leq \ell$. At every (ℓ, m) , we have a contribution to the position operators $u^\pm(\theta, \varphi)$ and momentum operators $p^\pm(\theta, \varphi)$ proportional to the partial wave function $Y_{\ell m}(\theta, \varphi)$. The signs of $u^\pm(\theta, \varphi)$ tell us whether we are in region *I* or region *II*. The signs of $p^\pm(\theta, \varphi)$ tell us whether we added or subtracted a particle from region *I* or region *II*.

- 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 $u^- > 0$.



The physical picture



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

Note that the in-particles will never get the opportunity to become truly hard particles.

Eq. (2) is to be seen as a “soft wall”-boundary condition near the origin of the Penrose diagram. Wave functions going in are reflected as wave functions going out. These again emerge as soft particles.

Thus, there is no firewall, ever.

In the previous slide, 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.

Note that the regions *III* and *IV* in the Penrose diagram (see slide 6) never play much of a role, even if an observer falling in region *III* would want to assure us that (s)he is still alive.

These regions are best to be seen as lying somewhere on the time-line *where time t is somewhere beyond infinity* (thus a mere repetition of the degrees of freedom we have seen before)

The catch?

Our "boundary condition at the origin" is in terms of momentum distributions $p^\pm(\theta, \varphi)$ and center-of-mass positions $u^\pm(\theta, \varphi)$ only.

But we would need the quantum wave functions as elements of Fock space (which would be specified by positions or momenta, but also by other quantum numbers!)

Such mapping should be unitary in Hilbert space. We are not certain that this can be done, but it is natural to look at how it is done in string theory. Our momentum distributions are like vertex insertions, although they are on the horizon instead of the string world sheet.

The antipodal identification

Regions I and II of the Penrose diagram are exact copies of one another. Often, it was thought that region II describes something like the 'inside' of a black hole. That cannot be right, since region II , like region I , has asymptotic regions. Hawking suggested that region II might be some other black hole, in an other universe, or far away in our universe. However, our 2×2 scattering matrix implies that the two regions are in contact with each other quantum mechanically. In ordinary branches of physics, such long-distance communication never takes place, and I don't think theories with such features make any sense.

It is far more natural to assume that region II describes the same black hole as region I . It must then represent some other part of the same black hole. Which other part? The local geometry stays the same, while the square of this $SO(3)$ operator must be the identity.

There is exactly one possibility: This is the $SO(3)$ operator $-\mathbb{I}$, which is:
the antipodal mapping.

Antipodal *identification* only holds for the central point (origin) of the Penrose diagram. Regions *I* and *II* are different regions of the universe. But by relating region *II* to region *I* by demanding that the angular coordinates are antipodes, means that now the mapping from Schwarzschild coordinates to Kruskal Szekeres coordinates is one-to-one. This now turns out to be an essential property of our coordinate transformations. Thus, we arrive at a new restriction for all general coordinate transformations:

In applying general coordinate transformations for quantized fields on a curved space-time background, to use them as a valid model for a physical quantum system, one must demand that the following constraint hold: the mapping must be one-to-one and differentiable.

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.

Black emptiness: **blue** regions are the accessible part of space-time; dotted lines indicate identification.

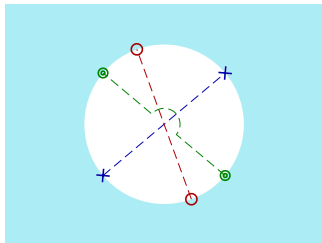
The white sphere within is *not* part of space-time. Call it a '**vacuole**'.

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.

Space coordinates change sign at the identified points

– *and also time changes sign*

(Note: time stands still at the horizon itself).



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 inhomogenous 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.**

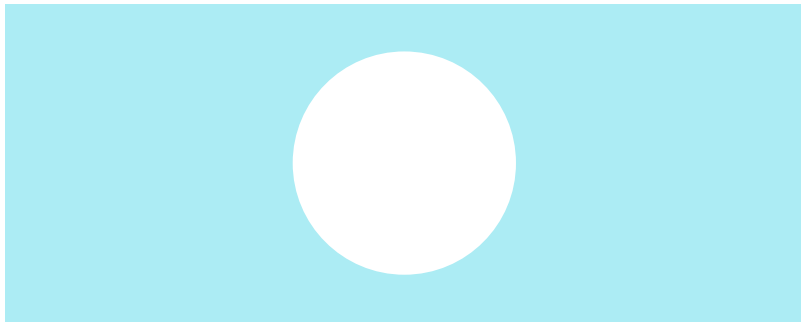
Demanding that the external observer chooses the point where the Hamilton density switches sign as being on the horizon, gives us a good practical definition for the entire Hamiltonian.

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.

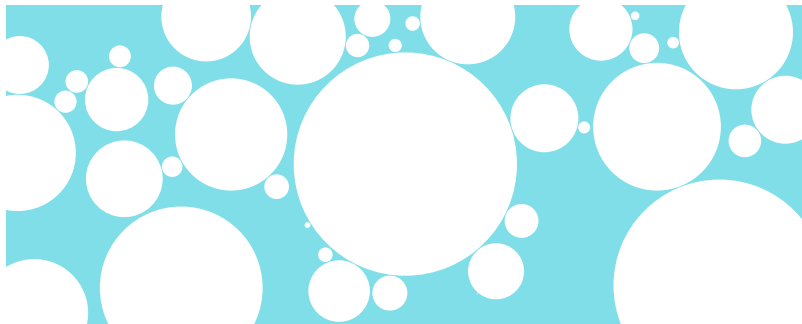
Virtual black holes and space-time foam (Summary)

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":



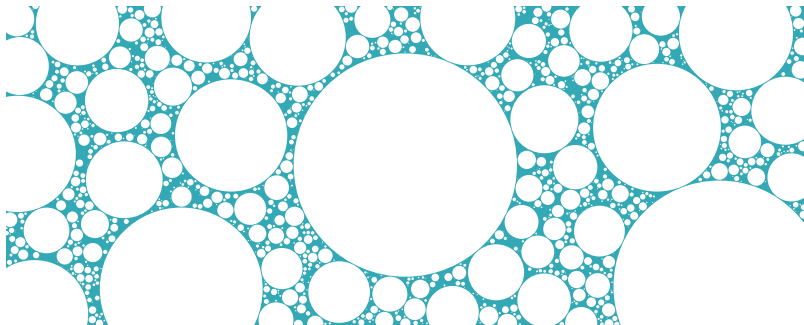
Virtual black holes and space-time foam (Summary)

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":



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":



Work is still in progress. More than happy to discuss these ideas with like-minded colleagues.

See: G. 't Hooft, arxiv:1612.08640 [gr-qc] + references there;
http://www.phys.uu.nl/~thoof/lectures/GtHBlackHole_2017.pdf

See also:

P. Betzios, N. Gaddam and O. Papadoulaki, The Black Hole S-Matrix from Quantum Mechanics, JHEP 1611, 131 (2016), arxiv:1607.07885.

----- ∞ -----