(ii) what is the *maximum* temperature of a globally stable configuration in a cavity with volume $V = 1\ \text{cm}^3$?

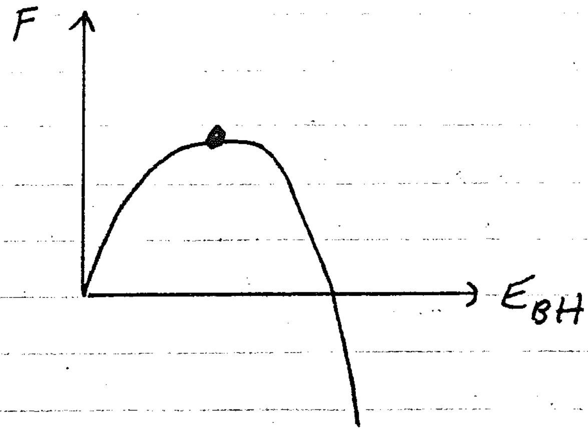
Remark: the maximum temperature is very high, because the total energy in the box must be huge before there is enough energy available to make a black hole that is cool enough to be in equilibrium with the remaining radiation.

Metastability

We have learned that sufficiently hot radiation in a box is unstable, because the entropy can be increased by spontaneous nucleation of a black hole. But the pure radiation is *locally* stable, and so there is a question of time scale. How long must one wait for nucleation of a black hole to be likely?

Let's assume that the box is so big that the total energy of the radiation in the box is much greater than the mass of a BH with temperature T; then nucleation of a BH with temperature T will lower the radiation temperature T by a negligible amount. If a black hole with temperature longer than the radiation temperature T occurs as a thermal fluctuation, it will quickly evaporate. But if a BH with temperature less than T appears, it will continue to accrete radiation (cooling the box) until the BH has absorbed a large portion of the total energy in the box.

We need to estimate, then, the free energy barrier for nucleation of a black hole with temperature T.



For a black hole in equilibrium with a thermal reservoir at inverse temperature β, the free energy F is given by

$$\beta F_{BH} = -S_{BH} + \beta E_{BH}$$
$$= -4\pi E_{BH}^2 + \beta E_{BH}$$

This has its maximum at $E_{BH} = \beta/8\pi$

— another way to see that a BH with temperature equal to that of the radiation must form.

The height of the free-energy barrier is

$$\beta F_{BH}\Big|_{max} = -\frac{1}{2}\frac{\beta^2}{8\pi} + \beta^2/8\pi = \beta^2/16\pi$$

So the nucleation rate is proportional to

$$(\text{Prob of nucleation per unit time and volume}) \propto \exp\left[\frac{-1}{16\pi T^2}\right]$$

in a large cavity containing radiation at Temperature T! (The leading contribution, for small $\hbar$, to the prefactor in front of the exponential has also been computed, by Gross, Perry, and Yaffe, Phys. Rev. D24(1982)330)

Of course, it is the Temperature in Planck units that appears in the exponential. For, say, $T = 300°K$, we have

$$\text{Rate} \sim \exp\left[-10^{-58}\right].$$

So nucleation of a black hole takes quite a while!

Now that we are considering a black hole in equilibrium with radiation in a cavity, let's return to the "easy" derivation of black hole radiance back on page 4.20, based on the periodicity of the Schwarzschild time coordinate (i.e., of <u>Euclidean time</u>).

First, consider the geometry of Schwarzschild metric, continued to $t = i\tau$. We have

$$d^2s_E = \left(1 - \frac{2M}{r}\right)d\tau^2 + \left(1 - \frac{2M}{r}\right)dr^2 + r^2 d\Omega^2$$

This geometry is easier to grasp if we introduce a coordinate

$$R = 4M\left(1 - \frac{2M}{r}\right)^{\frac{1}{2}}$$

$$\Rightarrow dR = \frac{4M^2}{r^2}\left(1 - \frac{2M}{r}\right)^{-\frac{1}{2}}$$

$$\Rightarrow d^2s_E = R^2\left(\frac{d\tau}{4M}\right)^2 + \left(\frac{r}{2M}\right)^4 dR^2 + r^2 d\Omega^2$$

This geometry is perfectly smooth at $R=0$ ($r=2M$), <u>if</u> we regard τ as an angular coordinate in the R-τ plane — that is $\frac{d\tau}{4M} = d\theta$, where θ has period 2π.

Hence, we should regard τ as a coordinate that is periodic mod $8\pi M$ on the Euclidean Schwarzschild geometry — which elucidates the periodicity of the

Schwarzschild - Kruskal transformation noted on page (4.20).

When we do the Euclidean rotation, we find that the (complexified) Schwarzschild coordinates exterior to the horizon cover all of a Euclidean manifold with the topology

$$\mathbb{R}^2 \times S^2$$
$$\uparrow \qquad\quad \uparrow$$
$$(R, \tau) \qquad (\theta, \phi)$$

except for a single point (that is, a single S^2) at $R=0$. We may complete the manifold by adding the missing point (S^2). Nothing in this complete Euclidean geometry corresponds to the regions of Kruskal P, F, II. In particular, there is no vestige of the singularity at $r=0$.

As described at the end of Chapter 3, we may construct the unique Euclidean Green function satisfying

$$\Box_E G_E(x, x') = -\frac{1}{\sqrt{g}} \delta^4(x-x')$$

on this Euclidean geometry that decays for $R \to \infty$. It is analytic on Euclidean Schwarzschild, and continued back to real Schwarzschild time t, is

just the *thermal* Green function in Region I, with (redshifted) $\beta = 8\pi M$ (for it is the boundary value of a Green function that is periodic in t with period $i\beta$, and is analytic in the strip).
— Hawking + Hartle
— Gibbons + Perry

It is natural to ask, how do we *interpret* this Green function; i.e., is it $\langle\Psi| T\, \phi(x)\, \phi(x') |\Psi\rangle$ in some state $|\Psi\rangle$? To specify this state consider the B.C. satisfied by the Green function.

We know that $G(x,x')$ is analytic in t in the (lower) strip

$$-i\beta < \operatorname{Im} t < 0 \qquad \beta = 8\pi M$$

Since

$$U = -\left(\frac{r}{2M}-1\right)^{\frac{1}{2}} e^{(r-t)/4M} = -e^{-(t-r_*)/4M}$$

$$V = \left(\frac{r}{2M}-1\right)^{\frac{1}{2}} e^{(r+t)/4M} = e^{(t+r_*)/4M}$$

(in Region I), this ensures that G is analytic

- In the lower half U plane, for $V=0$ (H^-)
- In the lower half V plane, for $U=0$ (H^+)

that is, imagine giving t an imaginary part $t \to t - i\epsilon$,

so that $$U = -|U| e^{i\sigma/4M}$$
$$V = |V| e^{-i\sigma/4M}$$

then, for $\sigma \in (0, 4\pi M)$, the U lhp is covered for $|V| = 0$, and the V lhp is covered for $|U| = 0$. (In fact, since we can vary σ in $(0, 8\pi M)$ we seem to have analyticity in all of U plane (for $|V| = 0$) except for cut on negative imaginary axis. Actually, this is true only if the point x' is kept in region I. Since $U, V \to -U, -V$ interchanges regions I and II, there is a singularity on the positive U axis if x' is in region II.)

We can interpret this analyticity property in terms of a notion of positive frequency in region I. Suppose we choose basis of solutions to KG eqn that are positive frequency wrt U on H^-. Boundary value data on H^- determines a solution to KG <u>if</u> we introduce, say, Dirichlet B.C. at some $r = r_0$. (This simulates putting the BH in a perfectly reflecting cavity.) These solutions will also be positive frequency wrt V on H^+.

To see this consider arbitrary positive frequency (wrt U) function on H^-.

This specifies a function analytic in lower U half plane. If we vary σ according to

$$U = -|U| e^{i\sigma/4M}$$
$$V = |V| e^{-i\sigma/4M}$$

by giving t an imaginary part, initial data is smooth on H^- for each fixed σ, and since Cauchy problem takes smooth data on initial value surface to smooth data on final surface, so we have a function analytic in lower half V plane on H^+.

We may now dispense with the cavity and specify a state

$$|H\rangle$$

such that $\quad a_i |H\rangle = 0$

if

u_i is pos. freq wrt U on H^-

or u_i is pos. freq wrt V on H^+

– This state H is called the "Hartle-Hawking vacuum."

$|H\rangle$ is one of three vacuum states that are frequently discussed on the BH background. The others are

$|B\rangle$ such that $a_i|B\rangle = 0$ if

u_i is pos freq wrt t

(or u_i pos freq wrt u on H^-
" " " v on $\mathscr{I}^-$)

This is the "Boulware vacuum", the state in which FIDOs detect no particles.

$|U\rangle$ such that $a_i|U\rangle = 0$ if

u_i is pos freq wrt U on H^-
" " " v on $\mathscr{I}^-$

This is the "Unruh vacuum" and corresponds to the realistic collapse situation – there are no incoming particles from $r=\infty$, and a thermal flux is coming from the (past) horizon.

The Green function argument for thermal emission from a black hole has the advantages

- It is simpler, in a sense, than Hawking's original argument.
- It is easier to generalize to interacting field theory, because analyticity properties of Green functions can be studied,

and are similar to free field theory. (At least in perturbation theory, where we can systematically study analyticity properties of Feynman diagrams — probably true nonperturbatively as well.)

The disadvantage of the argument is that it leaves unexplained why the state $|H\rangle$ is the correct one to consider, rather than $|B\rangle$ say. (And in fact, it is $|U\rangle$ that really corresponds to a BH evaporating in empty space; $|H\rangle$ describes BH in a bath.)

To better appreciate why $|H\rangle$ is to be preferred to $|B\rangle$, it is useful to consider how the (renormalized) energy-momentum tensor behaves near the horizon in these two states.

The Renormalized Stress-Tensor and Back Reaction

We have emphasized repeatedly that the notion of a "particle" suffers from ambiguities in curved spacetime. Whether particles are detected depends on the motion of the observer, and which observers are prefered cannot be determined locally, but only (sometimes) by some global criterion — e.g. if spacetime is asymptotically stationary.

It is useful, then, to characterize what a local observer can detect not in terms of (arbitrary) global notions of + and − frequency, but in terms of local observables $\mathcal{O}(x)$ that can be constructed from fields and are in principle measureable. Since $\langle \phi(x) \rangle = 0$ in a state with definite number of quanta, we might use e.g. $\phi(x)^2$, or other "composite operators" constructed from $\phi(x)$.

A particular local phenomenon that we might be interested in is back reaction -- how the quantum state of the fields feeds back and influences the spacetime geometry. Now, in QFT the Einstein equation

$$G_{\mu\nu} = -8\pi\, T_{\mu\nu}$$

becomes an operator equation; and it

is inevitable that, if the matter fields undergo quantum fluctuations, then the metric fluctuates also. A fully correct discussion of back reaction, then, requires us to go beyond quantized matter fields on a fixed background geometry, and to quantize gravity properly.

We will not attempt this; instead we continue to treat gravity classically, even though the source in the Einstein eqn is quantum-mechanical,

$$G_{\mu\nu} = -8\pi \langle T_{\mu\nu} \rangle .$$ "semiclassical gravity"

(this is not just expectation value of above operator equation, for $\langle G_{\mu\nu} \rangle \neq G_{\mu\nu}(\langle g \rangle)$, since $G_{\mu\nu}$ depends nonlinearly on g.) This semiclassical Einstein eqn may make sense if fluctuations in the operator $T_{\mu\nu}$ are small enough to neglect.

A local observer can in principle carry out measurements of the quantities

$\langle T_{\mu\nu} \rangle u^\mu u^\nu$ – energy density

$\langle T_{\mu\nu} \rangle u^\mu n^\nu$ – flux

where u^μ is observer's 4-velocity $u^2 = 1$ and $u \cdot n = 0$, $n^2 = -1$.

E.g., in the case of evaporating black hole, FIDO's near horizon should measure an energy flux into hole that is negative, since the BH is losing mass, and hence the horizon should be shrinking. This is an example of the sort of back reaction effect that we would like to understand.

A large part of the literature on QFT in curved spacetime concerns $\langle T_{\mu\nu} \rangle$, and some of this literature is rather technical. (see Chapter 6 of Birrell and Davies, and the book "Aspects of QFT in Curved Spacetime" by S. Fulling.) To discuss it, we must consider an aspect of QFT that we have mostly avoided up to now — renormalization. We will not be able to delve into this subject in great depth, but I will try to explain the basic concepts.

Composite Operators

The field $\phi(x)$ is an operator-valued distribution. And a product of distributions may be ill-defined. So a product of fields at the same spacetime point need not be a well-behaved field — does not yield an operator on Hilbert space when smeared with a smooth test function.

E.g. recall

$$G_+(x-y) = \langle 0|\,\phi(x)\,\phi(y)\,|0\rangle = \frac{-1}{4\pi^2 (x-y)^2}$$

in free massless field theory. It is singular as $x \to y$ (or, in fact, for x on the light cone of y).

A composite operator $\mathcal{O}$ constructed from ϕ and its derivatives must come equipped with a prescription that renders matrix elements $\langle\psi|\int \mathcal{O}(x)\,f(x)\,|\chi\rangle$ finite. — Prescription defines $(\mathcal{O})_{\text{ren}}$, renormalized operator

Example — can define $\phi(x)^2$ by the point splitting procedure

$$(\phi(x)^2)_{\text{ren}} = \lim_{x\to y}\Big[\phi(x)\,\phi(y) - \langle 0|\,\phi(x)\,\phi(y)\,|0\rangle\Big]$$

in general, in taking this limit, it may be required to average over orientation of $(x-y)^\mu$, with some suitable measure, so that $(\mathcal{O})_{\text{ren}}$ transforms as a tensor of the desired type

we subtract away a divergent c-no.

The renormalized operator that we have constructed is said to be <u>normal-ordered</u>. — we have in effect, by making the subtraction, moved all $a(k)$'s to the right and $a(k)^+$'s to the left. The resulting renormalized operator has the property

$$\langle 0 | (\phi(x)^2)_{ren} | 0 \rangle = 0$$

and has finite matrix elements between Fock space states.

More generally, we construct a renormalized composite operator $(\mathcal{O})_{ren}$ from a formal composite expression $\mathcal{O}$ by the procedure:

- <u>Regulate</u> the operator.
  We modify how $\mathcal{O}$ behaves "at short distances" so that $(\mathcal{O})_{reg}$ has finite matrix elements — e.g. point splitting in above example.
  Regularization alone is an unsuitable way to define operator because...
  (i) the matrix elements of $(\mathcal{O})_{reg}$ are very sensitive to the artificial cutoff, e.g., $\epsilon = x-y$ in point splitting case.
  (ii) Regulation spoils tensor properties of the operator — e.g. scalar becomes "biscalar" above.

- Subtract the divergent part.
  We remove a piece of the operator matrix element that becomes divergent when regulation is removed. The subtracted part may be a c-no., or may have to be regarded as the matrix element of a local operator of the same or lower dimension as $\mathcal{O}$ ("operator mixing" under renormalization) with the same quantum nos as $\mathcal{O}$ (perhaps including $\mathcal{O}$ itself).

- Remove the regulator.
  E.g., we take $x \to y$, in the point spitting method, after divergent part is subtracted away.

We may regulate and subtract in a myriad of different ways (different "schemes" for defining renormalized operator). All different schemes are related in a simple way — they differ by the finite part of the subtraction in the second step. Any well-formulated physics question must have an answer that does not depend on the scheme used to define the renormalized operator.

Stress Tensor in Curved Spacetime

For a massless free scalar field in flat spacetime, the "canonical" stress tensor is
$$T_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi - \frac{1}{2}\eta_{\mu\nu}\partial_\lambda \phi \partial^\lambda \phi$$

(page 2.6). It is a composite operator and requires renormalization.

It is natural to normal order it — we subtract so that
$$\langle 0 | (T_{\mu\nu})_{ren} | 0 \rangle = 0$$
— no stress-energy in the vacuum.
But subtractions that make $(T_{\mu\nu})_{ren}$ finite in flat spacetime donot suffice in a general spacetime.

The subtractions that must be made in a general spacetime are "c-no." subtractions that do not depend on the state in which $\langle T_{\mu\nu} \rangle$ is evaluated, but do depend on the background spacetime. Furthermore, whatever terms are subtracted must be conserved tensors. And they are local in spacetime, because they arise from very short-distance fluctuations that probe only local structure.

We can see what the form of the subtractions will be by dimensional analysis. If ϵ is the short distance scale inherent in the regulator, then, since $(T_{\mu\nu})$ has dimensions of $(\text{length})^{-4}$, the expansion of $(T_{\mu\nu})_{reg}$ for ϵ small has the form

$$\langle T_{\mu\nu}\rangle_{reg} \sim \underbrace{\frac{1}{\epsilon^4} + \frac{1}{\epsilon^2} + \ln\epsilon}_{=\text{infinite part}}$$

$$+ \; O(\epsilon^2) \; - \text{finite part.}$$

Since the infinite part, in each order in ϵ, is a conserved tensor, we have, by dimensional analysis:

$\frac{1}{\epsilon^4}$ term $\propto$ (conserved tensor of dimension 0)

$$= g_{\mu\nu}$$

$\frac{1}{\epsilon^2}$ term $\propto$ (conserved tensor of dimension 2)

$$= G_{\mu\nu} \equiv R_{\mu\nu} - \tfrac{1}{2} R g_{\mu\nu}$$

– two derivatives of the metric

$\ln\epsilon$ term $\propto$ (conserved tensor of dimension 4)

= two independent terms with four derivatives of the metric

These terms are (page 161 in Birrell and Davies)

$$H^{(1)}_{\mu\nu} = 2R_{;\mu\nu} - 2g_{\mu\nu}\Box R - \tfrac{1}{2}g_{\mu\nu}R^2 + 2RR_{\mu\nu}$$

$$H^{(2)}_{\mu\nu} = 2R_{\mu;\nu\alpha}^{\alpha} - \Box R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}\Box R + 2R_\mu{}^\alpha R_{\alpha\nu} - \tfrac{1}{2}g_{\mu\nu}R^{\alpha\beta}R_{\alpha\beta}$$

what we find then, is that the infinite subtractions in definition of $(T_{\mu\nu})_{ren}$ can be absorbed into the parameters of the Einstein equation

$$R_{\mu\nu} - \tfrac{1}{2}R g_{\mu\nu} + \Lambda_0 g_{\mu\nu} = -8\pi G_0 \langle (T_{\mu\nu})_{reg}\rangle$$

= "bare" values of Newton's constant and cosmological constant, before renormalization

$$= -8\pi G_0 \Big[\underbrace{\langle (T_{\mu\nu})_{ren}\rangle}_{\text{"finite" part}} + \alpha^{(0)}g_{\mu\nu} + \alpha^{(2)}G_{\mu\nu} \underbrace{+ \alpha_1^{(4)}H^{(1)}_{\mu\nu} + \alpha_2^{(4)}H^{(2)}_{\mu\nu}}_{\text{"infinite part"}} \Big]$$

If we put the "infinite" part on LHS of equation, this has the form

$$G_{\mu\nu} + \Lambda_{ren}\, g_{\mu\nu} + \underbrace{(4\text{-derivative terms})}_{\text{new terms induced by renorm.}} = -8\pi\, G_{ren} \langle (T_{\mu\nu})_{ren}\rangle$$

"renormalized" parameters that we actually measure.

This renormalization can be described in an alternative (and perhaps more illuminating) language. Suppose we define a QFT by introducing an explicity cutoff mass M. This means that the quantum fluctuations of the fields with wavelength $< M^{-1}$ are not included (the theory is regulated) E.g., we might imagine $M \sim M_{Planck}$; it is, in any event, large compared to the energy scale relevant to observations that we want to discuss. The theory has an action – ("bare action")

$$S = S_{grav,0} + S_{matter,0}$$

where $$S_{grav,0} = \int d^4x \sqrt{g} \left[\frac{1}{16\pi G_0}(R - 2\Lambda_0)\right]$$

Now we let the cutoff M "float" down to a new scale $\mu \ll M$. In the process, to keep low energy physics invariant, we incorporate the effects of short-distance quantum fluctuations into renormalization of the parameters of the theory. So the same low-energy physics can be described by a renormalized action

$$S(\mu) = S_{grav}(\mu) + S_{matter}(\mu),$$

where

$$S_{grav}(\mu) = \int d^4x \sqrt{g} \left[\frac{1}{16\pi G_{ren}(\mu)} \left(R - \Lambda_{ren}(\mu) \right) + (\text{4-derivative terms}) + \cdots \right].$$

On dimensional grounds, we expect the renormalized and bare parameters to differ by amounts of order

$$\delta\left(\frac{\Lambda}{8\pi G}\right) \sim M^4$$

$$\delta\left(\frac{1}{16\pi G}\right) \sim M^2$$

$$\delta(\text{4-deriv coupling}) \sim \ln(M/\mu)$$

The cosmological constant, in particular, receives an enormous renormalization from the short distance fluctuations (contribution to vacuum energy from zero-point fluctuations of the field). Naively, then, bare Λ and renormalization of Λ must very nearly cancel in order that $\Lambda_{ren}/8\pi G_{ren} \ll M^4$.

Failing this, we would expect

$$\left(\Lambda/8\pi G\right)_{ren} \sim M_{Pl}^4,$$

while in fact we know from expansion of universe --

$$\left(-\Lambda/8\pi G\right)_{\text{ren}} \lesssim 10^{-29} g/cm^3 \sim 10^{-122} M_{Pl}^4$$

— a spectacular disagreement between theory and experiment — "The Cosmological Constant problem". Still a big mystery.

The planck mass, i.e. $\frac{1}{16\pi G}$,

also gets renormalized, but the effects of fluctuations with wavelengths $\ll M_P^{-1}$ are unimportant.

The 4-derivative terms get induced by renormalization if they are not present to begin with. The terms $H_{\mu\nu}^{(1,2)}$ written above arise from

$$S^{(1)} = \int d^4x \sqrt{g}\, R^2$$

$$S^{(2)} = \int d^4x \sqrt{g}\, R_{\alpha\beta} R^{\alpha\beta}$$

when we vary with respect to $g_{\mu\nu}$ to derive an eqn of motion. We need not consider

$$\int d^4x \sqrt{g}\, R_{\alpha\beta\gamma\delta} R^{\alpha\beta\gamma\delta}, \text{ because}$$

$$\int d^4x \sqrt{g}\left(R_{\alpha\beta\gamma\delta} R^{\alpha\beta\gamma\delta} + R^2 - 4 R_{\alpha\beta} R^{\alpha\beta}\right)$$

is integral of a total derivative (topological invariant) and does not contribute to field equation.

The induced 4-derivative terms are actually logarithmically sensitive to the "floating cutoff" μ; this means that these renormalizations are not completely dominated by short-distance fluctuations, as renormalizations of Λ and G are. The logarithmic dependence arises because all length scales contribute democratically to the renormalization, so that the leading contribution to the renormalization comes from fluctuations at length scales $\gg M^{-1}$. (The renormalized 4-deriv. couplings "run" with μ; more about this below, in connection with the conformal anomaly.)

~~As noted above, different "renormalization schemes" for $(T_{\mu\nu})_{ren}$ correspond to different choices for the finite parts of the subtractions. We now understand~~

Note that the effects of the 4-derivative terms in the Einstein equation are typically highly suppressed, if curvature is small in planck units. If the coefficients of these terms are of order one, the corresponding corrections are of order $(L_{planck}/L)^2$, if L is the length scale that characterizes the curvature.

As noted above, different "renormalization schemes" for $\langle T_{\mu\nu}\rangle_{ren}$ correspond to different choices for the *finite* parts of the subtractions. We now understand that different schemes actually differ in an essentially trivial way — by moving a term from $T_{\mu\nu}$ on the RHS of the Einstein eqn over to the LHS. This reshuffling has no effect on any physical predictions.

Stress - Tensor in Flat Spacetime

Even when there is no curvature, there is an ambiguity in the energy momentum tensor of our massless free scalar field.

$$(T_{\mu\nu})_{canonical} = \partial_\mu\phi\partial_\nu\phi - \eta_{\mu\nu}\left(\tfrac{1}{2}\partial_\lambda\phi\partial^\lambda\phi\right)$$

is the object satisfying $\partial_\mu T^\mu{}_\nu = 0$ that is derived from translation invariance by the Noether procedure. The conserved quantities are

$$P_\mu = \int T_{0\mu}\, d^3x.$$

But we may change $T_{\mu\nu}$ by a total derivative term without changing the P_μ's:

$$(T_{\mu\nu})_{new} = (T_{\mu\nu})_{canonical} - \xi(\partial_\mu \partial_\nu - \eta_{\mu\nu}\partial^2)\phi^2$$

satisfies $\partial_\mu T^{\mu\nu} = 0$, for any value of ξ, and P_μ does not depend on ξ.

Note that the trace is

$$(T^\mu{}_\mu)_{new} = -\partial_\lambda \phi \partial^\lambda \phi + 3\xi \partial^2 \phi^2$$
$$= (-1+6\xi)\partial_\lambda \phi \partial^\lambda \phi$$

(using eqn of motion $\partial^2\phi = 0$.)

So that $(T^\mu{}_\mu)_{new} = 0$ for $\xi = \frac{1}{2}$.

This choice defines the "new improved" stress Tensor.

What is the significance of $T^\mu{}_\mu = 0$? We can understand the meaning of the parameter ξ better if we imagine "turning on" the curvature. Keeping terms with at most two derivatives, the action of a massless scalar field is

$$S_{matter} = \int d^4x \sqrt{g}\left(\tfrac{1}{2}\partial_\mu \phi \partial_\nu \phi g^{\mu\nu} - \tfrac{1}{2}\xi R \phi^2\right)$$

where ξ is a free parameter. KG eqn then becomes

$$(\Box + \xi R)\phi = 0.$$

We can extract flat space $T_{\mu\nu}$ by varying S_{matter} wrt $g_{\mu\nu}$, and then setting $g_{\mu\nu} = \eta_{\mu\nu}$,

$$\delta S_{matter} = -\frac{1}{2}\int d^4x \sqrt{g}\,\left(T_{\mu\nu}\,\delta g^{\mu\nu}\right).$$

Then

$$\left(T_{\mu\nu}\right)_{g=\eta} = \partial_\mu\phi\,\partial_\nu\phi - \tfrac{1}{2}\eta_{\mu\nu}\,\partial_\lambda\phi\,\partial^\lambda\phi - \xi\left(\partial_\mu\partial_\nu - \eta_{\mu\nu}\partial^2\right)\phi^2$$

The free parameter arises in $T_{\mu\nu}$ because there really is freedom in how we couple the scalar field to a curved background.

The trace of $T_{\mu\nu}$ determines how S_{matter} transforms under a conformal rescaling of the metric:

$$g(x) \quad g_{\mu\nu}(x) \to \Omega^2(x)\, g_{\mu\nu}(x)$$

$$= (1+2\delta\Omega)\, g_{\mu\nu},$$

for an infinitesimal conformal transformation $\Omega = 1+\delta\Omega$.

$$\Rightarrow \delta S_{matter} = -\int d^4x\ \delta\Omega(x)\, T^\mu_{\ \mu}(x).$$

The coupling to matter is conformally invariant if $T^\mu_{\ \mu} = 0$.

From perspective of flat space QFT, the significance of $T^\mu{}_\mu = 0$ is that

$$D_\mu = x^\nu \, (T_{\mu\nu})_{new}$$

is then conserved, and $\int d^3x \, D_0$ is the operator that generates a global scale transformation.

Conformal Anomaly

Even if we choose $\xi = \frac{1}{6}$ in coupling massless scalar to curved background, so that coupling is conformally invariant at the classical level, the conformal invariance is broken by quantum effects. This is called the "conformal anomaly".

The origin of the conformal anomaly is easy to understand; it arises from the logarithmic running of couplings induced by renormalization. In other words, although there is no mass scale in the classical theory, we have no choice but to introduce an implicit mass scale when we remove divergences $\propto \ln \epsilon$.

Formally, when we make a change of length scale

$$g_{\mu\nu} \to (1+2\delta\Omega)\, g_{\mu\nu},$$

physics is not really left invariant unless we simultaneously adjust the length scale at which renormalized couplings are defined,

$$\mu \to (1+\delta\Omega)\mu.$$

So –

$$0 = \int \sqrt{g}\, \delta\Omega \left(-\left(T^\mu_\mu\right)_{ren} + \mu \frac{\partial}{\partial\mu} \mathcal{L}_{ren} \right)$$

(where $S_{ren} = \int \sqrt{g}\, \mathcal{L}_{ren}$)

or

$$\left(T^\mu_\mu\right)_{ren} = \mu \frac{\partial}{\partial\mu} \mathcal{L}_{ren}$$

– this is the conformal anomaly.

In other words, if logarithmic renorm. is

$$S = \int d^4x \sqrt{g} \left[\cdots + (\ln\mu)\, \alpha^{(4)} \mathcal{O}^{(4)} + \cdots \right]$$

term arising from log divergence

Then

$$\left(T^\mu_\mu\right)_{ren} = \alpha^{(4)} \mathcal{O}^{(4)}$$

term with 4 derivatives of metric.

Vacuum Polarization in Rindler Spacetime

As usual, this is a good warmup for the BH case. We wish to compare how back reaction behaves near horizon in the Rindler vacuum and the Minkowski vacuum

Consider

$$\left[\langle 0, R_{in} | T_{\mu\nu} | 0, R_{in}\rangle - \langle 0, M_{in} | T_{\mu\nu} | 0, M_{in}\rangle\right] u^\mu u^\nu$$

(where u^μ = 4-velocity of Rindler observer). This quantity is unaffected by renormalization, since the subtractions cancel between the two terms. It can be computed as a sum over Rindler modes (done by Candelas and Deutsch) but we know the result without doing a new computation, because we know that a Rindler observer sees no quanta in $|0, R_{in}\rangle$, and sees a thermal spectrum in $|0, M_{in}\rangle$. Hence ---

$$= -\frac{1}{\xi^4}\int \frac{d\omega\,\omega^2}{2\pi^2}\frac{\omega}{e^{2\pi\omega}-1}$$

— just the energy density of a thermal bath at temperature $T = 1/2\pi\xi$, but with a minus sign.

Now, in this case, it is obvious how to renormalize $T_{\mu\nu}$ — we make subtractions so that there is no back reaction in Minkowski vacuum on flat space; that is,

$$\langle 0, Min | (T_{\mu\nu})_{ren} | 0, Min \rangle = 0.$$

Thus

$$(u^\mu u^\nu)_{Rin} \langle 0, Rin | (T_{\mu\nu})_{ren} | 0, Rin \rangle = -\frac{\pi^2}{30}\left(\frac{1}{2\pi\xi}\right)^4$$

(thermal)

The Rindler observer in the Minkowski vacuum, close to the horizon, sees a very hot thermal bath, with very large energy density. Yet he finds that this bath exerts no back reaction on the spacetime. He concludes that, in addition to the positive contribution to the energy density due to the thermal radiation, there is also a compensating "vacuum polarization" correction that is large and negative.

The Rindler observer in the Rindler vacuum sees the vacuum polarization, uncompensated by thermal radiation. He finds a large negative energy density, which exerts a strong back reaction on the spacetime. This divergent back reaction at the horizon is perceived by freely falling observers as well as Rindler observers; that is

$$\langle 0, Rin | (T_{\mu\nu})_{ren} | 0, Rin \rangle (u^\mu u^\nu)_{FFO}$$

also blows up at horizon.

(We have $\quad u^\mu_{Rin} = \left(\frac{d\eta}{d\tau}, \frac{d\xi}{d\tau}, \vec{0}\right) = \left(\frac{1}{\xi}, 0, \vec{0}\right)$

$$u^\mu_{FFO} = \left(\frac{1}{\xi}\cosh\eta, \sinh\eta, \vec{0}\right)$$

if FFO is a $\bar{z}$=constant – see page 3.20D.)

The divergent vacuum polarization may be interpretted as follows: Because of the ∞ red shift at the horizon, the Rindler modes freeze there, that is, in the Rindler vacuum, quantum fluctuations are strongly suppressed near the horizon, relative to the fluctuations that occur in the Minkowski vacuum. Hence the negative vacuum energy. The quantum fields desporately want to fluctuate at the horizon, but they are prevented from doing so, and so exert a strong force on the horizon.

Vacuum Polarization in Black Hole Spacetime

Since Boulware state $|B\rangle$ is analog of $|0, R\,in\rangle$ and Hawking-Hartle state $|H\rangle$ is analog of $|0, M\,in\rangle$, it is natural to consider

$$\left[\langle B| T_{\mu\nu} |B\rangle - \langle H| T_{\mu\nu} |H\rangle\right](u^\mu u^\nu)_{FIDO}$$

— again, a quantity unaffected by renormalization of $T_{\mu\nu}$. This can be expressed as a mode sum that cannot be evaluated analytically, but the asymptotic form close to the horizon can be extracted (Candelas).

Again, we can easily interpret the result. The FIDO detects no quanta in state $|B\rangle$, and a thermal bath (close to the horizon) in the state $|H\rangle$ so

$$= \frac{-\pi^2}{30}\left(\frac{1}{8\pi M}\right)^4\left(1-\frac{2M}{r}\right)^{-2}$$

— the thermal energy density at the local temperature $T = (8\pi M)^{-1}\left(1-\frac{2M}{r}\right)^{-\frac{1}{2}}$, but with a minus sign.

This difference diverges at the horizon, and we must determine whether the divergence occurs in the state $|B\rangle$, the state $|H\rangle$ or both. The Rindler analogy suggests that the Hartle-Hawking state $|H\rangle$ behaves smoothly at the horizon, while the vacuum polarization in the Boulware state $|B\rangle$ is divergent and negative.

To resolve this, we must renormalize $T_{\mu\nu}$ by making the appropriate subtractions. (To the relevant order of approximation, we need not worry about ambiguities in the finite part of the subtractions, as $R_{\mu\nu}$, $H^{(1)}_{\mu\nu}$, $H^{(2)}_{\mu\nu}$ vanish if the background satisfies the vacuum Einstein equation $R_{\mu\nu}=0$.) One finds that, indeed

$$\langle H|(T_{\mu\nu})_{ren}|H\rangle = \text{finite at horizon}$$

while leading behavior in Boulware state is

$$\langle B|(T_{\mu\nu})_{ren}|B\rangle(u^\mu u^\nu)_{FIDO}$$
$$= -\frac{\pi^2}{30}\left(\frac{1}{8\pi M}\right)^4\left(1-\frac{2M}{r}\right)^{-2} \quad \text{as } r\to 2M$$

The FIDO in the Hartle-Hawking state sees a thermal bath with energy density diverging at the horizon, yet the back reaction is finite, because the divergence is cancelled by a large negative vacuum polarization correction to the energy density. (This is consistent with the perspective of an EFO, who sees no thermal radiation and a smoothly behaving vacuum polarization at the horizon.)

The FIDO in the Boulware state sees the large negative vacuum polarization contribution to the energy density near the horizon, uncompensated by any thermal radiation. The strong gravitational field at the horizon suppresses the vacuum fluctuations of the quantum fields near the horizon; hence the large negative energy density. The fields therefore exert a strong back-reaction force on the geometry. FFO's also perceive a divergent vacuum polarization at the horizon.

In the Hartle-Hawking state, there is no net flux of energy across the horizon; the black hole is accreting radiation at the same rate that it is emitting, and so remains in equilibrium with the thermal bath. In the Unruh state, however, the black hole is losing mass, and the energy of the escaping radiation is increasing. Since energy is conserved in the static spacetime exterior to the horizon, energy must be entering this region through the horizon. Except that it is really more appropriate to say that a negative energy flux is escaping this region through the horizon.

To understand how this works in

more detail, we must note first of all that the globally conserved quantity in the BH background is the redshifted energy, or the "energy at infinity." For a test particle moving on a geodesic this is

$$E_\infty = p_0 = g_{00} p^0 .$$

E.g. a particle of mass m dropped from rest at $r=\infty$ will be moving ultrarelativistically as measured by a FIDO close to the horizon, but will add only mass m to the mass of the BH, after it has descended below the stretched horizon. Similarly, a field quantum with red-shifted frequency ω_∞ adds mass ω_∞ to the BH when it is absorbed.

For each mode of the field, the locally measured energy at horizon is finite, and hence the red shifted energy is zero, if the mode is thermally occupied.

So the "energy at ∞" carried by a mode of redshifted frequency ω_∞ is

$$(E_\infty)_{\text{mode}} = (n - n_{\text{thermal}})\, \omega_\infty$$

this is vacuum polarization correction, near the horizon.

Now, in the Unruh state, we know that modes upcoming from the past horizon are precisely thermally occupied; therefore, these modes carry no E_∞ away:

$$n_\uparrow = n_{thermal} \Rightarrow (E_\infty)_\uparrow = 0$$

Modes that are escaping down through the future horizon have occupation number that comes from two sources: upcoming quanta that have been reflected back by the potential barrier, and quanta incoming from $\mathscr{I}^-$ that have successfully surmounted the potential barrier. So –

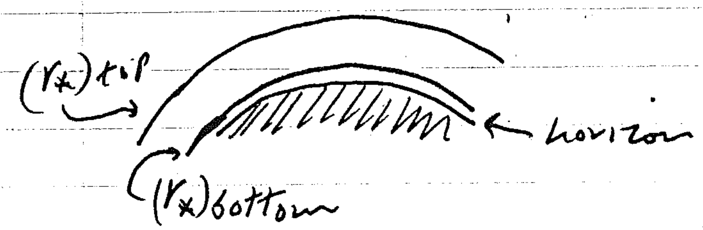
$$n_\downarrow = \underbrace{|r|^2}_{\text{reflection probability}} n_\uparrow + \underbrace{(1-|r|^2)}_{\text{transmission probability}} n_{in}$$

For these modes, too, vacuum polarization contributes an energy equal to that of a thermal bath, with a negative sign, and so modes crossing the future horizon carry energy at ∞

$$(E_\infty)_\downarrow = (n_\downarrow - n_{thermal})\,\omega_\infty$$

$$= (1-|r|^2)(n_{in} - n_{thermal})$$

(using $n_\uparrow = n_{thermal}$) To find the rate at which energy escapes region I through the future horizon, we now sum over ω_∞, weighted by a rate at which modes with frequency ω_∞ are escaping through the horizon.



To do the mode counting, imagine putting two concentric spherical shells around the BH, both very close to the horizon. Consider radiation modes in the cavity between the two shells. The radial behavior of the modes is

$$\sim e^{-i\omega(t-r_*)}$$

and so, as the inner shell is pushed closer to the horizon, we can replace sum over radial modes by an integral $\left((r_*)_{bottom} \to -\infty\right)$:

$$\sum_{modes} \to \frac{L}{2\pi} \sum_{\ell m} \int d\omega$$

where $L = (r_*)_{top} - (r_*)_{bottom}$

Since a wave-packet propagates from top to the bottom of the cavity in Schwarzschild time $\Delta t = L$, the rate per unit Schwarzschild time at which energy escapes is

$$\frac{\Delta M}{\Delta t} = \sum_{\ell}(2\ell+1)\int\frac{d\omega}{2\pi}\left(1-|r_{\omega\ell}|^2\right)\left(n^{in}_{\omega\ell}-n^{thermal}_{\omega\ell}\right)$$

In the Hartle-Hawking state, we have $n^{in} = n^{thermal}$, and so no mass loss occurs. But in the Unruh state, $n^{in} = 0$, and this expression coincides exactly with the black hole luminosity on page 4.50.

Thus, with vacuum polarization properly accounted for, we find a flux of negative energy through the future horizon H^+ that exactly compensates for the energy radiated to $\mathscr{I}^+$, in the Unruh state.

Near the horizon, nearly all propagating outgoing modes get reflected back; this endows the BH in the Unruh state with a "thermal atmosphere" as described on page (4.62). Thus, back reaction is finite at H^+ in the Unruh state, as in the Hartle-Hawking state. However, on the past horizon H^- in the Unruh state, we have $n_\uparrow = n_{thermal}$ and $n_\downarrow = 0$. So the Unruh state has divergent negative energy density on H^-. This is similar to Boulware state on H^-, except that only half of the modes, rather than all, are empty.

(This description of the energy flux at H^+ in the Unruh state comes from "The Membrane Paradigm", and Frolov + Thorne, Phys. Rev. D39 (1989) 2125.)

The Final State of the Evaporating Black Hole

As black hole evaporates and loses mass, it will eventually reduce its mass to $M \sim M_{planck}$. At this point, its size is of order L_{planck}, its "temperature" (as computed semiclassically) is of order M_{planck}. Quantum fluctuations in geometry now become important, and our semiclassical theory breaks down. What happens next?

We don't know. Two reasonably plausible possibilities are --

- the black hole disappears completely, leaving no trace
- A stable remnant with $M \sim M_{planck}$ remains, an exotic, stable "elementary" particle

If the black hole disappears completely, what implications does this have?

Breakdown of Global Conservation Laws

For example, baryon number conservation (or B-L, which is conserved in the standard model) would not be respected by the formation and evaporation of a black hole, even if satisfied by all other processes. Since a black hole has no baryonic hair, black holes made from a collapsing matter star evaporate in the same way as black holes made from a collapsing antimatter star; in the semiclassical theory, both produce baryons and antibaryons in equal abundances.

Even if baryonic and antibaryonic black holes behave differently when $M \sim M_{Planck}$ and semiclassical theory breaks down, it is too late to rescue the law of baryon conservation, for a solar mass black hole made from 10^{57} baryons that has already radiated nearly all of its mass.

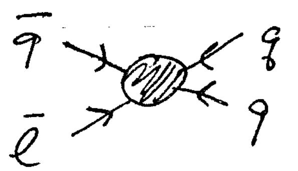
What if black hole evaporation is not complete, and a Planck size remnant remains? Can baryon conservation then hold? We might regard the remnant as a highly exotic tightly bound nucleus with $B \sim 10^{57}$, say. But for this to make sense, there should be many species of stable BH, each with a different value of B, and all, presumably, with $M \sim M_{Planck}$.

That does not sound plausible.

If black hole physics can change baryon number, there is no reason to expect that <u>virtual</u> black holes (quantum gravity) cannot induce the decay of the proton. A process

quark + quark $\rightarrow$ antiquark + antilepton

(the simplest B-changing process allowed by the gauge symmetries of the standard model) would then have an <u>amplitude</u>, on dimensional grounds.



$$\sim \frac{1}{M_{planck}^2}$$

Thus, in order of magnitude

$$\text{Proton Decay Rate} \sim \frac{M_{proton}^5}{M_{planck}^4} \sim (10^{45}\,yr)^{-1}.$$

To see one decay a year, one would need to watch about $(10)^6$ $(Km)^3$ of water.

So, our failure to observe proton decay should not be construed as evidence that black holes do not violate conservation of baryon number.

The Loss of Quantum Coherence

In describing the radiation emitted by an evaporating black hole, as seen by a distant observer, we used a density matrix $\rho_{\mathcal{I}^+}$. In a sense, this was a matter of convenience, since the observer had no access to information about the quantum state of radiation that had reached the future horizon H^+.

But if the black hole eventually evaporates completely, and the horizon disappears, this is no longer just a matter of convenience. It seems that a pure quantum state at $\mathcal{I}^-$ can evolve into a mixed state at $\mathcal{I}^+$. There is an intrinsic, unavoidable loss of phase information concerning the initial quantum state, and a corresponding intrinsic generation of entropy. The phase information cannot be retained even in principle.

If this is so, the foundations of quantum mechanics must be modified. The fundamental dynamical object becomes, not a wave function, but a density matrix.

Can this conclusion be avoided, so that quantum mechanics as we know it may survive?

Perhaps the radiation emitted in the late stages of BH evaporation has quantum-mechanical correlations with the radiation emitted in the early stages. But there is no sign of such correlations in the semiclassical analysis of BH radiance, which yields an exactly thermal density matrix for the outgoing radiation. In order for such correlations to be established, it seems that the black hole would need to have some kind of nonclassical "quantum hair", so that the radiation emitted early would leave an imprint on the BH that could influence the radiation emitted late.

If evaporation halts, leaving a stable remnant, then perhaps the remnant could attain a quantum state that is highly correlated with the radiation that has been emitted. But if there is no intrinsic generation of entropy, then it seems that the Planck-size remnant must be capable of carrying an enormous amount of information. It would need to have access to a number of "internal states" of order $\exp(\frac{1}{4}A_{\text{initial}})$, where A_{initial} is the area of the horizon when the BH first forms. This is hard to imagine

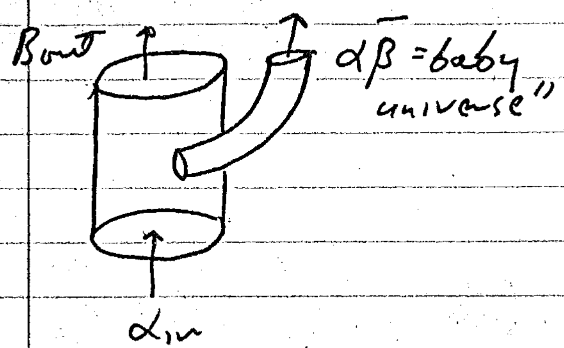
If the "no-hair" theorem fails to apply quantum-mechanically, then accretion of a particle might change the "quantum state" of the black hole, so that the subsequent emission would be correlated with what was absorbed. In this way, loss of quantum coherence, and the intrinsic increase in entropy, might in principle be avoided. Black hole radiation, then, when analyzed with greater care, would not be precisely thermal, but would be capable of carrying complex correlations, and hence much information.

(That black holes destroy quantum coherence has been emphasized by Hawking; that this conclusion might be too hasty has been stressed by 't Hooft.)

Topology Change in Quantum Gravity

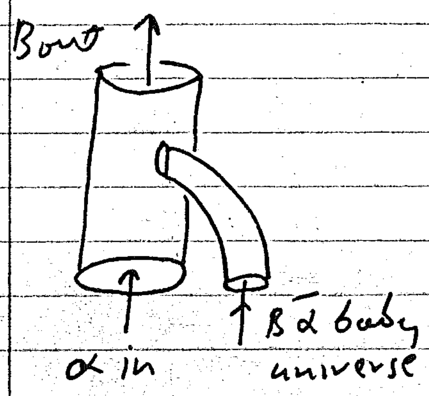
Hawking claims that the formation and subsequent complete evaporation of a black hole has another important implication, that fluctuations in the topology of spacetime must be allowed in quantum gravity.

I am not sure that I understand the connection, but if I interpret the claim correctly, the idea is that, if an object with quantum numbers denoted by α collapses to form a black hole, and then evaporates to produce an object with quantum numbers β. Then the

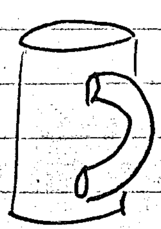


change in quantum numbers is carried away by a "baby universe" with quantum numbers $\alpha\bar{\beta}$.

(E.g. it carries away the change in the baryon number.) This baby universe is a closed 3-manifold, completely disconnected from our universe, and completely inaccessible to measurement by us. Hence the intrinsic loss of information. Since a closed universe carries vanishing energy, angular momentum, and charge, this process is consistent with the notion that black holes have at least a few varieties of hair.



We can just as well imagine that a BH event $\alpha_{in} \to \beta_{out}$ is accompanied by the absorption by our universe of a baby universe with quantum numbers $\beta\bar{\alpha}$.



Putting together $\alpha_{in} \to \beta_{out}$ with the time reversed process $\beta_{in} \to \alpha_{out}$, we describe a history of the universe in which a handle, or "wormhole" is attached to spacetime. This is the aforementioned fluctuation in topology.

In this scenario, any black hole event is in principle reversible. A black hole is allowed to evaporate in all possible ways consistent with its mass, angular momentum, and charge. The radiation looks thermal because thermal radiation is overwhelmingly the most probable state, if all microscopic states are allowed.

Will the question of the final state of the evaporating black hole, and the issue of loss of quantum coherence, ever be resolved? This is a genuine quantum gravity question, and we may find the answer once we have a sufficient grasp of quantum gravity (string theory?).

In the meantime, I expect that deeper insights into these questions can be derived from closer scrutiny of the semiclassical theory presented in this course.